



RESEARCH ARTICLE

Methodological evaluation of individual cognitive prediction based on the brain white matter structural connectome

Guozheng Feng^{1,2,3}  | Yiwen Wang^{1,2,3} | Weijie Huang^{1,2,3} | Haojie Chen^{1,2,3} | Zhengjia Dai⁴ | Guolin Ma⁵ | Xin Li^{1,2} | Zhanjun Zhang^{1,2}  | Ni Shu^{1,2,3}

¹State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China

²BABRI Centre, Beijing Normal University, Beijing, China

³Beijing Key Laboratory of Brain Imaging and Connectomics, Beijing Normal University, Beijing, China

⁴Department of Psychology, Sun Yat-sen University, Guangzhou, China

⁵Department of Radiology, China-Japan Friendship Hospital, Beijing, China

Correspondence

Ni Shu, State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, China.
Email: nshu@bnu.edu.cn

Funding information

Fundamental Research Funds for the Central Universities, Grant/Award Number: 2017XTCX04; International Cooperation and Exchange of the National Natural Science Foundation of China, Grant/Award Number: 81820108034; National Key Research and Development Project of China, Grant/Award Number: 2018YFC1315200; National Natural Science Foundation of China, Grant/Award Numbers: 81671761, 81871425; State Key Laboratory of Cognitive Neuroscience and Learning, Grant/Award Number: CNLYB2001

Abstract

An emerging trend is to use regression-based machine learning approaches to predict cognitive functions at the individual level from neuroimaging data. However, individual prediction models are inherently influenced by the vast options for network construction and model selection in machine learning pipelines. In particular, the brain white matter (WM) structural connectome lacks a systematic evaluation of the effects of different options in the pipeline on predictive performance. Here, we focused on the methodological evaluation of brain structural connectome-based predictions. For network construction, we considered two parcellation schemes for defining nodes and seven strategies for defining edges. For the regression algorithms, we used eight regression models. Four cognitive domains and brain age were targeted as predictive tasks based on two independent datasets (Beijing Aging Brain Rejuvenation Initiative [BABRI]: 633 healthy older adults; Human Connectome Projects in Aging [HCP-A]: 560 healthy older adults). Based on the results, the WM structural connectome provided a satisfying predictive ability for individual age and cognitive functions, especially for executive function and attention. Second, different parcellation schemes induce a significant difference in predictive performance. Third, prediction results from different data sets showed that dMRI with distinct acquisition parameters may plausibly result in a preference for proper fiber reconstruction algorithms and different weighting options. Finally, deep learning and Elastic-Net models are more accurate and robust in connectome-based predictions. Together, significant effects of different options in WM network construction and regression algorithms on the predictive performances are identified in this study, which may provide important references and guidelines to select suitable options for future studies in this field.

KEYWORDS

brain structural connectome, cognitive prediction, individual difference, machine learning, methodological evaluation, white matter

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *Human Brain Mapping* published by Wiley Periodicals LLC.

1 | INTRODUCTION

Combined with different neuroimaging features, regression algorithms have been increasingly applied to predict individual cognitive functions (Gabrieli et al., 2015; Sui et al., 2020). Specifically, regression-based multivariate models have provided a powerful and widely used approach to predict human behavior from neuroimaging features. Models spanning multiple brain regions (features) have been developed to reveal the neurobiology of cognition and behavior (Rosenberg et al., 2018). Compared with traditional univariate brain-behavior mapping, the multivariate model shows stronger statistical power and spatial mapping capability (Woo et al., 2017). In addition, predictive models usually employ a nested cross-validation (CV) strategy to avoid the problem of overfitting and to achieve generalization across different datasets (Gabrieli et al., 2015; Rosenberg et al., 2018; Sui et al., 2020; Woo et al., 2017).

Using different brain imaging modalities, different types of neuroimaging features can be extracted. Brain connectome-based features and predictive models exhibit great potential for individual fingerprint identification, and they have been gradually applied to the individual prediction of cognitive functions (Rosenberg et al., 2016; Sui et al., 2020). Both functional and structural brain networks can be constructed using different brain magnetic resonance imaging (MRI) techniques. Then, graph theory analysis is performed to quantify the integration and segregation of brain networks. To date, the functional connectome has been widely applied to predict fluid intelligence (Jiang et al., 2020), attention (Gao et al., 2020; Rosenberg et al., 2016), processing speed (Fountain-Zaragoza et al., 2019), working memory (Jangray et al., 2018), creativity (Beaty et al., 2018), and visuospatial functions (Chen et al., 2019); however, the potential utility of the structural connectome for establishing individual predictions remains largely unknown. Compared with the functional connectome, the structural connectome has been less studied as a neuroimaging predictor of individual cognition and behavior.

Using the diffusion MRI (dMRI) tractography technique, the macroscale white matter (WM) structural connectome has been constructed at an individual level in vivo (Bullmore & Sporns, 2009; Jbabdi et al., 2015; Rubinov & Sporns, 2010; Sporns et al., 2005). The topological organization of the brain structural connectome is sensitive to development and aging and exhibits substantial individual differences in both normal and disease populations. Previous brain network studies using dMRI have shown a significant correlation between connectomics features (edge or topological properties) and processing speed or executive function in normal elderly individuals and patients with MCI (Fornito et al., 2015; Madole et al., 2020; Palop & Mucke, 2016), suggesting the potential of brain connectome-based markers to predict individual cognitive performance. Notably, our previous study trained a linear regression model based on the WM structural connectome, which predicts executive and attention functions in normal elderly individuals (Li et al., 2020). Recently, a study adopted the WM structural network to construct an enriched functional network, which improved the network consistency and predictive power of cognitive function (Kim et al., 2021). Moreover, a

cutting-edge study found that ultrahigh-resolution WM connectomes yielded accurate predictive performance both on a range of behavioral measures and individualized fingerprints (Mansour et al., 2021), encouraging further investigation of brain-behavior studies with structural connectomes.

The brain connectome-based prediction pipeline consists of two core parts: brain network construction and machine learning-based prediction. However, many methodological choices require further evaluation in this field, especially for structural connectome-based predictions. First, how do different methods of WM network construction affect the cognitive prediction performance? Zhong et al. (2015) found that diverse construction methods affect the individual differences in network measures. Dhamala et al. (2021) evaluated the extent to which template selection and tractography (probabilistic and deterministic) methods influence prediction results. However, the effects of different strategies for defining edges have not yet been considered. Second, how do different regression algorithms affect the predictive performance? With structural connectome-based features, we must examine which regression algorithms are more suitable for individual predictions. These methodological pipelines have been extensively evaluated for functional MRI (fMRI) data-based prediction (Cui & Gong, 2018; He et al., 2020; Pervaiz et al., 2020; Scheinost et al., 2019). Finally, researchers have not evaluated which network construction methods and regression algorithms should be used for different cognitive domains. To our knowledge, no study has systematically evaluated the WM structural connectome-based individual predictions pipeline and discussed the potential effects.

In the present study, we focused on the methodological evaluation of brain structural connectome-based cognitive predictions pipeline by assessing different network construction methods with dMRI data and different machine learning regression algorithms. For network construction, we considered two parcellation schemes for defining nodes and seven strategies for defining edges. We used eight regression models for the regression algorithms, including both linear and nonlinear models. Four cognitive domains and brain age were targeted as predictive tasks based on two independent data sets (Beijing Aging Brain Rejuvenation Initiative [BABRI]: 633 healthy older adults; Human Connectome Projects in Aging [HCP-A]: 560 healthy older adults).

2 | MATERIALS AND METHODS

2.1 | Participants

In the present study, two independent data sets were used. One cohort is the BABRI, which included 633 cognitively normal elderly Chinese participants (age range of 45 to 86 years, mean age of 65.5 ± 6.9 years, 393 females) who were recruited by community public health centers. The detailed inclusion and exclusion criteria for the participants have been published (Yang et al., 2021). All participants signed an informed consent form approved by the Institutional

Review Board of the Beijing Normal University Imaging Center for Brain Research and the study conformed to the principles of the Declaration of Helsinki.

The other cohort is Lifespan HCP-A 1.0 data release (<https://www.humanconnectome.org/study/hcp-lifespan-aging>), which is currently screening 560 healthy older adults (age range of 36 to 100 years, mean age of 57.46 ± 14.09 years, 322 females), after matching according to the official demographic information table and selection and quality control during MRI preprocessing. The detailed inclusion and exclusion criteria have been published (Bookheimer et al., 2019).

2.2 | Cognitive composite score

The following four common cognitive domains were included to ensure the consistency between the two data sets: executive function, attention, language, and memory. In the BABRI data set, these four domains were selected from the comprehensive neuropsychological battery of the BABRI (Yang et al., 2021). In the HCP-A data set, the cognitive scores were assessed with the National Institutes of Health (NIH) Toolbox Cognition Battery (<https://www.healthmeasures.net/explore-measurement-systems/nih-toolbox>) and other cognitive tasks (<https://www.humanconnectome.org/study/hcp-lifespan-aging/documentation>). The subscores for each cognitive domain in the HCP-A

data set were matched with those in the BABRI data set. For each cognitive domain, the main composite score was calculated as the sum of z scores from the neuropsychological tests belonging to this domain, while the z score was obtained by subtracting the mean from the raw score and dividing by the SD ($z \text{ score} = [\text{raw} - \text{mean}]/\text{SD}$). Subjects were discarded due to the lack of information for specific cognitive domains. The demographic information and cognitive characteristics of participants included in the BABRI and HCP-A data sets are presented in Table 1. Detailed descriptions of the neuropsychological testing in the two data sets are provided in Text S1, Supporting Information. The histograms showing the distributions of age and cognitive scores are presented in Figure S1.

2.3 | Imaging acquisition

2.3.1 | BABRI data set

The MRI data were acquired with a Siemens Trio 3 T scanner with a 16-channel phased array head coil at the Imaging Center for Brain Research, Beijing Normal University. MRI scanning included the collection of 3D T1-weighted structural MRI with a 1 mm isotropic voxel size (repetition time [TR] = 1900 ms, echo time [TE] = 3.44 ms, inversion time [TI] = 900 ms, flip angle = 9° , field of view [FOV] = $256 \times 256 \text{ mm}^2$, and 176 sagittal slices) and diffusion-

TABLE 1 The demographics and cognitive characteristics of study samples

		Samples (F/M)	Mean \pm SD	Range	
BABRI	Age (years)	393/240	65.54 ± 6.91	45.00 ~ 86.00	
	Executive function	SCWT-C	387/239	0.23 ± 0.59	-3.30 ~ 1.69
		TMT-B			
	Attention	SDMT	383/237	0.24 ± 0.56	-1.78 ~ 2.42
		TMT-A			
	Language	CVFT	390/236	0.27 ± 0.62	-1.60 ~ 2.07
		BNT			
Memory	AVLT	389/239	0.30 ± 0.69	-1.33 ~ 2.42	
	ROCF-delay				
HCP-A	Age (years)	322/238	57.46 ± 14.09	36.00 ~ 100.00	
	Executive function	DCCS	297/209	-0.21 ± 2.13	-4.80 ~ 9.30
		FICAT			
		TMT-B			
	Attention	PCPS	297/209	-0.13 ± 1.51	-3.82 ~ 5.26
		TMT-A			
	Language	CC	299/208	0.00 ± 1.00	-3.17 ~ 3.60
	Memory	RAVLT	293/204	0.15 ± 1.78	-4.69 ~ 4.74
PSMT					

Note: (1) BABRI: SDMT, Symbol Digit Modalities Test; TMT-A, Trail Making Test-A; SCWT-C, Stroop Color and Word Test C; TMT-B, Trail Making Test-B; AVLT, Auditory-Verbal Learning Test; ROCF-delay, Rey-Osterrieth Complex Figure Test; CVFT, Category Verbal Fluency Test; BNT, Boston Naming Test. (2) HCP-A: DCCS, Dimensional Change Card Sort Test; FICAT, Flanker Inhibitory Control and Attention Test; TMT-B, Trail Making Test-B; PCPS, Pattern Completion Processing Speed Test; TMT-A: Trail Making Test-A; CC, Crystallized Cognition, which is derived by Picture Vocabulary and Oral Reading Recognition; RAVLT, Ray Auditory Verbal Learning Test; PSMT, Picture Sequence Memory Test.

weighted MRI (DWI) with a 2 mm isotropic voxel size (30 diffusion directions with $b = 1000 \text{ s/mm}^2$ and an image with $b = 0 \text{ s/mm}^2$, $TR = 9500 \text{ ms}$, $TE = 92 \text{ ms}$, flip angle = 90° , $FOV = 256 \times 256 \text{ mm}^2$, and 70 axial slices).

