## Supplementary Material

### Methods and materials

### Quality assessment of published DTI studies

Two of the authorsindependently rated each included study for quality and completeness using a 12-point checklist adapted from previous published meta-analyses[1], modified to reflect critical variables important in assessing DTI studies[2, 3]. The 12-point checklist was divided into 3 categories: participants (items 1–4), methods for image acquisition and analysis (items 5–10), and results and conclusions (items 11 and 12) (**Table S1**). Each item received a score of 1, 0.5 or 0 if the criteria were fully, partially or not met, respectively.

### Voxel-wise meta-analysis of published DTI data

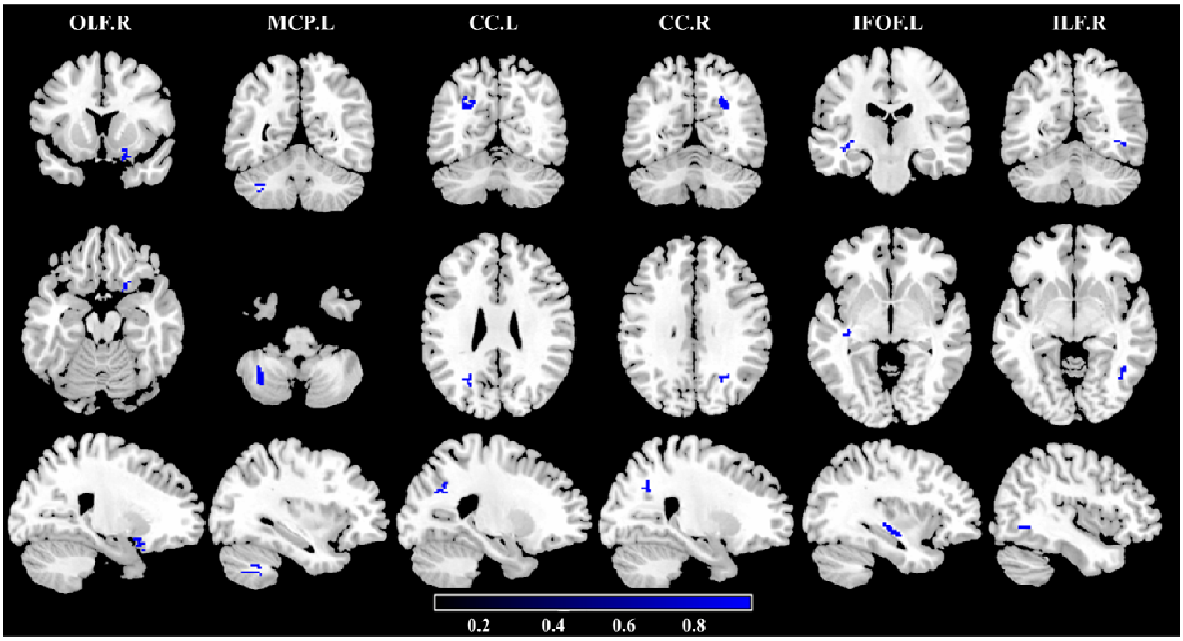
We analyzed FA differences in WM between patients with PD and HC using AES-SDM, a voxel-based meta-analytic approach that uses the reported peak coordinates to recreate original maps of the effect size of FA differences in WM between patients and controls, rather than just assessing the probability or likelihood of a peak[4]. The AES-SDM method incorporates the positive features of earlier methods (e.g. activation likelihood estimation). It reconstructs positive and negative differences in the same signed differential map to avoid any voxel appearing significant in opposite directions and usesthe effect size to combine reported peak coordinates with statistical parametric maps [5]. Complementary analyses, such as jackknife sensitivity, subgroup and meta-regression analyses, can help assess the robustness and heterogeneity of the results. The kernels of the previous versions of ES-SDM are isotropic (i.e. the effect size of a voxel close to a peak would depend only on the effect size of the peak and the Euclidian distance between the voxel and the peak); AES-SDM introduces a novel improvement: it adopts anisotropic kernels to assign different values to different neighboring voxels based on the spatial correlation between them (i.e. highly correlated voxels are estimated to have larger effect sizes, while uncorrelated voxels are estimated to have smaller or null effect sizes) [6]. The basic assumption is that correlated voxels are more likely to be from the same brain region. The recreation of effect size maps using anisotropic kernels can be more accurate than using isotropic kernels, thus allowing more complete and accurate meta-analyses.

