A multivariate pattern analysis study of the HIV-related white matter anatomical structural connections alterations

Zhenchao Tang[1†], Zhenyu Liu[2†], Ruili Li[3], Xinwei Cui[4], Hongjun Li[*3] Enqing Dong[*1], Jie Tian[*2]

1 School of Mechanical, Electrical & Information Engineering, Shandong University, Weihai, Shandong Province, 264209, China

2 Key Laboratory of Molecular Imaging, Chinese Academy of Sciences, Beijing, 100190, China

3 Beijing YouAn Hospital, Capital Medical University, Beijing, 100069, China

4 Cooperative Innovation Center of Internet Healthcare, Zhengzhou University, Zhengzhou, China 450052

Corresponding Authors:

Jie Tian, E-mail: tian@ieee.org, TEL: 86-10-82628760;

Hongjun Li, E-mail: lihongjun00113@126.com, TEL: 86-10-62527995;

Enqing Dong, E-mail: enqdong@sdu.edu.cn, TEL: 86-0631-5688512;

†Equal contributors: Zhenchao Tang and Zhenyu Liu;

*Common corresponding authors: Jie Tian, Hongjun Li and Enqing Dong.

Abstract

It's widely known that HIV infection would cause white matter integrity impairments. Nevertheless, it is still unclear that how the white matter anatomical structural connections are affected by HIV infection. In the current study, we employed a multivariate pattern analysis to explore the HIV-related white matter connections alterations. Forty antiretroviral-therapy-naïve HIV patients and thirty healthy controls were enrolled. Firstly, an Automatic Anatomical Label (AAL) atlas based white matter structural network, a 90 × 90 FA-weighted matrix, was constructed for each subject. Then, the white matter connections deprived from the structural network were entered into a lasso-logistic regression model to perform HIV-control group classification. Using leave one out cross validation, a classification accuracy (ACC) of 90% (P=0.002) and areas under the receiver operating characteristic curve (AUC) of 0.96 was obtained by the classification model. This result indicated that the white matter anatomical structural connections contributed greatly to HIV-control group classification, providing solid evidence that the white matter connections were affected by HIV infection. Specially, 11 white matter connections were selected in the classification model, mainly crossing the regions of frontal lobe, Cingulum, Hippocampus, and Thalamus, which were reported to be damaged in previous HIV studies. This might suggest that the white matter connections adjacent to the HIV-related impaired regions were prone to be damaged.

Description of purpose

In current study, we aimed to employ lasso-logistic regression model to select the white matter connections and use them as features to perform HIV-control group classification. By investigating the selected features, we hoped to explore the HIV-related white matter connections alterations.

Methods

Subjects and Data acquisition

This study was approved by the Beijing You'an Hospital of Capital Medical University. All participants had provided written informed consent. Forty antiretroviral-therapy-naïve HIV patients and thirty healthy controls were enrolled in the current study. All the HIV patients were naïve to antiretroviral therapy, and the HIV associated Dementia (HAD) stage was 0. The current CD4 cell count was collected for every patients. Demographic and clinical information of HIV patients and healthy controls was provided in Table 1.

Table1. Demographic and clinical information of the subject enrolled in current study

	HIV patients	Healthy control	P-value
Age (Year)	34.0±8.7	32.2±7.5	0.374
Gender (Male/Female)	31/9	17/13	0.063
Current CD4 Count (Cell/μL)	227.3±190.1	-	-
HAD stage	0	-	-

Age and Current CD4 Count were presented as mean ± standard deviation. P-values were calculated by independent-samples T test for age and Chi-square test for Sex. HAD: HIV associated Dementia.

All MRI were obtained on a 3.0T Siemens scanner at the Beijing You'an Hospital of Capital Medical University. T1-weighted images were acquired with a spoiled gradient recall sequence. The parameters were: repetition time (TR) = 1900ms; echo time (TE) = 2.52ms; field of view (FOV) = 250 × 250mm; acquisition matrix = 256 × 246; number of slices = 176; slice thickness = 1mm; flip angle = 9°. Scan time was 4 min 18s. For DTI, we used single shot echo-planar imaging (EPI) sequence in contiguous axial planes covering the whole brain. Diffusion weighted images (DWI) were acquired along 20 non-collinear and non-coplanar directions with b = 1000 s/mm^2 and one b = 0 s/mm^2 image. The parameters were: TR/TE = 3300/90ms; FOV = 230 × 230 mm; acquisition matrix = 128 × 128; flip angle=90°; number of slices = 25; slice thickness = 4mm. Scan time was 3 min 39s.

Data preprocessing and White matter network construction

We employed the FMRIB Software Library (FSL, FMRIB, Oxford, UK) to preprocess the DTI data. Firstly, the EddyCorrect Tool from FSL was used to correct for head motion and eddy current distortions by aligning all raw DWI volumes to the b0 image. Secondly, the brain mask was extracted from the b0 image by the Brain Extraction Tool (BET) provided in FSL. Then, a diffusion tensor, or ellipsoid, was model at each voxel by Dtifit Tool. Based on the eigenvalues of the tensor, Fractional Anisotropy (FA), Mean Diffusivity (MD), Axial Diffusivity (AD) and Radius Diffusivity (RD) values were calculated on a voxel by voxel basis.