2.3.2 | HCP-A data set

The MRI data were acquired at four imaging sites using a matched Siemens Prisma 3 T scanner with a 32-channel head coil. The high-quality MRI scans included T1-weighted structural images with a 0.8 mm isotropic voxel ($TR = 2500 \text{ ms}$, $TE = 1.81/3.6/5.39/7.18 \text{ ms}$, $TI = 1000 \text{ ms}$, flip angle = 8° , $FOV = 256 \times 256 \text{ mm}^2$, and 208 sagittal slices) and DWI images with a 1.5 mm isotropic voxel size (185 diffusion directions with both $b = 1500 \text{ s/mm}^2$ and $b = 3000 \text{ s/mm}^2$ and 14 DWIs with $b = 5 \text{ s/mm}^2$, $TR = 3230 \text{ ms}$, $TE = 89.20 \text{ ms}$, flip angle = 78° , $FOV = 210 \times 210 \text{ mm}^2$, and 92 axial slices).

2.4 | Image preprocessing

The preprocessing procedures for diffusion magnetic resonance imaging (dMRI) data comprised the correction of the eddy current and motion artifacts, estimation of the diffusion tensor elements, and calculation of the fractional anisotropy (FA). The eddy current distortions and motion artifacts in the dMRI data were corrected by applying an affine alignment of each DWI image to the b_0 image using the `eddy_correct` command in the FMRIB's Diffusion Toolbox (FDT) toolbox of FMRIB Software Library (FSL) (<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT>). For the HCP dataset, we applied minimal preprocessing pipelines (Glasser et al., 2013) by referring to the publicly available code from <https://github.com/Washington-University/HCPpipelines>, including intensity normalization, echo planar imaging distortion correction, eddy current, motion artifacts, and gradient nonlinearity correction. The diffusion tensor elements were estimated by solving the Stejskal and Tanner equations, and the FA value of each voxel was calculated using the `dtifit` command in the FDT toolbox of FSL.

For the estimation of fiber orientations in each voxel, the ball-and-stick model estimated from `bedpost` was used (Behrens et al., 2003, 2007; Jbabdi et al., 2012), which is a Bayesian estimation of diffusion parameters using sampling techniques for modeling crossing fibers within each voxel. The `bedpostx_gpu` (Hernandez et al., 2013) command in the FDT toolbox was adopted to quickly estimate multiple fiber orientations (three fibers modeled per voxel) based on the preprocessed dMRI data. Approximately 423 s were required for the analysis of each subject in the BABRI data set and approximately 1551 s per subject in the HCP-A data set.

2.5 | WM network construction

The brain network consists of two main elements: nodes and edges. In this study, we applied two parcellation schemes with different brain

atlases and seven edge definition strategies to evaluate the effect of different WM network construction methods on the performance of cognitive predictions. In addition, two diffusion models (the tensor model and ball-and-stick model) were employed to reconstruct whole-brain WM fiber streamlines. Therefore, for each participant, 14 distinct WM structural networks were constructed with different methods (Figure S2).

2.5.1 | Network node definition

In the present study, we used two popular brain atlases to define network nodes: Automated Anatomical Labeling with 90 cortical and subcortical regions (AAL90) (Tzourio-Mazoyer et al., 2002) and the Human Brainnetome Atlas with 246 brain regions (BNA246) (Fan et al., 2016). Briefly, the dMRI b_0 image was aligned to the native T1 image, and then the native T1 image was normalized to the ICBM-152 T1 template in MNI space using the Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB) Linear Image Registration Tool (FLIRT of FSL, <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FLIRT>) (Jenkinson et al., 2002) and FMRIB Nonlinear Image Registration Tool (FNIRT of FSL, <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FNIRT>). Inverse transformation matrices derived from the aforementioned steps were applied to the brain atlases, and then we obtained whole-brain parcellation with cortical and subcortical areas as the network nodes.

2.5.2 | Network edge definition

Deterministic and probabilistic tractography were performed to define network edges, and different WM network matrices were constructed using different weighting strategies.

Deterministic tractography

Two distinct diffusion models (tensor model and ball-and-stick model) were used to perform deterministic tractography. Deterministic tractography with a single tensor model was performed using the Diffusion Toolkit (<http://www.trackvis.org/dtk/>) and the command-line `dti_tracker` to reconstruct whole-brain fiber tracts. Using the Camino toolbox (<http://camino.cs.ucl.ac.uk/>), a command-line track was used to reconstruct fibers with a ball-and-stick model estimated from `bedpostx`, and command-line `procstreamlines` were adopted to remove false-positive fibers. Based on the deterministic tractography results, a binary network (BN), FA-weighted network (FA), and fiber number weighted network (FN) were constructed for each subject. BN represents the presence or absence of fiber bundles between two regions; if the number of fiber streamlines is greater than 0, it is 1; otherwise, it is 0. Other BN strategies are provided in Text S4. FA weight is defined as the average FA value of the voxels traversed along the connected fibers between two regions. The FN weight is the number of fiber streamlines connecting two brain regions.

Probabilistic tractography

Based on the bedpost estimation of fiber orientations in each voxel, probabilistic tractography was implemented to estimate the probability of connectivity between two regions (Behrens et al., 2007). The probability p_{ij} from region of interest (ROI) i to ROI j was defined as the number of fibers passing through ROI j divided by the total number of fibers sampled from ROI i . For each voxel within the seed region, 5000 fibers were sampled. Importantly, the probability p_{ij} is not necessarily equivalent to the probability p_{ji} because of the tractography dependence on the seeding location, and thus the probability between ROIs i and j is defined as the average of $p_{ij} + p_{ji}$. The `probtrackx2_gpu` command in the FDT toolbox (Hernandez-Fernandez et al., 2019) utilizes GPU acceleration to produce sample streamlines. For the AAL90-based network, approximately 563 s were required to produce results for each subject in the BABRI data set and approximately 3618 s for each subject in the HCP-A data set. The BNA246-based network took approximately 1250 s per subject in the BABRI data set and approximately 4948 s for subjects in the HCP-A data set. Considering that removing spurious connections with relatively low probability and ensuring comparability of features between probabilistic tractography (PT)-derived WM networks and deterministic tractography (DT)-derived networks, we chose the following strategy: two brain regions were considered unconnected if the mean connectivity probability across the subjects was >2 SDs below a given threshold (Cao et al., 2013). The threshold should be selected to ensure that the sparsity of mean probability network is approximate to that of deterministic networks.

2.6 | Machine learning prediction framework

2.6.1 | Connectome-based features

Both edge and node features based on the brain connectome were considered. As the WM structure network is a symmetric matrix, only the upper/lower triangular matrix represented edge features. Additionally, five common nodal metrics for brain connectome research were included as node features: clustering coefficient (Chen et al., 2021), shortest path length (Boot et al., 2020), nodal efficiency (Li et al., 2020), local efficiency (Tuladhar et al., 2016), and degree centrality (Liao et al., 2017), which were calculated with GREYNA software (<http://www.nitrc.org/projects/gretna/>) (Wang et al., 2015). See Text S2 for detailed definitions of nodal network metrics. Then, both edge and node features were concatenated and flattened into a one-dimensional vector. Standard min-max scaling was employed in the training set to estimate the scale values (min and max values) and applied to the testing set (Figure 1c and f).

2.6.2 | Regression algorithms

Seven common regression algorithms were evaluated in this study, including basic ordinary least squares (OLS) regression, least absolute shrinkage and selection operator least absolute shrinkage and selection operator (LASSO) regression, ridge regression, Elastic-Net, linear support vector regression (LSVR), relevance vector regression (RVR),

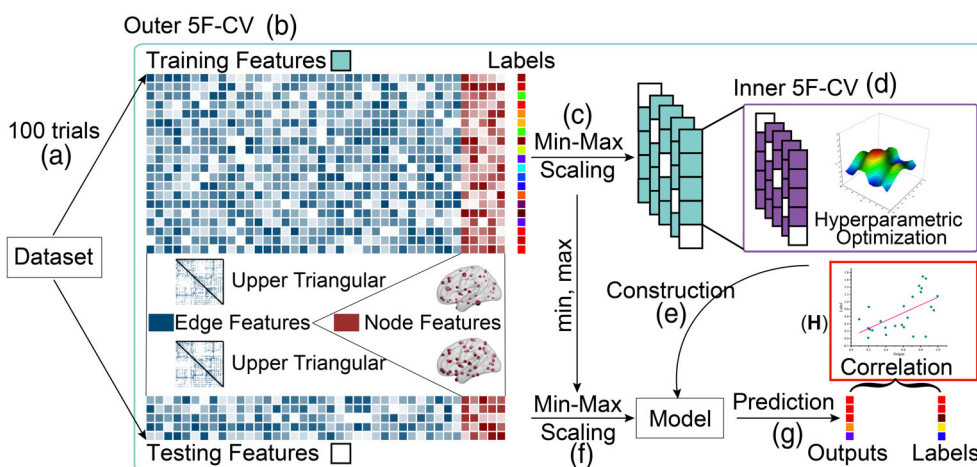


FIGURE 1 Prediction framework used in this study. (a) One hundred trials were performed per group to avoid model bias caused by data partitioning. (b) Outer 5F-CV. The data set is randomly divided into five folds on average, with four folds serving as training features and the remaining one fold serving as the testing feature; the process was repeated five times. (c) Min-max scaling. The training features are executed by scaling normalization, and min and max values are obtained after fitting. (d) Inner 5F-CV. The training features are further randomly divided into five folds, of which four folds are used for model training under the hyperparameter candidate combinations, and the remaining one fold is used to test the model accuracy and determine the optimal hyperparameter. This step was omitted if the regression algorithm did not involve hyperparameter selection. (e) Model construction. Based on the determined optimal hyperparameter combination, the training model is constructed according to the training set divided in (b). (f) Application of min-max scaling. The testing features from (b) are subjected to scaling normalization using min and max values from (c). (g) Prediction. The testing features from (b) are input to generate prediction outputs. (h) Model evaluation. The predicted outputs and corresponding labels were calculated using Pearson's correlation analysis to obtain the prediction accuracy of the model

and partial least squares regression (PLSR). Usually, the dimension of neuroimaging features is much larger than the sample size of the training set, which may lead to the problem of overfitting. Regularization terms in the objective functions of some regression algorithms automatically select features, reduce redundancy, simplify the model, and enhance generalization. The PLSR algorithm solves the overfitting problem using a component extraction strategy from the original data. Notably, a deep learning method, the multi-layer perceptron neural network (MLP), was also examined in the experiment. Compared with some complex network structures, MLP has a simple structure and high performance (Tolstikhin et al., 2021). Detailed descriptions of the aforementioned regression algorithms are provided in Text S3.

2.6.3 | Nested cross-validation framework

In the present study, a nested n -fold ($n = 5$) CV (5F-CV) framework was adopted, which consists of both outer CV and inner CV. In the outer CV, the data set is randomly divided into n folds, one-fold is used for the testing set, and the remaining $n-1$ folds are used for the training model. By repeating the process n times, n predictive models are obtained, and the averaged predictive accuracy across n times is the accuracy of the final model. Inner CV is used to search for optimal hyperparameters. Within the candidate set of hyperparameters, four-fifths of the training set was used to train the model under different hyperparameters, and the remaining one-fifth of the training set was used to determine the optimal parameters (Figure 1b and d). An inner CV step was skipped if the regression algorithm did not contain hyperparameters.