**Table S1.** Imaging Methodology Quality Assessment Checklist (Compiled from 1 to 3).

| |
|---|
| **Category 1: Participants** |
| 1. Patients were evaluated prospectively, specific diagnostic criteria were applied, and demographic data were reported |
| 2. Healthy comparison participants were evaluated prospectively; psychiatric and medical illnesses were excluded |
| 3. Important variables (e.g. age, gender, illness duration, H&Y stage, UPDRS scores, medication status, LEDD) were checked either via stratification or statistics |
| 4. Sample size per group: $\geq 20$ scores 1, $\geq 10$ scores 0.5 |
| **Category 2: Methods for image acquisition and analysis** |
| 5. Magnet field strength: 3T scores 1, 1.5T scores 0.5 |
| 6. Number of diffusion gradient directions: $\geq 20$ scores 1, $\geq 12$ scores 0.5 |
| 7. b-value ($s/mm^2$): $\geq 1{,}000$ scores 1, $< 1{,}000$ scores 0.5 |
| 8. Whole brain analysis was automated without a previously defined region |
| 9. Coordinates of decreased or increased FA reported in a standard space |
| 10. Analysis pipeline and measurements were clearly enough described that they could be reproduced. |
| **Category 3: Results and conclusions** |
| 11. Statistical results were corrected for multiple comparison scores 1, uncorrected scores 0.5 |
| 12. Conclusions were consistent with the results obtained, and the limitations were discussed |

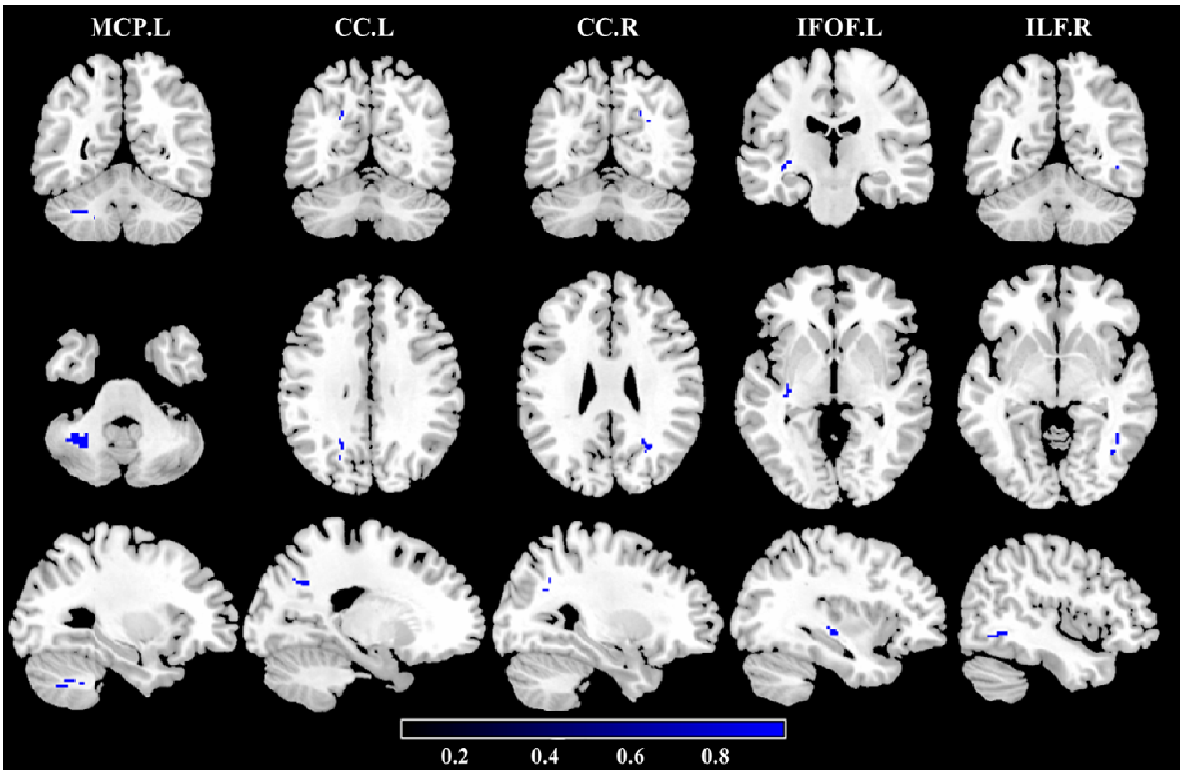
**Key:** Score: 0, criterion not met; 0.5, criterion partially met; 1, criterion fully met.