White matter network was constructed by Diffusion toolkit (http://trackvis.org/dtk/) and PANDA (http://www.nitrc.org/projects/panda/). Firstly, the whole brain deterministic tractography were reconstructed employing the Fiber Assignment by Continuous Tracking (FACT) algorithm, which was incorporated in the Diffusion toolkit. Fiber tracking was terminated where the FA<0.2 or the angle between two consecutive eigenvectors orientations was greater than 45°. Subsequently, the white matter structural network was constructed with PANDA based on the Automatic Anatomical Label (AAL) atlas, which parcellated the whole brain into 90 regions. These regions were taken as the nodes of the network, and the white matter streamlines between the regions was taken as the edges of the network. Any pair of nodes were considered anatomically connected if at least three white matter streamlines were found between them. The mean Fractional Anisotropy (FA) values along the white matter streamlines were taken as the edge weight. Thus, we constructed a FA-weighted brain structural network, which was a 90 × 90 matrix, for each subject. For the reason that the white matter connection matrix was symmetric, the matrix actually provided 4005 FA-weighted connections. These FA-weighted connections were used as features in the following HIV-control group classification.

HIV-control group classification

Based on the FA-weighted connections obtained above, we employed lasso-logistic regression to perform the HIV-control group classification. The HIV patients were labeled as 1, and the healthy controls were labeled as 0. The 4005 FA-weighted connections together with gender (Male was taken as 1, and female as 0) and age were fed into the classification model to make prediction of the labels. Gender and age were added in the classification model to exclude the potential impact of age or gender

difference between two groups. Glmnet (http://statweb.stanford.edu/~tibs/lasso.html) was used to construct the classification model. In the training stage, leave one out cross validation was used to estimation the performance of classification. The regularization parameter λ in lasso-logistic regression was set as: the minimal value of λ was 0.1 and the number of λ sequence was 100. The λ minimized the misclassification error was selected as the optimal regularization parameter. In the validation stage, the classification model was constructed with the optimal regularization parameter. After the receiver operating characteristic curve (ROC) curve was plotted, the area under the ROC (AUC) were calculated to evaluate the classification performance. The classification accuracy (ACC), which was defined as the percent of the accurately classified samples, was also obtained to testify the effectiveness of the classification model.

Permutation analysis

By means of Permutation analysis, we estimated how likely we were to observe the same classification performance by chance. We permutated the labels of the subjects randomly for 500 times, and carried out the classification process with each set of permuted labels. The P-values of the reported classification accuracy were defined as: $P = (1+N_{Better\ Prediction}) / (1+N)$. N was the number of permutation; $N_{Better\ Prediction}$ was the number of greater ACC observed in all the permutations.

Results

The multivariate pattern analysis revealed obvious white matter anatomical structural connections alterations between the HIV patients and Healthy controls. After lasso-logistic regression, 11 white matter anatomical structural connections were selected in the classification model (Details of the connections can be found in Table2). The white matter anatomical structural connections selected in the classification model greatly contributed to the HIV-control classification. In other word, HIV infection caused obvious white matter connection alterations.

A classification accuracy of 90% (P=0.002) and AUC of 0.96 was obtained by the classification model. The ROC curve was displayed in Fig.1

Table2. White matter anatomical structural connections selected in the classification model

Feature Number	ROI1	ROI2	β
47	Frontal_Mid_Orb_R	Frontal_Mid_R	-0.158
157	Supp_Motor_Area_L	Precentral_L	0.953
198	Olfactory_L	Frontal_Sup_Orb_L	1.327
254	Frontal_Sup_Medial_L	Supp_Motor_Area_R	-0.230
285	Frontal_Med_Orb_L	Frontal_Sup_Orb_R	-0.020
564	Cingulum_Mid_R	Cingulum_Mid_L	-0.559
861	Amygdala_R	Hippocampus_R	1.311
1766	Parietal_Sup_R	Occipital_Mid_R	-0.797
2622	Putamen_L	SupraMarginal_L	0.118
2939	Thalamus_R	Frontal_Mid_Orb_R	-0.607
3657	Temporal_Mid_R	Temporal_Pole_Sup_R	-0.775

Feature Number was the sequence number of the selected white matter anatomical structural connection in all the 4007 features. ROI1 and ROI2 were the end points of the respective white matter anatomical structural connections. β was the model parameter of the corresponding features in the classification model.

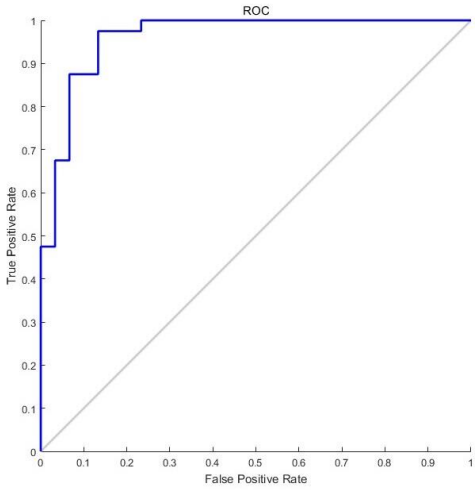

Fig 1. The ROC curve of HIV-Control group classifiction (AUC=0.96, ACC=90%)

New or breakthrough work to be presented

In the current study, we employed lasso-logistic regression model to perform HIV-control group classification. White matter structural connections were entered as feature to the classification model. The classification model achieved relatively satisfying performance, providing solid evidence that the white matter connections were affected by HIV infection. The regions that the altered white matter connections crossed were consistent with previous studies, indicating the effectiveness of the multivariate pattern analysis method.

Conclusions

The multivariate pattern analysis revealed obvious white matter connections alterations in HIV infection. The altered white matter connections mainly crossed the regions which had been reported to be impaired in previous HIV study, suggesting that the white matter connections adjacent to the HIV-related impaired regions were prone to be damaged.

Announcement

This work is not being and has not been submitted for any publication or presentation elsewhere.