2.6.4 | Construction and evaluation of predictive models

First, the optimal hyperparameters were determined based on the inner CV (Figure 1d), and the predictive model was determined using the training set of outer CV (Figure 1e). Then, the model was applied to the test set to generate predicted outputs (Figure 1g). Finally, Pearson's correlation coefficients between predicted outputs and actual labels were calculated to assess the model performance (Figure 1h).

2.6.5 | Experimental setup and implementation

Based on the extracted features from 14 WM networks with distinct construction methods, 8 regression algorithms were used to predict 5 cognition-related measures across 2 independent datasets. Each experiment included 100 trials; therefore, a total of $14 \times 8 \times 5 \times 100 \times 2 = 112,000$ trials were performed in this study. The experimental settings of the regression algorithms are summarized in detail below.

For OLS and RVR, inner CV was excluded because of the lack of hyperparameters. For LASSO and Ridge, regularization strength $\alpha \in \{x | x = 2^n, n \in \mathbf{Z}, n \in [-10, 5]\}$ was used to improve the conditioning of the problem and reduce the variance of the estimates. Larger values specify stronger regularization. For Elastic-Net, the constant $\alpha \in \{x | x = 2^n, n \in \mathbf{Z}, n \in [-10, 5]\}$ was the regularization strength, and the tradeoff parameter $\beta \in \{x | x = 0.1n, n \in \mathbf{Z}, n \in [0, 10]\}$ was applied to the L1 penalty and L2 penalty. A raster search was performed to find the optimal combination among the 16×11 hyperparametric combinations. For LSVR, the regularization parameter $c \in \{x | x = 2^n, n \in \mathbf{Z}, n \in [-5, 10]\}$ was inversely proportional to the regularization strength. For PLSR, $n \in \{x | x \in [1, 6], x \in \mathbf{Z}\}$ was the number of components. For MLP, three hidden layers with neural units [256 128 64] and four hidden layers with neural units [512 256 128 64] were adopted. The dropout ratio $d \in \{x | x = 0.1n, n \in \mathbf{Z}, n \in [0, 5]\}$ was applied to all hidden layers, enhancing generalization and simplifying the model. The leaky rectified linear unit nonlinearity activation function limited the output amplitude to avoid losing negative weight information. The mean absolute error loss and adaptive moment estimation gradient optimizer with a learning rate $lr = \{0.001, 0.0001, 0.00001\}$ and no batch size strategy were incorporated to calculate and update the model parameters.

OLS, LASSO, Ridge, Elastic-Net, LSVR, and PLSR were implemented in scikit-learn library 0.24.0 (<https://scikit-learn.org/>), RVR was executed referencing <https://github.com/AmazaspShumik/sklearn-bayes>, and MLP was implemented in pytorch 1.9.0 (<https://pytorch.org/>) using an NVIDIA RTX 3090.

2.7 | Statistical analyses

Pearson's correlation coefficients and mean absolute error (MAE) between predicted scores and actual scores were calculated to evaluate the predictive performance. To simplify the description, Pearson's correlation coefficient is used as the main measure for the subsequent descriptions of the prediction results. Since correlations of prediction were not normally distributed, group differences in predictive performances between two brain atlases (AAL90 and BNA246) were compared using a two-sample Wilcoxon rank-sum test. The pairwise comparisons among different edge definition methods and regression algorithms after inverted rank sorting were performed with a two-sample Wilcoxon rank-sum test. The Bonferroni correction was used to correct for multiple comparisons. All the aforementioned statistical analyses were performed using R (<https://www.r-project.org/>).

3 | RESULTS

3.1 | Predictive performances across different cognitive domains

In this study, the predictive performance brain age and cognitive scores across four domains was comprehensively analyzed. For each

predictive measure, 112 candidate combinations (2 node definition atlases, 7 edge definition methods, and 8 regression algorithms) were obtained, and 100 trials were conducted per combination. The predictive performance of each combination was reported as the average value across 100 trials. As shown in Table 2, the mean, std, min, median, and max values of prediction precision (Pearson's correlation coefficient) were calculated for each measure. The prediction results for executive function and attention were most stable among the four cognitive domains in both the BABRI and HCP-A data sets. For memory, prediction results were good in HCP-A but not in BABRI. Moreover, prediction results for language were relatively poor in both data sets. Regarding brain age, the prediction accuracy in the HCP-A dataset was significantly higher than that in the BABRI data set.

3.2 | The effect of the node definition on predictive power

As shown in Figure 2, split violin plots show the predictive performances obtained with different node definition atlases in the BABRI (Figure 2a) and HCP-A (Figure 2b) data sets. Group comparisons

between the two atlases showed that the BNA246 atlas significantly outperformed the AAL90 atlas in predicting brain age and all cognitive measures ($p < .001$, Bonferroni correction) in both data sets.

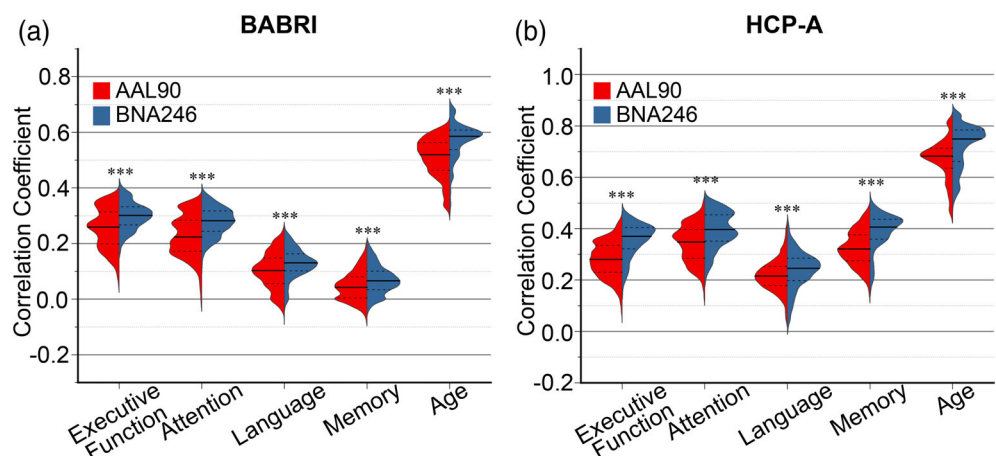
3.3 | The effect of the edge definition on predictive power

The axial split violin diagram in Figure 3a shows the predictive ability of the seven edge definition methods for different measures in the BABRI and HCP-A data sets. The heatmap in Figure 3b represents the pairwise comparisons across different edge definition methods after ranking (in reverse order) to obtain a statistical comparison. For BABRI, edge definitions based on a single tensor model were superior to the strategy based on the ball-and-stick model ($*_t > *_bs$), and DT was superior to PT, especially in predicting executive function and attention. For HCP-A, the edge definitions based on the ball-and-stick model were better than those based on the single tensor model ($*_bs > *_t$). Meanwhile, PT was better than DT, showing excellent predictive performance. In general, the edge definition with FA weight exhibited excellent performance in predicting executive function and attention cognitive tasks, where FA_t was the best for BABRI and

TABLE 2 The predictive performance for age and four cognitive composition scores

	Predicted measure	R (Pearson correlation)		MAE	
		Mean \pm SD	Min \sim max	Mean \pm SD	Min \sim max
BABRI	Executive function	0.28 \pm 0.06	0.09 \sim 0.40	0.45 \pm 0.06	0.37 \sim 0.86
	Attention	0.25 \pm 0.07	0.02 \sim 0.37	0.45 \pm 0.05	0.38 \sim 0.89
	Language	0.12 \pm 0.06	-0.01 \sim 0.22	0.53 \pm 0.05	0.46 \sim 0.93
	Memory	0.06 \pm 0.05	-0.02 \sim 0.17	0.60 \pm 0.07	0.52 \sim 1.15
	Age	0.54 \pm 0.07	0.33 \sim 0.69	4.83 \pm 0.71	3.67 \sim 7.49
HCP-A	Executive function	0.32 \pm 0.07	0.11 \sim 0.45	1.63 \pm 0.25	1.28 \sim 2.75
	Attention	0.37 \pm 0.07	0.17 \sim 0.50	1.19 \pm 1.18	0.93 \sim 2.20
	Language	0.23 \pm 0.06	0.06 \sim 0.40	8.35 \pm 5.11	6.72 \sim 59.53
	Memory	0.36 \pm 0.07	0.18 \sim 0.50	1.41 \pm 0.19	1.13 \sim 2.24
	Age	0.69 \pm 0.08	0.48 \sim 0.84	8.89 \pm 3.40	5.48 \sim 30.96

FIGURE 2 The effect of network node definitions on predictive power. (a) Split violin plot of prediction results for the BABRI cohort. (b) Split violin plot of prediction results for the HCP-A cohort. Group comparisons between two atlases were performed with the two-sample Wilcoxon rank-sum test ($***p < 0.001$, Bonferroni correction)



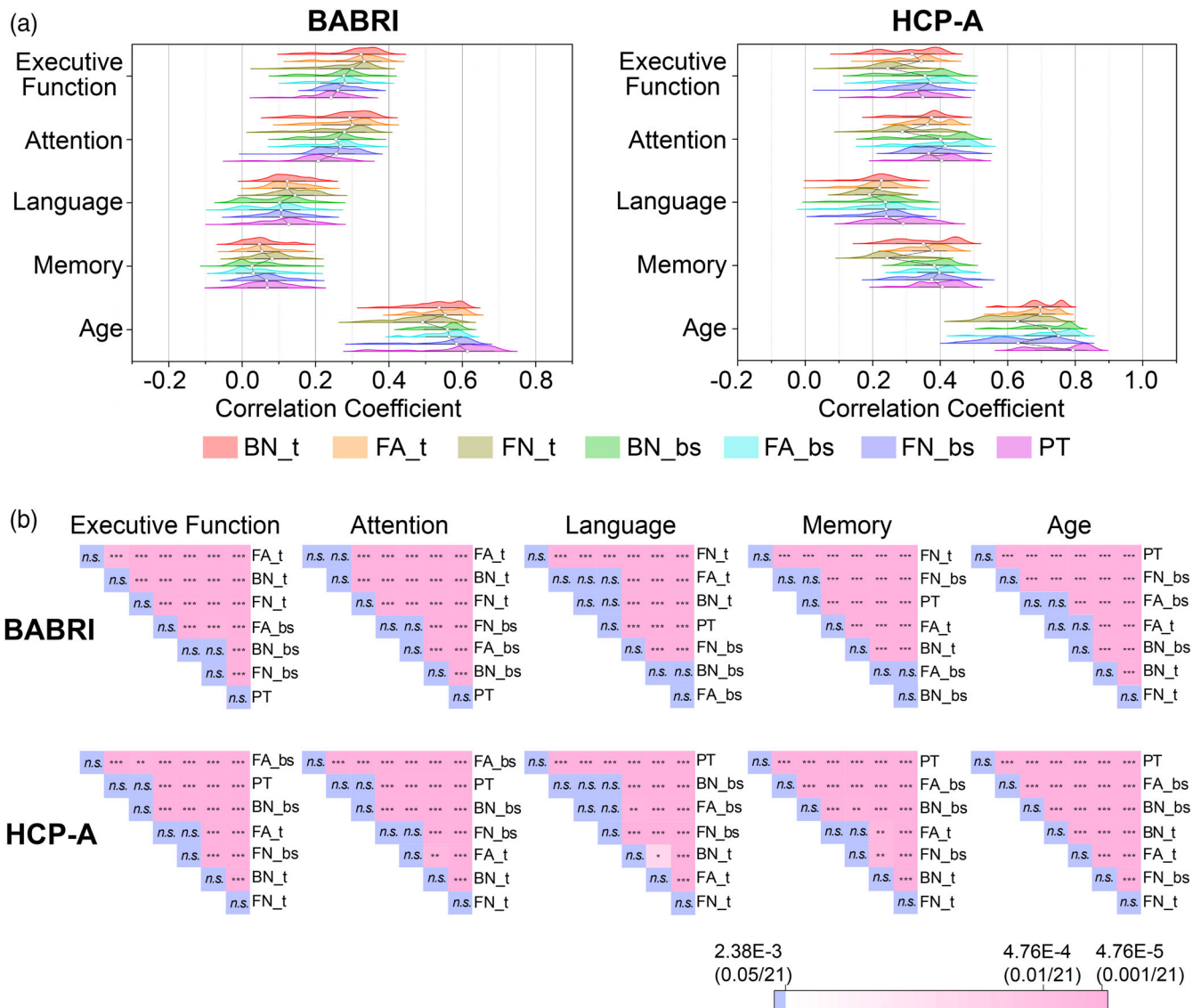


FIGURE 3 The effect of network edge definitions on predictive power. (a) The left panel shows the results from the BABRI cohort, and the right panel shows the results from the HCP-A cohort. The median point line of each half violin diagram is used to visually identify the size of the comparison value. In the legend below, different colors correspond to different edge definitions. (b) the pairwise comparison of edge definition methods after ranking (in reverse order) with a two-sample Wilcoxon rank-sum test among various measures in the BABRI and HCP-A data sets. The heatmaps denote p values from the two-sample Wilcoxon rank-sum test ($***p < 0.001/n$; $**p < 0.01/n$; $*p < 0.05/n$, $n = 21$, Bonferroni correction), where pink represents significant p values, and blue represents nonsignificant p values. The abbreviation “_t” refers to the single tensor model, and “_bs” refers to the ball-and-stick model. Abbreviations of the form “BN_t” refers to using BN edge definition based on single tensor models

FA_bs was more suitable for HCP-A. For predicting brain age, both datasets consistently indicated that PT performed better than DT.