**Figure S1.** Regions showing decreased FA in the subgroup analysis by medication status.



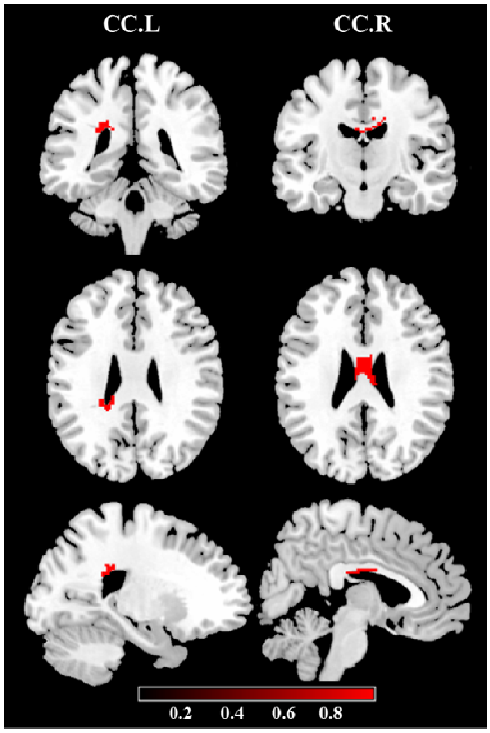

The subgroup analysis of medication-free patients showed decreased FA in the white matter of the right OLF(the leftmost column). The subgroup analysis of medicated patients showed decreased FA in the MCP, CC, IFOFand ILF(the subsequent columns labeled left to right). Abbreviation: CC, corpus callosum; FA, fractional anisotropy; IFOF, inferior fronto-occipital fasciculus; ILF, inferior longitudinal fasciculus; L, Left; MCP, middle cerebellar peduncles; OLF, olfactory cortex; R, right.

**Figure S2.** Regions showing decreased FA in the subgroup analysis of medicated patients with LEDD < 400 mg/day.



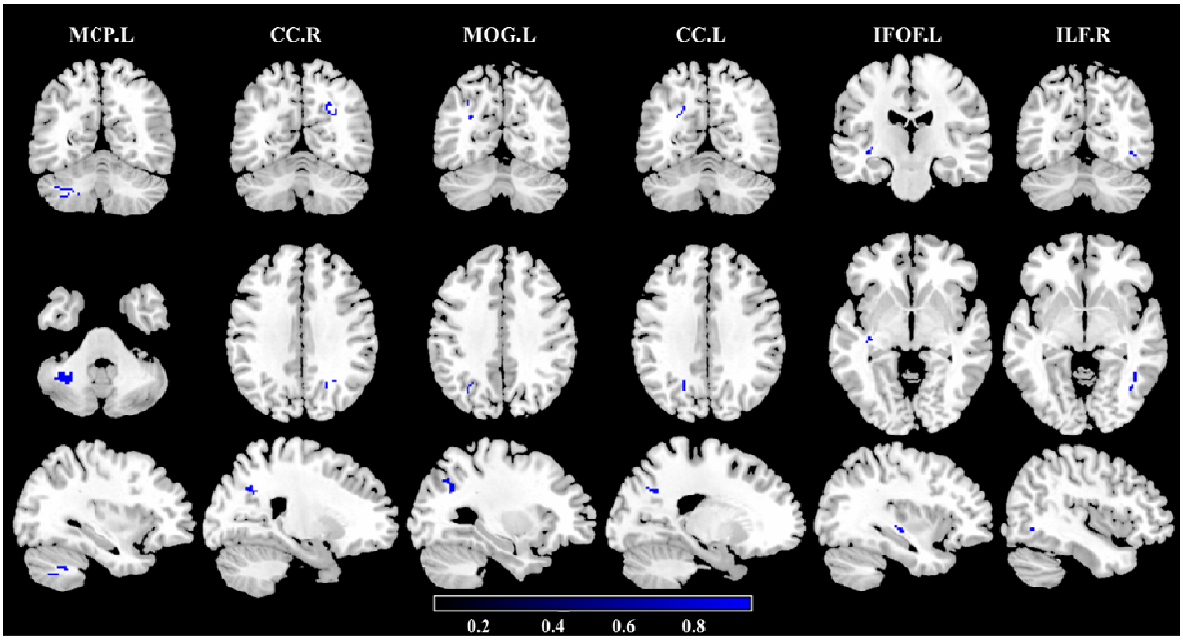

The subgroup analysis of the medicated patients with LEDD < 400 mg/day showed decreased FA in the MCP, CC, IFOF and ILF. Abbreviation: CC, corpus callosum; FA, fractional anisotropy; IFOF, inferior fronto-occipital fasciculus; ILF, inferior longitudinal fasciculus; L, Left; LEDD, levodopa equivalent daily dose; MCP, middle cerebellar peduncles; R, right.