3.4 | The effect of the regression algorithm on the predictive power

The violin map of the eight regression algorithms in Figure 4a illustrates distributions of prediction results for different measures in the BABRI and HCP-A data sets. Pairwise comparisons across different algorithms are presented in heatmaps (Figure 4b) after ranking

(in reverse order). Based on the results, MLP achieved the best prediction results for all prediction tasks. Among the seven traditional regression algorithms, Elastic-Net exhibited an optimal and robust prediction ability in most prediction tasks.

3.5 | Recommended optimal combination

Figure 5 provides a visualization of all possible combinations in the structural connectome-based cognitive prediction pipeline. The options with higher red proportions are consistent with previous

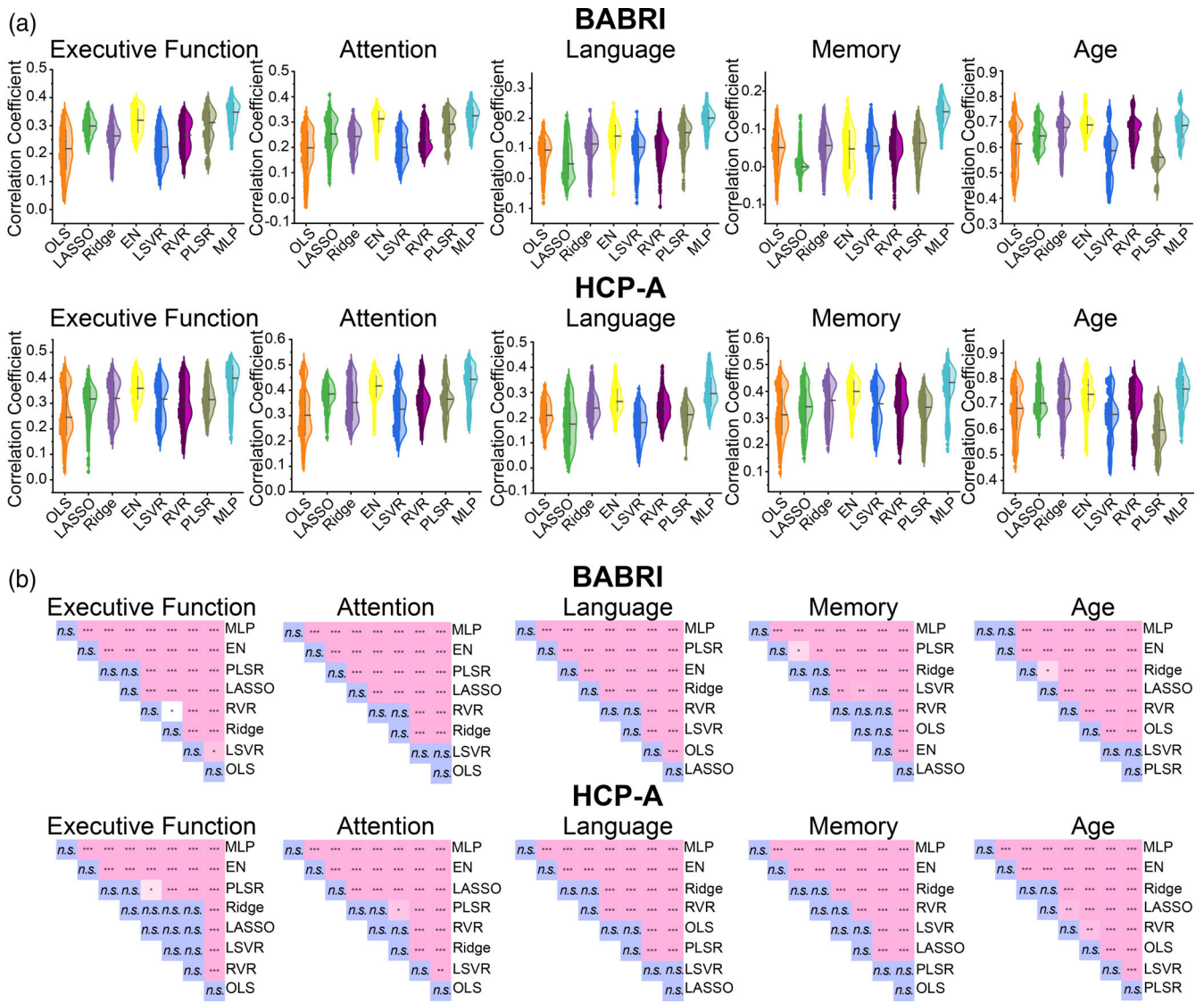


FIGURE 4 The effect of regression algorithms on predictive power. (a) The violin maps denote the results predicted by eight regression algorithms for different measures in the BABRI and HCP-A cohorts. (b) The pairwise comparison of regression algorithms after ranking (in reverse order) with the two-sample Wilcoxon rank-sum test ($***p < 0.001/n$; $**p < 0.01/n$; $*p < 0.05/n$, $n = 28$, Bonferroni correction) among various measures in the BABRI and HCP-A cohorts. The heatmaps denote p values from the two-sample Wilcoxon rank-sum test, pink indicates significant p values, and blue represents nonsignificant p values. The abbreviation “EN” refers to the elastic-net algorithm

statistical results. Table 3 shows the recommended combination for each measure in each data set and the mean \pm variance of 100 trials of each combination. Notably, MLP appears in the optimal combination for all measures. Considering its complex parameter adjustment procedure and weak interpretability (Cichy & Kaiser, 2019), cost-effective combinations including only traditional regression algorithms are also presented for reference.

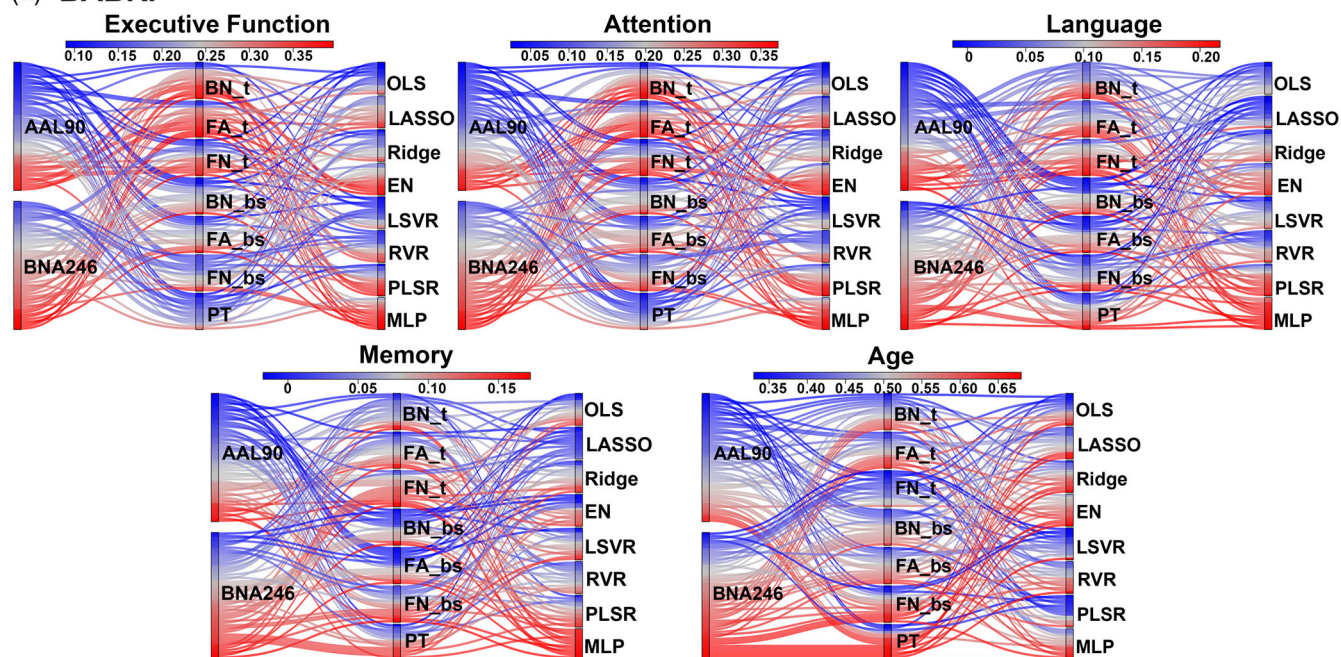
3.6 | Cross-validation in independent datasets

Using two independent data sets, the generalizability, specificity, and interpretability of the predictive models were investigated.

3.6.1 | Generalizability

To assess the generalizability of predictive models, we selected executive function and attention scores with the highest prediction accuracy in both data sets. The composite scores from different data sets were adjusted to a uniform distribution based on the z score. Subsequently, a cost-effective combination was used to train the model based on one data set, and the independent test was performed in the other data set. Four models were generated and applied for external independent validation. As shown in Figure 6, all four models significantly predicted the corresponding cognitive scores in the independent test data set.

(a) BABRI



(b) HCP-A

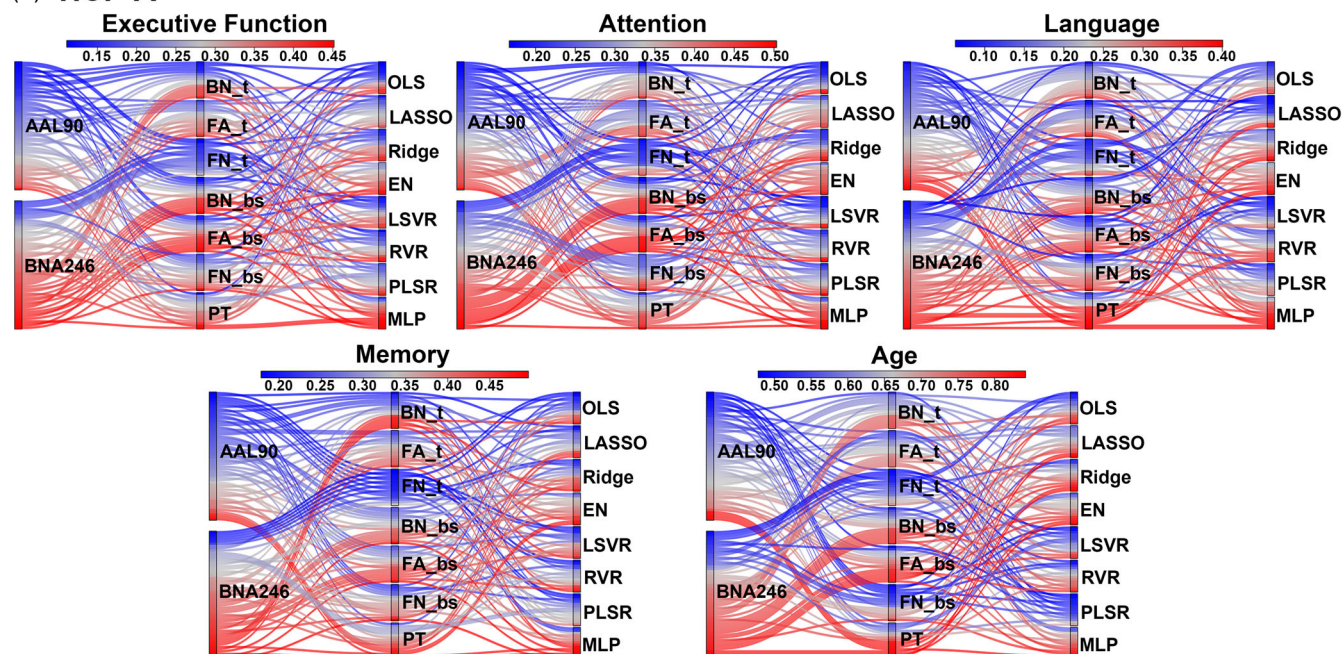


FIGURE 5 A visualization of all possible combinations of options in the pipeline. The vertical sorting is selected relative to the prediction performance for each subblock in ascending order of prediction correlation/accuracy. Each line of the combination is color-coded according to prediction performance. In each subblock, the color proportion is clearly distinguished, where a higher red proportion indicates the inclusion of more high-precision combinations, conversely, a higher blue proportion indicates the inclusion of more low-precision combinations. The abbreviation “_t” refers to the single tensor model, “_bs” refers to the ball-and-stick model. Abbreviations of the form “BN_t” refers to using BN edge definition based on single tensor models. And “EN” refers to the elastic-net algorithm

3.6.2 | Specificity

To investigate the specificity of the predictive models, the models for executive function and attention were also applied to predict

language and memory. We generated predictive models according to 1000 permutations of predictive variables and obtained a null distribution for model performance. Effective specificity was described, and the original prediction accuracy was less than 99%. Table 4 shows

TABLE 3 The recommended optimal combination

	Predicted measure	High-precision combination	Cost-effective combination
BABRI	Executive function	246_FA_t_MLP (0.40 ± 0.01)	246_FA_t_Elastic-Net (0.38 ± 0.01)
	Attention	246_FA_t_MLP (0.37 ± 0.01)	246_FA_t_Elastic-Net (0.35 ± 0.02)
	Language	90_FN_t_MLP (0.22 ± 0.02)	90_FN_t_PLSR (0.21 ± 0.01)
	Memory	246_PT_MLP (0.17 ± 0.02)	246_FN_bs_LSVR (0.10 ± 0.02)
	Age	246_PT_MLP (0.69 ± 0.01)	246_PT_Elastic-Net (0.68 ± 0.01)
HCP-A	Executive function	246_FA_bs_MLP (0.45 ± 0.01)	246_FA_bs_Elastic-Net (0.42 ± 0.02)
	Attention	246_FA_bs_MLP (0.50 ± 0.01)	246_FA_bs_PLSR (0.49 ± 0.02)
	Language	246_PT_MLP (0.40 ± 0.02)	246_PT_Elastic-Net (0.35 ± 0.03)
	Memory	246_FN_bs_MLP (0.50 ± 0.02)	246_PT_Elastic-Net (0.47 ± 0.02)
	Age	246_PT_MLP (0.84 ± 0.01)	246_PT_Ridge (0.84 ± 0.01)

Note: All combinations are denoted by abbreviations, for example, “246_FA_t_MLP” refers to using BNA246 for node definition, FA_t for edge definition and MLP for prediction.

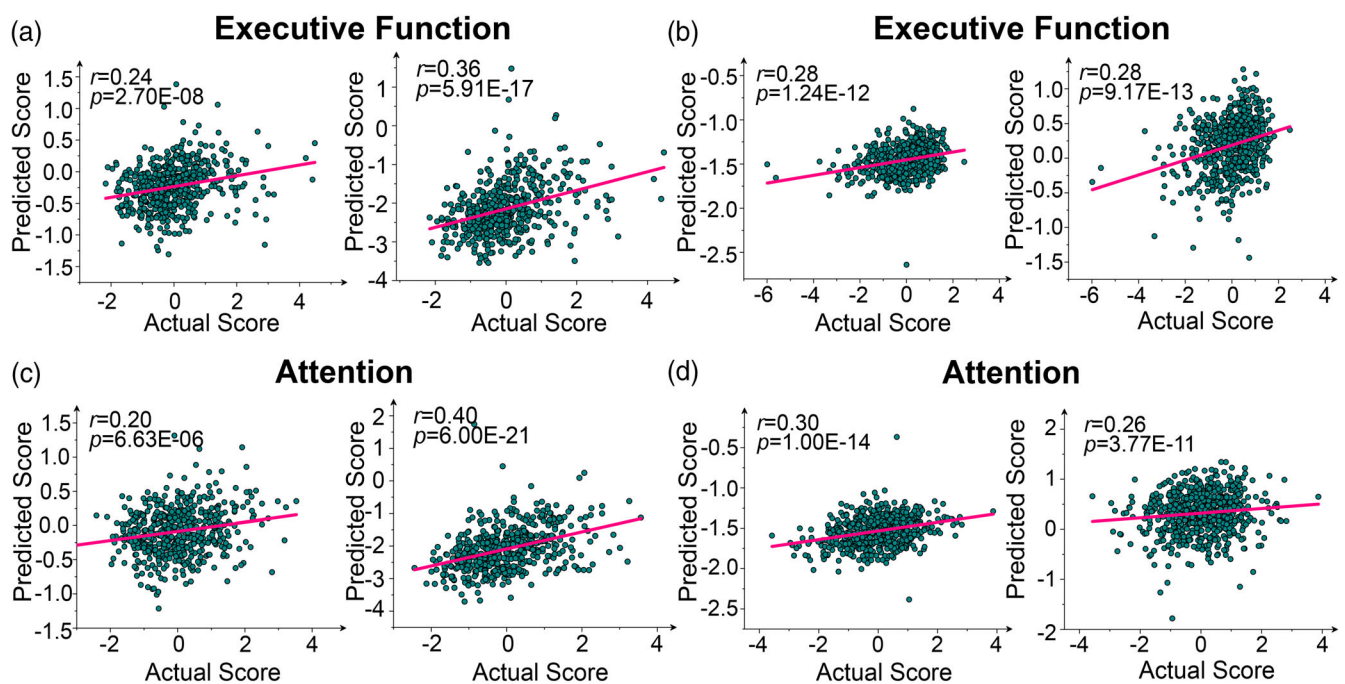


FIGURE 6 External independent validation of predictive models. (a) (b) The validation of executive function, and (c) and (d) show the validation of attention. (a) Training on BABRI using 246_FA_t_Elastic-net, and testing on HCP-A using 246_FA_t (left panel) and 246_FA_bs (right panel). (b) Training on HCP-A using 246_FA_bs_Elastic-net, and testing on BABRI using 246_FA_t (left panel) and 246_FA_bs (right panel). (c) Training on BABRI using 246_FA_t_Elastic-net, and testing on HCP-A using 246_FA_t (left panel) and 246_FA_bs (right panel). (d) Training on HCP-A using 246_FA_bs_PLSR, and testing on BABRI using 246_FA_t (left panel) and 246_FA_bs (right panel). Above all combinations are denoted by abbreviations, for example, “246_FA_t_MLP” refers to using BNA246 for node definition, FA_t for edge definition, and MLP for prediction

the specific model predicting other cognitive scores with distinct FA features derived from two diffusion models. Model specificity was observed for the prediction of language but not for memory.

3.6.3 | Interpretability

Brain regions with positive and negative contributions for the prediction were determined by calculating the sum of all positive and negative weights of selected features, including both nodal metrics and

connections for each region. The top 20% of brain regions contributing to predicting executive function and attention are shown in Figure 7. The regions for predicting executive function in both data sets were congruently located in the superior frontal gyrus, middle frontal gyrus, inferior frontal gyrus, precentral gyrus, middle temporal gyrus, inferior parietal lobule, precuneus, hippocampus, basal ganglia, and thalamus. For predicting attention, the contributed regions in both data sets were mainly located in the superior frontal gyrus, middle frontal gyrus, paracentral lobule, inferior temporal gyrus, superior parietal lobule, precuneus, lateral occipital cortex, hippocampus, and basal ganglia.

TABLE 4 The specificity of predictive models

	Predictive model	Language		Memory	
		FA_t	FA_bs	FA_t	FA_bs
Executive function	BABRI_FA_t	0.03 (32.5%)	0.06 (70.9%)	0.24 (100%)	0.32 (99.5%)
	HCP-A_FA_bs	0.11 (76.0%)	-0.08 (81.9%)	0.15 (99.9%)	0.06 (84.0%)
Attention	BABRI_FA_t	0.12 (96.6%)	0.07 (74.1%)	0.12 (98.8%)	0.03 (16.5%)
	HCP-A_FA_bs	0.01 (4.6%)	0.04 (36.8%)	0.19 (100%)	0.34 (100%)

Note: (*) means probability that original prediction accuracy is greater than random prediction accuracy. All combinations are denoted by abbreviations, for example, "BABRI_FA_t" refers to using FA_t for edge definition in BABRI.

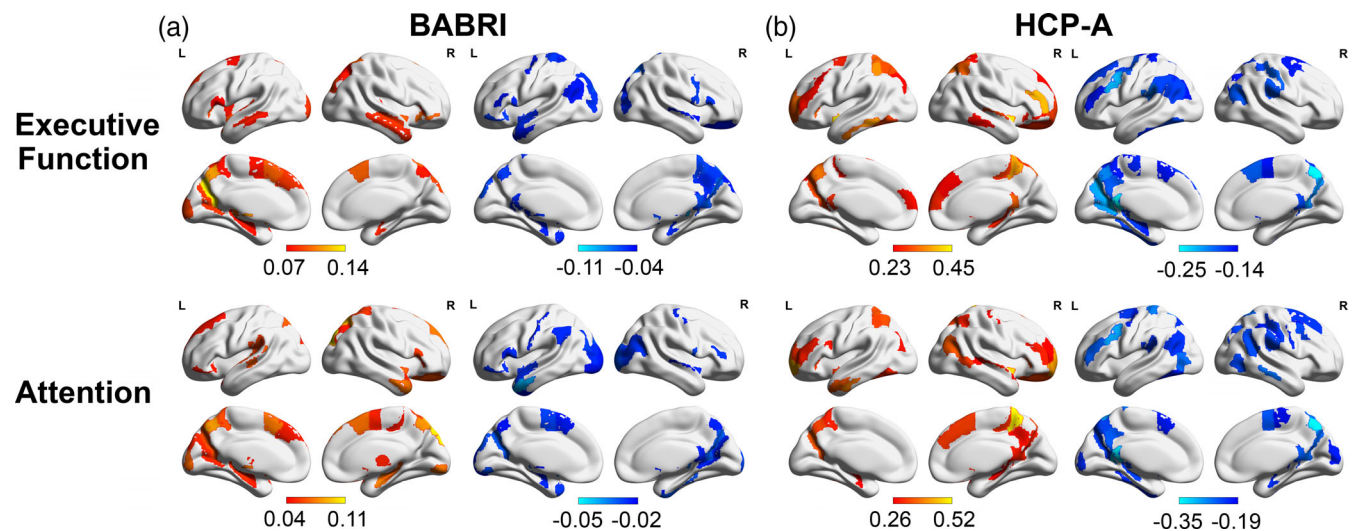


FIGURE 7 The distribution of brain regions contributing to the cognitive prediction. The top 20% of brain regions contributing to the prediction of executive function and attention are shown with lateral and medial views of the right (RH) and left (LH) hemispheres in the BABRI (a) and HCP-A (b) data sets. Warmer colors indicate positive weights, and cooler colors indicate negative weights. The results were visualized using BrainNet viewer software (<https://www.nitrc.org/projects/bnv/>) (Xia et al., 2013)

4 | DISCUSSION

To our knowledge, this study is the first to systematically evaluate the pipeline of WM structural connectome-based individual predictions and discuss the potential effects in two independent large cohorts. First, we confirmed that the WM connectome contributes to the prediction of multiple cognitive functions and age. Second, we showed that the prediction performance is influenced by WM network construction, including node and edge definitions. Third, different regression algorithms affect the predictive performance to varying degrees. Fourth, through CV in independent data sets, we investigated the generalizability, specificity, and interpretability of predictive models.