**Figure S3.** Regions showing increased FA in the subgroup analysis of medicated patients with LEDD ≥ 400 mg/day.



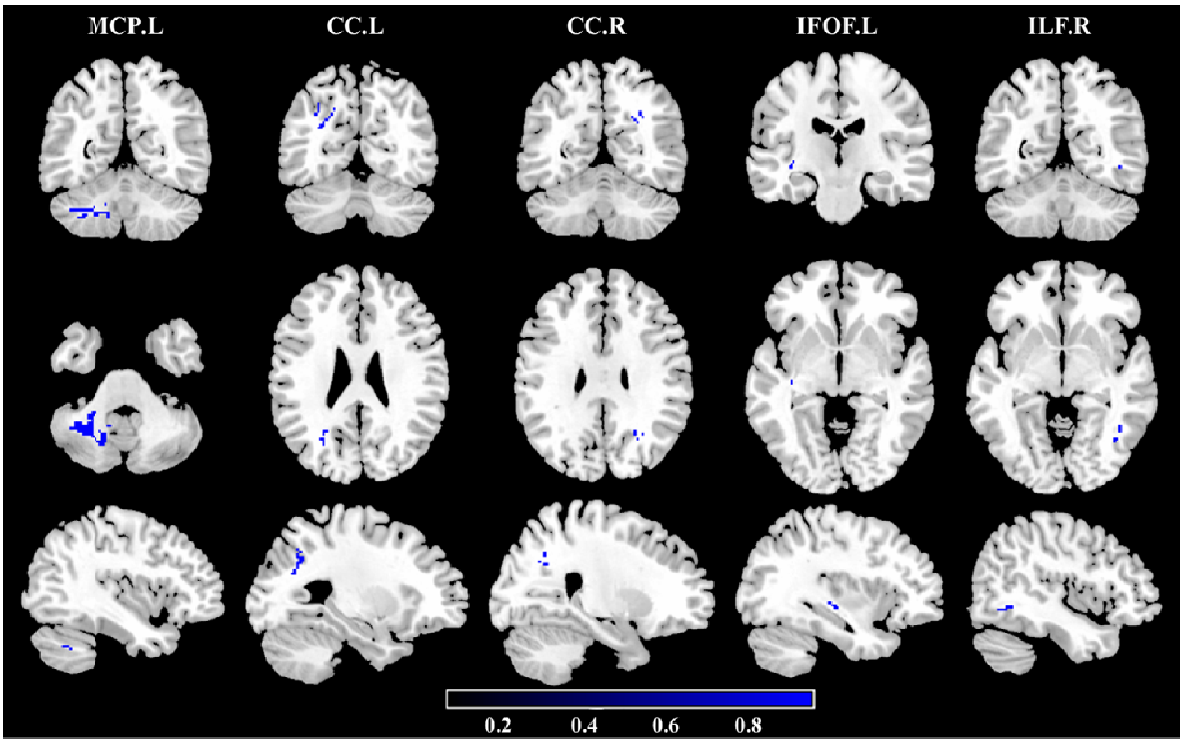

The subgroup analysis of the medicated patients with LEDD ≥ 400 mg/day showed increased FA in the CC. Abbreviation: CC, corpus callosum; FA, fractional anisotropy; L, Left; LEDD, levodopa equivalent daily dose; R, right.