4.1 | WM connectome features tend to predict cognitive function and age

Cognitive function and age present pronounced individual differences in measures of brain connectivity. Previous dMRI studies have found significant associations between microstructural connectivity patterns

and age (Cole & Franke, 2017) and various cognitive functions (Bennett & Madden, 2014; Coelho et al., 2021; Palop & Mucke, 2016), such as processing speed (Kochunov et al., 2016), fluid intelligence (Zimmermann et al., 2018), language (Lin et al., 2020), executive function (Li et al., 2020; Shu et al., 2012), and attention (Li et al., 2020). Although the WM connectome is undeniably related to age and cognition, few studies have analyzed its ability to predict cognitive functions (Kim et al., 2021; Li et al., 2020; Madole et al., 2020; Mansour et al., 2021). In this study, we predicted four cognitive measures and age using WM connectome features (edge and node features). Consistent results were obtained from both cohorts and revealed comparable individual predictive power compared with advanced methods (Beck et al., 2021; Fountain-Zaragoza et al., 2019; Gao et al., 2020; Kim et al., 2021; Madole et al., 2020; Mansour et al., 2021), which is associated with the correlation between WM and cognitive function and age from the perspective of prediction. Executive function and attention had higher predictive power for other cognitive functions, supporting significant statistical results that executive function and attention are related to WM in individuals with age-related normal degeneration or traumatic brain injury (Bai

et al., 2020; Bubb et al., 2018; Chopra et al., 2018; Cristofori et al., 2015; Madole et al., 2020; Webb et al., 2020). The predictive accuracy of age was higher than that of cognitive functions, and age appears to be more sensitive to changes in brain structure. On the other hand, the conflicting results for memory prediction might be influenced by the difference between the cognitive measures used in the two data sets, and this difference might also explain the inter-cohort difference in the accuracy of predictions of age and different cognitive domains.

4.2 | The effect of network construction methods on predictive power

4.2.1 | Node definition

Prior template selection can affect individual neuroimaging analyses (Zalesky et al., 2010; Zhong et al., 2015). From the perspective of prediction, a recent study (Dhamala et al., 2021) compared the predictive power using two different parcellations that differ in nodal resolution and derivation to construct a WM network, suggesting that high-resolution parcellation might be better for predicting cognition. However, nonhomologous templates weaken the result of comparison when one template is functionally defined and the other is anatomically defined. In our study, the AAL90 parcellation is anatomically defined, and the fine-grained BNA246 parcellation is defined based on both anatomical and functional connections. These two prevalent parcellations are widely used in research supported by WM (Bi et al., 2021; Cao et al., 2013; Huang et al., 2021; Li et al., 2020; Zhong et al., 2015). Our results indicate that BNA246 is significantly superior to AAL90 ($p < .001$, Bonferroni correction) in predicting various measures, indicating that finer division-derived features that more plausibly contribute are useful for obtaining prediction results, although higher-resolution node definitions have theoretical overfitting threats (for more details see Text S5). In addition, this result also seems to be influenced by the ability to depict individual differences (Zalesky et al., 2010; Zhong et al., 2015).

4.2.2 | Edge definition

The edge definition of the structural connectome hinges on tractography (Yeh et al., 2021). Deterministic tractography based on a tensor model is generally considered plagued by fiber crossing, while probabilistic tractography can address this problem (Behrens et al., 2007) but is a trade-off with exorbitant computational time considerations, resulting in the prohibition of related research using large samples. A compromise solution, deterministic tractography based on a ball-and-stick model, is characterized by acceptable time consumption and multidirectional fiber tracking. Multifiber deterministic tractography is highlighted for mapping the connectome (Sarwar et al., 2019) on simulated dMRI compared with other avenues. As the evaluation of these three avenues from a predictive perspective has

not been clearly conducted to date, a presupposition is that a specific central data set may be applied to a suitable fiber reconstruction model. All three methods were included as fiber reconstruction options in the present study. We found that the single tensor model generally outperformed the ball-and-stick model in predicting cognitive function in the BABRI cohort, while the opposite conclusion was obtained for the HCP-A cohort. This model preference across cohorts may be interpreted as influenced by acquisition parameters, since dMRI of HCP-A included hundreds of gradient directions and higher voxel resolution. Interestingly, the prediction performance of the single tensor model was weaker than that of the ball-and-stick model for predicting age in both the BABRI and HCP-A cohorts, possibly because changes in the brain with age are brain-wide and subtle, whereas the ball-and-stick model reconstructs more fiber bundles and captures these subtle changes more easily.

On the other hand, we analyzed the effects of different weighting strategies (BN, FA, and FN) by focusing on execution function and attention that exhibited higher predictive accuracy. Many studies usually choose different weighting strategies with distinct neurophysiological significance for a particular problem, and the inconsistent results seem to be caused by different weighting strategies (Zhong et al., 2015). In the present study, FA > BN > FN was relatively consistent based on the applicable model across the two cohorts. This result supports the validity of FA for capturing individual differences in the structural integrity of WM from the perspective of forecasting cognition (Bennett & Madden, 2014; Coelho et al., 2021; Qi et al., 2018; Sui et al., 2018). Furthermore, the comparison (Figure 3b) revealed that the effect of model selection was greater than that of the weighting strategy. The weighting strategies with the lowest prediction accuracy in the optimal model were greater than or equal to the weighted strategies with the highest prediction accuracy in the sub-optimal model.

4.3 | The effect of regression algorithms on predictive power

Many studies have used a single machine learning algorithm to predict measures without considering the prediction bias associated with different regression algorithms (Cui & Gong, 2018; He et al., 2020; Pervaiz et al., 2020). A comprehensive experimental analysis based on different algorithms has an important role as a reference and guideline for future prediction research in the computational neuroscience field on the selection of algorithms. We showed that various algorithms have heterogeneous performance, where the deep learning algorithm displays its considerable fitting ability compared with traditional regression algorithms and Elastic-Net reveals the most robust performance among traditional regression algorithms. In addition, we also observed an overall trend in which algorithms predict different measures across data sets from different centers in similar patterns, which may indicate that the performance of the algorithm is robust for data sets from different centers. In particular, emerging deep learning models, such as BrainNetCNN (Kawahara et al., 2017) and GCNN (He

et al., 2020), are deemed to change the conventional pattern of the field of cognitive prediction. However, critics draw attention to their limitations, the divergence between the limited range of tasks a deep learning model can currently perform, and the many cognitive functions lacking explanations (Cichy & Kaiser, 2019; Kay, 2018).

4.4 | Generalizability, specificity, and interpretability of the predictive model

Many exploratory studies are based on relatively small samples and tend to report higher precision with a positive bias (Gabrieli et al., 2015; Sui et al., 2020). The results that are statistically rigorous must be based on larger samples, especially for neurodiversity studies that focus on individual differences, because sufficient sampling is required not only for the entire population but also for individual diversity within that population (Woo et al., 2017). Furthermore, the lack of completely independent samples for prospective testing will also lead to data dependence bias. Using the aforementioned guidelines, we conducted CV of the prediction model in two completely independent large data sets to examine the generalizability, specificity, and interpretability of the model.

4.4.1 | Generalizability

Model generalizability has always been a key problem in the field of multicenter cognitive prediction, and the validation of external independent samples is generally considered the most objective test method. A hidden rule is that training and testing across multicenter samples are built in the same feature from a common pipeline. In our experiment, a very unexpected finding was that the model trained using the FA_t feature in the BABRI data set for predicting the performance of FA_t was significantly lower than that of FA_{bs}, and the identical result was observed simultaneously for the prediction of executive function (a) and attention (c) in the HCP-A cohort. However, when the attentional model trained with the FA_{bs} feature from the HCP-A cohort, the predictive performance for BABRI_FA_t was slightly higher than that for BABRI_FA_{bs} (d), and no analogous result was observed for the executive function prediction (b). Overall, this generalizability is not restricted by the selection of the fiber tracking model in the scope of consideration of the experiment in this study. At the same time, this finding reveals that more accurate follow-up analysis results might be obtained by adopting the most suitable data processing method for the data set from each center when simultaneously predicting data from multiple centers.

4.4.2 | Specificity

We evaluated the domain specificity of the constructed predictive model described in Table 4. None of the four models predicted performance on the language task, as the predictive precision was less than

99% of the random accuracy of 1000 permutation tests, indicating the specificity of the model. In contrast, no definite specificity was observed for predicting memory tasks. The result is that memory and attention functions are likely to be predictive of their shared processes with processing speed (Gao et al., 2020), while attention is usually significantly correlated with executive function.

4.4.3 | Interpretability

Machine learning not only provides models with higher predictive accuracy but also, more importantly, helps us understand and interpret the necessary and sufficient representation basis of the brain (Sui et al., 2020; Woo et al., 2017). Its heuristic techniques have the exceptional advantage of simplifying models and giving humans intuitively readable feature weights to provide potential biological interpretability. Consistent with previous studies (Bennett & Madden, 2014; Li et al., 2020; Lin et al., 2020), our visualized feature mapping showed multiple brain regions (frontal, parietal, temporal, limbic system, etc.) with strong contributions to executive function and attention across distinct cohorts.

4.5 | Methodological issues

Several methodological issues in this study must be addressed. First, our parcellation-level WM connectomes may not be sufficiently precise to represent fine features for predicting individual cognitive functions and brain age. The vertex-level WM connectome with ultrahigh resolution has been shown to have potential advantages for individual predictions (Mansour et al., 2021). However, it was not considered in this study due to the inherent computational complexities of handling a high-dimensional connectome. Second, we should examine combining more fiber reconstruction methods (e.g., other model-based methods or model-free methods) and tracking algorithms (e.g., machine learning) to reconstruct accurate fibers focusing on data from different centers (Jeurissen et al., 2019; Yeh et al., 2021) and explore its convergence and divergence of predictive performance in the future. Third, the predictive efficacy of different sets of graph parameters should be further examined. Finally, our general results are based on two data sets of healthy elderly individuals, and further studies are needed to explore whether the results can be generalized to other age groups (adolescents or young adults).

5 | CONCLUSIONS

In the present study, we systematically evaluated different options in the pipeline of brain WM connectome-based cognitive predictions in two large data sets. Both WM network construction methods and regression algorithms influence the predictive performances for cognitive functions and brain age in elderly subjects. The results indicated that the effects of different options in the predictive pipeline should be

considered and evaluated in future studies. Therefore, our study may provide important methodological references and guidelines for future research on brain WM connectome-based individual predictions.

ACKNOWLEDGMENTS

The authors thank all the volunteers for their participation in the study and anonymous reviewers for their insightful comments and suggestions. This work was supported by the National Natural Science Foundation of China (No. 81671761 and 81871425); Funds for International Cooperation and Exchange of the National Natural Science Foundation of China (No. 81820108034); National Key Research and Development Project of China (No. 2018YFC1315200); Fundamental Research Funds for the Central Universities (No. 2017XTCX04); Open Research Fund of the State Key Laboratory of Cognitive Neuroscience and Learning (No. CNLYB2001). Data in this publication were provided (in part) by the Human Connectome Project-Aging (HCP-A), which is supported by the National Institute On Aging of the National Institutes of Health under Award Number U01AG052564. The content in this publication is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

CONFLICT OF INTEREST

There are no conflicts of interest including any financial, personal, or other relationships with people or organizations for any of the authors related to the work described in the article.

AUTHOR CONTRIBUTIONS

Guozheng Feng: Software, Data curation, Writing - original draft, Writing - review & editing. Yiwen Wang: Data curation, Writing - original draft, Writing - review & editing. Weijie Huang: Writing - Software, Data curation, Supervision. Haojie Chen: Software, Data curation. Zhengjia Dai: Writing - review & editing, Supervision. Guolin Ma: Writing - review & editing. Xin Li: Writing - review & editing. Zhanjun Zhang: Writing - review & editing. Ni Shu: Software, Data curation, Writing - review & editing, Supervision.

DATA AVAILABILITY STATEMENT

The data sets generated and analyzed in the present study will be made available from the corresponding author to other scientists on request in anonymized format and according to data protection policy in the ethics agreement. Code availability The Python, MATLAB, and R code for MRI data processing and experiments is available from the authors upon request.

ORCID

Guozheng Feng  <https://orcid.org/0000-0002-6937-8592>
Zhanjun Zhang  <https://orcid.org/0000-0001-7266-4218>

REFERENCES

- Bai, L., Bai, G., Wang, S., Yang, X., Gan, S., Jia, X., Yin, B., & Yan, Z. (2020). Strategic white matter injury associated with long-term information processing speed deficits in mild traumatic brain injury. *Human Brain Mapping*, 41(15), 4431–4441. <https://doi.org/10.1002/hbm.25135>
- Beaty, R. E., Kenett, Y. N., Christensen, A. P., Rosenberg, M. D., Benedek, M., Chen, Q., Fink, A., Qiu, J., Kwapił, T., Kane, M. J., & Silvia, P. J. (2018). Robust prediction of individual creative ability from brain functional connectivity. *Proceedings of the National Academy of Sciences of the United States of America*, 115(5), 1087–1092. <https://doi.org/10.1073/pnas.1713532115>
- Beck, D., de Lange, A. G., Maximov, I. I., Richard, G., Andreassen, O. A., Nordvik, J. E., & Westlye, L. T. (2021). White matter microstructure across the adult lifespan: A mixed longitudinal and cross-sectional study using advanced diffusion models and brain-age prediction. *NeuroImage*, 224, 117441. <https://doi.org/10.1016/j.neuroimage.2020.117441>
- Behrens, T. E., Berg, H. J., Jbabdi, S., Rushworth, M. F., & Woolrich, M. W. (2007). Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage*, 34(1), 144–155. <https://doi.org/10.1016/j.neuroimage.2006.09.018>
- Behrens, T. E., Woolrich, M. W., Jenkinson, M., Johansen-Berg, H., Nunes, R. G., Clare, S., Matthews, P. M., Brady, J. M., & Smith, S. M. (2003). Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance in Medicine*, 50(5), 1077–1088. <https://doi.org/10.1002/mrm.10609>
- Bennett, I. J., & Madden, D. J. (2014). Disconnected aging: Cerebral white matter integrity and age-related differences in cognition. *Neuroscience*, 276, 187–205. <https://doi.org/10.1016/j.neuroscience.2013.11.026>
- Bi, Q., Wang, W., Niu, N., Li, H., Wang, Y., Huang, W., Chen, K., Xu, K., Zhang, J., Chen, Y., Wei, D., Cui, R., Shu, N., & Zhang, Z. (2021). Relationship between the disrupted topological efficiency of the structural brain connectome and glucose hypometabolism in normal aging. *NeuroImage*, 226, 117591. <https://doi.org/10.1016/j.neuroimage.2020.117591>
- Bookheimer, S. Y., Salat, D. H., Terpstra, M., Ances, B. M., Barch, D. M., Buckner, R. L., Burgess, G. C., Curtiss, S. W., Diaz-Santos, M., Elam, J. S., Fischl, B., Greve, D. N., Hagy, H. A., Harms, M. P., Hatch, O. M., Hedden, T., Hodge, C., Japardi, K. C., Kuhn, T. P., ... Yacoub, E. (2019). The lifespan human connectome project in aging: An overview. *NeuroImage*, 185, 335–348. <https://doi.org/10.1016/j.neuroimage.2018.10.009>
- Boot, E. M., Mc van Leijzen, E., Bergkamp, M. I., Kessels, R. P. C., Norris, D. G., de Leeuw, F. E., & Tuladhar, A. M. (2020). Structural network efficiency predicts cognitive decline in cerebral small vessel disease. *NeuroImage: Clinical*, 27, 102325. <https://doi.org/10.1016/j.nicl.2020.102325>
- Bubb, E. J., Metzler-Baddeley, C., & Aggleton, J. P. (2018). The cingulum bundle: Anatomy, function, and dysfunction. *Neuroscience and Biobehavioral Reviews*, 92, 104–127. <https://doi.org/10.1016/j.neubiorev.2018.05.008>
- Bullmore, E., & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews. Neuroscience*, 10(3), 186–198. <https://doi.org/10.1038/nrn2575>
- Cao, Q., Shu, N., An, L., Wang, P., Sun, L., Xia, M. R., Wang, J.-H., Gong, G.-L., Zang, Y.-F., Wang, Y.-F., & He, Y. (2013). Probabilistic diffusion tractography and graph theory analysis reveal abnormal white matter structural connectivity networks in drug-naïve boys with attention deficit/hyperactivity disorder. *The Journal of Neuroscience*, 33(26), 10676–10687. <https://doi.org/10.1523/JNEUROSCI.4793-12.2013>
- Chen, Q., Baran, T. M., Turnbull, A., Zhang, Z., Rebok, G. W., & Lin, F. V. (2021). Increased segregation of structural brain networks underpins enhanced broad cognitive abilities of cognitive training. *Human Brain Mapping*, 42(10), 3202–3215. <https://doi.org/10.1002/hbm.25428>
- Chen, Q., Beaty, R. E., Cui, Z., Sun, J., He, H., Zhuang, K., Ren, Z., Liu, G., & Qiu, J. (2019). Brain hemispheric involvement in visuospatial and verbal divergent thinking. *NeuroImage*, 202, 116065. <https://doi.org/10.1016/j.neuroimage.2019.116065>
- Chopra, S., Shaw, M., Shaw, T., Sachdev, P. S., Anstey, K. J., & Cherbuin, N. (2018). More highly myelinated white matter tracts are associated

- with faster processing speed in healthy adults. *NeuroImage*, 171, 332–340. <https://doi.org/10.1016/j.neuroimage.2017.12.069>
- Cichy, R. M., & Kaiser, D. (2019). Deep neural networks as scientific models. *Trends in Cognitive Sciences*, 23(4), 305–317. <https://doi.org/10.1016/j.tics.2019.01.009>
- Coelho, A., Fernandes, H. M., Magalhaes, R., Moreira, P. S., Marques, P., Soares, J. M., Amorim, L., Portugal-Nunes, C., Castanho, T., Santos, N. C., & Sousa, N. (2021). Signatures of white-matter microstructure degradation during aging and its association with cognitive status. *Scientific Reports*, 11(1), 4517. <https://doi.org/10.1038/s41598-021-83983-7>
- Cole, J. H., & Franke, K. (2017). Predicting age using neuroimaging: Innovative brain ageing biomarkers. *Trends in Neurosciences*, 40(12), 681–690. <https://doi.org/10.1016/j.tins.2017.10.001>
- Cristofori, I., Zhong, W., Chau, A., Solomon, J., Krueger, F., & Grafman, J. J. N. (2015). White and gray matter contributions to executive function recovery after traumatic brain injury. *Neurology*, 84(14), 1394–1401.
- Cui, Z., & Gong, G. (2018). The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features. *NeuroImage*, 178, 622–637. <https://doi.org/10.1016/j.neuroimage.2018.06.001>
- Dhamala, E., Jamison, K. W., Jaywant, A., Dennis, S., & Kuceyeski, A. (2021). Distinct functional and structural connections predict crystallised and fluid cognition in healthy adults. *Human Brain Mapping*, 42, 3102–3118. <https://doi.org/10.1002/hbm.25420>
- Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., Laird, A. R., Fox, P. T., Eickhoff, S. B., Yu, C., & Jiang, T. (2016). The human Brainnetome atlas: A new brain atlas based on connective architecture. *Cerebral Cortex*, 26(8), 3508–3526. <https://doi.org/10.1093/cercor/bhw157>
- Fornito, A., Zalesky, A., & Breakspear, M. (2015). The connectomics of brain disorders. *Nature Reviews Neuroscience*, 16(3), 159–172. <https://doi.org/10.1038/nrn3901>
- Fountain-Zaragoza, S., Samimy, S., Rosenberg, M. D., & Prakash, R. S. (2019). Connectome-based models predict attentional control in aging adults. *NeuroImage*, 186, 1–13. <https://doi.org/10.1016/j.neuroimage.2018.10.074>
- Gabrieli, J. D. E., Ghosh, S. S., & Whitfield-Gabrieli, S. (2015). Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience. *Neuron*, 85(1), 11–26. <https://doi.org/10.1016/j.neuron.2014.10.047>
- Gao, M., Wong, C. H. Y., Huang, H., Shao, R., Huang, R., Chan, C. C. H., & Lee, T. M. C. (2020). Connectome-based models can predict processing speed in older adults. *NeuroImage*, 223, 117290. <https://doi.org/10.1016/j.neuroimage.2020.117290>
- Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J. R., Van Essen, D. C., Jenkinson, M., & WU-Minn HCP Consortium. (2013). The minimal preprocessing pipelines for the human connectome project. *NeuroImage*, 80, 105–124. <https://doi.org/10.1016/j.neuroimage.2013.04.127>
- He, T., Kong, R., Holmes, A. J., Nguyen, M., Sabuncu, M. R., Eickhoff, S. B., Bzdok, D., Feng, J., & Yeo, B. T. T. (2020). Deep neural networks and kernel regression achieve comparable accuracies for functional connectivity prediction of behavior and demographics. *NeuroImage*, 206, 116276. <https://doi.org/10.1016/j.neuroimage.2019.116276>
- Hernandez, M., Guerrero, G. D., Cecilia, J. M., Garcia, J. M., Inuggi, A., Jbabdi, S., Behrens, T. E. J., & Sotiropoulos, S. N. (2013). Accelerating fibre orientation estimation from diffusion weighted magnetic resonance imaging using GPUs. *PLoS One*, 8(4), e61892. <https://doi.org/10.1371/journal.pone.0061892>
- Hernandez-Fernandez, M., Reguly, I., Jbabdi, S., Giles, M., Smith, S., & Sotiropoulos, S. N. (2019). Using GPUs to accelerate computational diffusion MRI: From microstructure estimation to tractography and connectomes. *NeuroImage*, 188, 598–615. <https://doi.org/10.1016/j.neuroimage.2018.12.015>
- Huang, W., Li, X., Li, X., Kang, G., Han, Y., & Shu, N. (2021). Combined support vector machine classifier and brain structural network features for the individual classification of amnesic mild cognitive impairment and subjective cognitive decline patients. *Frontiers in Aging Neuroscience*, 13, 687927. <https://doi.org/10.3389/fnagi.2021.687927>
- Jangraw, D. C., Gonzalez-Castillo, J., Handwerker, D. A., Ghane, M., Rosenberg, M. D., Panwar, P., & Bandettini, P. A. (2018). A functional connectivity-based neuromarker of sustained attention generalizes to predict recall in a reading task. *NeuroImage*, 166, 99–109. <https://doi.org/10.1016/j.neuroimage.2017.10.019>
- Jbabdi, S., Sotiropoulos, S. N., Haber, S. N., Van Essen, D. C., & Behrens, T. E. (2015). Measuring macroscopic brain connections in vivo. *Nature Neuroscience*, 18(11), 1546–1555. <https://doi.org/10.1038/nn.4134>
- Jbabdi, S., Sotiropoulos, S. N., Savio, A. M., Grana, M., & Behrens, T. E. (2012). Model-based analysis of multishell diffusion MR data for tractography: How to get over fitting problems. *Magnetic Resonance in Medicine*, 68(6), 1846–1855. <https://doi.org/10.1002/mrm.24204>
- Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2), 825–841. <https://doi.org/10.1006/nimg.2002.1132>
- Jeurissen, B., Descoteaux, M., Mori, S., & Leemans, A. (2019). Diffusion MRI fiber tractography of the brain. *NMR in Biomedicine*, 32(4), e3785. <https://doi.org/10.1002/nbm.3785>
- Jiang, R., Calhoun, V. D., Fan, L., Zuo, N., Jung, R., Qi, S., Lin, D., Li, J., Zhuo, C., Song, M., Fu, Z., Jiang, T., & Sui, J. (2020). Gender differences in connectome-based predictions of individualized intelligence quotient and sub-domain scores. *Cerebral Cortex*, 30(3), 888–900. <https://doi.org/10.1093/cercor/bhz134>
- Kawahara, J., Brown, C. J., Miller, S. P., Booth, B. G., Chau, V., Grunau, R. E., Zwicker, J. G., & Hamarneh, G. (2017). BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment. *NeuroImage*, 146, 1038–1049. <https://doi.org/10.1016/j.neuroimage.2016.09.046>
- Kay, K. N. (2018). Principles for models of neural information processing. *NeuroImage*, 180, 101–109. <https://doi.org/10.1016/j.neuroimage.2017.08.016>
- Kim, M., Bao, J., Liu, K., Park, B. Y., Park, H., Baik, J. Y., & Shen, L. (2021). A structural enriched functional network: An application to predict brain cognitive performance. *Medical Image Analysis*, 71, 102026. <https://doi.org/10.1016/j.media.2021.102026>
- Kochunov, P., Rowland, L. M., Fieremans, E., Veraart, J., Jahanshad, N., Eskandar, G., Du, X., Muellerklein, F., Savransky, A., Shukla, D., Sampath, H., Thompson, P. M., & Hong, L. E. (2016). Diffusion-weighted imaging uncovers likely sources of processing-speed deficits in schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America*, 113(47), 13504–13509. <https://doi.org/10.1073/pnas.1608246113>
- Li, X., Wang, Y., Wang, W., Huang, W., Chen, K., Xu, K., Zhang, J., Chen, Y., Li, H., Wei, D., Shu, N., & Zhang, Z. (2020). Age-related decline in the topological efficiency of the brain structural connectome and cognitive aging. *Cerebral Cortex*, 30(8), 4651–4661. <https://doi.org/10.1093/cercor/bhaa066>
- Liao, X., Vasilakos, A. V., & He, Y. (2017). Small-world human brain networks: Perspectives and challenges. *Neuroscience and Biobehavioral Reviews*, 77, 286–300. <https://doi.org/10.1016/j.neubiorev.2017.03.018>
- Lin, Y. C., Baete, S. H., Wang, X., & Boada, F. E. (2020). Mapping brain-behavior networks using functional and structural connectome fingerprinting in the HCP dataset. *Brain and Behavior: A Cognitive Neuroscience Perspective*, 10(6), e01647. <https://doi.org/10.1002/brb3.1647>