**Figure S4.** Regions showing decreased FA in the subgroup analysis of early stage patients.



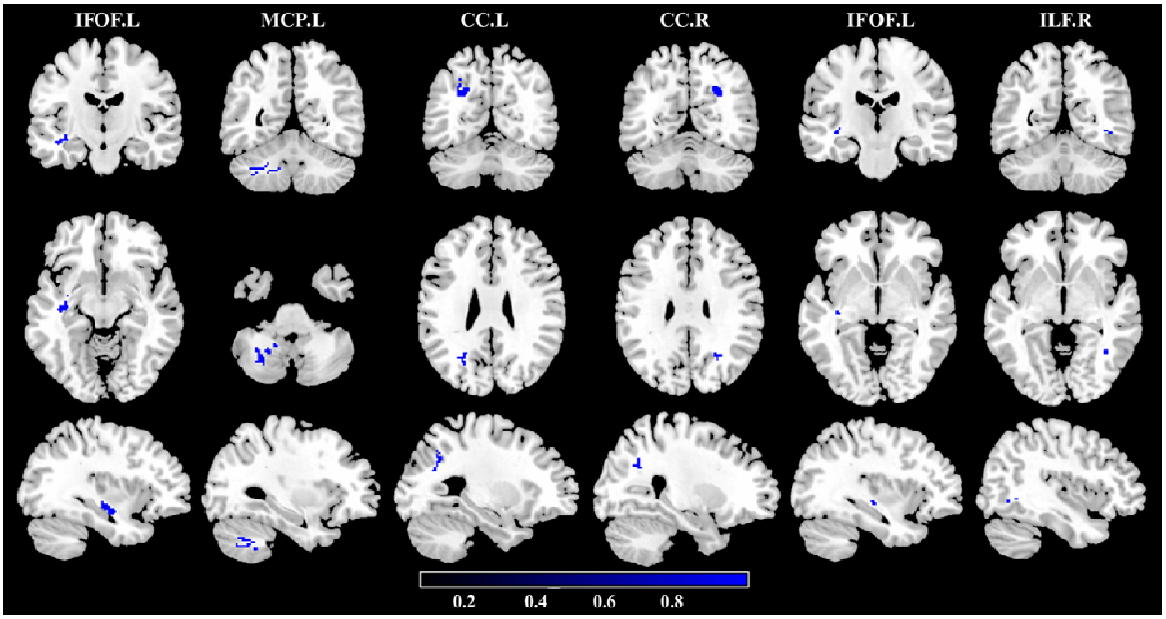

The subgroup analysis of the early stage patients showed decreased FA in the MCP, CC, MOG, IFOF and ILF. Abbreviation: CC, corpus callosum; FA, fractional anisotropy; IFOF, inferior fronto-occipital fasciculus; ILF, inferior longitudinal fasciculus; L, Left; MCP, middle cerebellar peduncles; MOG, middle occipital gyrus; R, right.

**Figure S5.** Regions showing decreased FA in the subgroup analysis of studies using a voxel-based analysis method.




The subgroup analysis of the studies using a voxel-based analysis method showed decreased FA in the MCP, CC, IFOF and ILF. Abbreviation: CC, corpus callosum; FA, fractional anisotropy; IFOF, inferior fronto-occipital fasciculus; ILF, inferior longitudinal fasciculus; L, Left; MCP, middle cerebellar peduncles; R, right.

**Figure S6.** Regions showing decreased FA in subgroup analysis by the number of diffusion directions.




The subgroup analysis of ≥ 30 diffusion directions showed decreased FA in left IFOF (the leftmost column). The subgroup analysis of < 30 diffusion directions showed decreased FA in the MCP, CC, IFOF and ILF (the subsequent columns labeled left to right). Abbreviations: CC, corpus callosum; FA, fractional anisotropy; IFOF, inferior fronto-occipital fasciculus; ILF, inferior longitudinal fasciculus; L, Left; MCP, middle cerebellar peduncles; R, right.

**References**

1. Shepherd AM, Matheson SL, Laurens KR, Carr VJ, Green MJ. Systematic meta-analysis of insula volume in schizophrenia. Biol Psychiatry 2012;72(9):775-784.
2. Chen L, Hu X, Ouyang L, He N, Liao Y, Liu Q, Zhou M, Wu M, Huang X, Gong Q. A systematic review and meta-analysis of tract-based spatial statistics studies regarding attention-deficit/hyperactivity disorder. Neurosci Biobehav Rev 2016;68:838-847.
3. Jiang J, Zhao YJ, Hu XY, Du MY, Chen ZQ, Wu M, Li KM, Zhu HY, Kumar P, Gong QY. Microstructural brain abnormalities in medication-free patients with major depressive disorder: a systematic review and meta-analysis of diffusion tensor imaging. J Psychiatry Neurosci 2017;42(3):150-163.
4. Radua J, Via E, Catani M, Mataix-Cols D. Voxel-based meta-analysis of regional white-matter volume differences in autism spectrum disorder versus healthy controls. Psychol Med 2011;41(7):1539-1550.
5. Radua J, Mataix-Cols D, Phillips ML, El-Hage W, Kronhaus DM, Cardoner N, Surguladze S. A new meta-analytic method for neuroimaging studies that combines reported peak coordinates and statistical parametric maps. Eur Psychiatry 2012;27(8):605-611.
6. Radua J, Rubia K, Canales-Rodriguez EJ, Pomarol-Clotet E, Fusar-Poli P, Mataix-Cols D. Anisotropic kernels for coordinate-based meta-analyses of neuroimaging studies. Front Psychiatry 2014;5:13.