- Madole, J. W., Ritchie, S. J., Cox, S. R., Buchanan, C. R., Hernandez, M. V., Maniega, S. M., Wardlaw, J. M., Harris, M. A., Bastin, M. E., Deary, I. J., & Tucker-Drob, E. M. (2020). Aging-sensitive networks within the human structural connectome are implicated in late-life cognitive declines. *Biological Psychiatry*, *89*, 795–806. <https://doi.org/10.1016/j.biopsych.2020.06.010>
- Mansour, L. S., Tian, Y., Yeo, B. T. T., Cropley, V., & Zalesky, A. (2021). High-resolution connectomic fingerprints: Mapping neural identity and behavior. *NeuroImage*, *229*, 117695. <https://doi.org/10.1016/j.neuroimage.2020.117695>
- Palop, J. J., & Mucke, L. (2016). Network abnormalities and interneuron dysfunction in Alzheimer disease. *Nature Reviews. Neuroscience*, *17*(12), 777–792. <https://doi.org/10.1038/nrn.2016.141>
- Pervaiz, U., Vidaurre, D., Woolrich, M. W., & Smith, S. M. (2020). Optimising network modelling methods for fMRI. *NeuroImage*, *211*, 116604. <https://doi.org/10.1016/j.neuroimage.2020.116604>
- Qi, S., Calhoun, V. D., van Erp, T. G. M., Bustillo, J., Damaraju, E., Turner, J. A., Du, Y., Yang, J., Chen, J., Yu, Q., Mathalon, D. H., Ford, J. M., Voyvodic, J., Mueller, B. A., Belger, A., McEwen, S., Potkin, S. G., Preda, A., Jiang, T., & Sui, J. (2018). Multimodal fusion with reference: Searching for joint neuromarkers of working memory deficits in schizophrenia. *IEEE Transactions on Medical Imaging*, *37*(1), 93–105. <https://doi.org/10.1109/TMI.2017.2725306>
- Rosenberg, M. D., Casey, B. J., & Holmes, A. J. (2018). Prediction complements explanation in understanding the developing brain. *Nature Communications*, *9*(1), 589. <https://doi.org/10.1038/s41467-018-02887-9>
- Rosenberg, M. D., Finn, E. S., Scheinost, D., Papademetris, X., Shen, X., Constable, R. T., & Chun, M. M. (2016). A neuromarker of sustained attention from whole-brain functional connectivity. *Nature Neuroscience*, *19*(1), 165–171. <https://doi.org/10.1038/nn.4179>
- Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, *52*(3), 1059–1069. <https://doi.org/10.1016/j.neuroimage.2009.10.003>
- Sarwar, T., Ramamohanarao, K., & Zalesky, A. (2019). Mapping connectomes with diffusion MRI: Deterministic or probabilistic tractography? *Magnetic Resonance in Medicine*, *81*(2), 1368–1384. <https://doi.org/10.1002/mrm.27471>
- Scheinost, D., Noble, S., Horien, C., Greene, A. S., Lake, E. M., Salehi, M., Gao, S., Shen, X., O'Connor, D., Barron, D. S., Yip, S. W., Rosenberg, M. D., & Constable, R. T. (2019). Ten simple rules for predictive modeling of individual differences in neuroimaging. *NeuroImage*, *193*, 35–45. <https://doi.org/10.1016/j.neuroimage.2019.02.057>
- Shu, N., Liang, Y., Li, H., Zhang, J., Li, X., Wang, L., He, Y., Wang, Y., & Zhang, Z. (2012). Disrupted topological organization in white matter structural networks in amnesic mild cognitive impairment: Relationship to subtype. *Radiology*, *265*(2), 518–527. <https://doi.org/10.1148/radiol.12112361>
- Sporns, O., Tononi, G., & Kötter, R. (2005). The human connectome: A structural description of the human brain. *PLoS Computational Biology*, *1*(4), e42. <https://doi.org/10.1371/journal.pcbi.0010042>
- Sui, J., Jiang, R., Bustillo, J., & Calhoun, V. (2020). Neuroimaging-based individualized prediction of cognition and behavior for mental disorders and health: Methods and promises. *Biological Psychiatry*, *88*(11), 818–828. <https://doi.org/10.1016/j.biopsych.2020.02.016>
- Sui, J., Qi, S., van Erp, T. G. M., Bustillo, J., Jiang, R., Lin, D., Turner, J. A., Damaraju, E., Mayer, A. R., Cui, Y., Fu, Z., Du, Y., Chen, J., Potkin, S. G., Preda, A., Mathalon, D. H., Ford, J. M., Voyvodic, J., Mueller, B. A., ... Calhoun, V. D. (2018). Multimodal neuromarkers in schizophrenia via cognition-guided MRI fusion. *Nature Communications*, *9*(1), 3028. <https://doi.org/10.1038/s41467-018-05432-w>
- Tolstikhin, I., Hounsfield, N., Kolesnikov, A., Beyers, L., & Dosovitskiy, A. (2021). MLP-mixer: An all-MLP architecture for vision. Tuladhar, A. M., van Dijk, E., Zwiers, M. P., van Norden, A. G., de Laat, K. F., Shumskaya, E., Norris, D. G., & de Leeuw, F. E. (2016). Structural network connectivity and cognition in cerebral small vessel disease. *Human Brain Mapping*, *37*(1), 300–310. <https://doi.org/10.1002/hbm.23032>
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, *15*(1), 273–289. <https://doi.org/10.1006/nimg.2001.0978>
- Wang, J., Wang, X., Xia, M., Liao, X., Evans, A., & He, Y. (2015). GREYNET: A graph theoretical network analysis toolbox for imaging connectomics. *Frontiers in Human Neuroscience*, *9*, 386. <https://doi.org/10.3389/fnhum.2015.00386>
- Webb, C. E., Rodrigue, K. M., Hoagey, D. A., Foster, C. M., & Kennedy, K. M. (2020). Contributions of white matter connectivity and BOLD modulation to cognitive aging: A lifespan structure-function association study. *Cerebral Cortex*, *30*(3), 1649–1661. <https://doi.org/10.1093/cercor/bhz193>
- Woo, C. W., Chang, L. J., Lindquist, M. A., & Wager, T. D. (2017). Building better biomarkers: Brain models in translational neuroimaging. *Nature Neuroscience*, *20*(3), 365–377. <https://doi.org/10.1038/nn.4478>
- Xia, M., Wang, J., & He, Y. (2013). BrainNet viewer: A network visualization tool for human brain connectomics. *PLoS One*, *8*(7), e68910. <https://doi.org/10.1371/journal.pone.0068910>
- Yang, C., Li, X., Zhang, J., Chen, Y., Li, H., Wei, D., Lu, P., Liang, Y., Liu, Z., Shu, N., Wang, F., Guan, Q., Tao, W., Wang, Q., Jia, J., Ai, L., Cui, R., Wang, Y., Peng, D., ... Beijing Aging Brain Rejuvenation Initiative Workgroup. (2021). Early prevention of cognitive impairment in the community population: The Beijing aging brain rejuvenation initiative. *Alzheimers Dementia*, *17*(10), 1610–1618. <https://doi.org/10.1002/alz.12326>
- Yeh, C. H., Jones, D. K., Liang, X., Descoteaux, M., & Connelly, A. (2021). Mapping structural connectivity using diffusion MRI: Challenges and opportunities. *Journal of Magnetic Resonance Imaging*, *53*(6), 1666–1682. <https://doi.org/10.1002/jmri.27188>
- Zalesky, A., Fornito, A., Harding, I. H., Cocchi, L., Yucel, M., Pantelis, C., & Bullmore, E. T. (2010). Whole-brain anatomical networks: Does the choice of nodes matter? *NeuroImage*, *50*(3), 970–983. <https://doi.org/10.1016/j.neuroimage.2009.12.027>
- Zhong, S., He, Y., & Gong, G. (2015). Convergence and divergence across construction methods for human brain white matter networks: An assessment based on individual differences. *Human Brain Mapping*, *36*(5), 1995–2013. <https://doi.org/10.1002/hbm.22751>
- Zimmermann, J., Griffiths, J. D., & McIntosh, A. R. (2018). Unique mapping of structural and functional connectivity on cognition. *The Journal of Neuroscience*, *38*(45), 9658–9667. <https://doi.org/10.1523/JNEUROSCI.0900-18.2018>

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Feng, G., Wang, Y., Huang, W., Chen, H., Dai, Z., Ma, G., Li, X., Zhang, Z., & Shu, N. (2022). Methodological evaluation of individual cognitive prediction based on the brain white matter structural connectome. *Human Brain Mapping*, *43*(12), 3775–3791. <https://doi.org/10.1002/hbm.25883>