



The SACT Template: A Human Brain Diffusion Tensor Template for School-age Children

Congying Chu^{1,7} · Haoran Guan¹ · Sangma Xie⁸ · Yanpei Wang¹ · Jie Luo¹ · Gai Zhao¹ · Zhiying Pan¹ · Mingming Hu¹ · Weiwei Men^{3,4} · Shuping Tan⁶ · Jia-Hong Gao^{3,4,5} · Shaozheng Qin^{1,2} · Yong He^{1,2} · Lingzhong Fan⁹ · Qi Dong¹ · Sha Tao¹

Received: 2 June 2021 / Accepted: 22 October 2021 / Published online: 29 January 2022
© Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences 2022

Abstract School-age children are in a specific development stage corresponding to juvenility, when the white matter of the brain experiences ongoing maturation. Diffusion-weighted magnetic resonance imaging (DWI), especially diffusion tensor imaging (DTI), is extensively used to characterize the maturation by assessing white matter properties *in vivo*. In the analysis of DWI data, spatial normalization is crucial for conducting inter-subject analyses or linking the individual space with the reference space. Using tensor-based registration with an appropriate diffusion tensor template presents high accuracy regarding spatial normalization. However, there is a lack of a standardized diffusion tensor template dedicated to school-age children with ongoing brain development. Here, we established the school-age children diffusion tensor

(SACT) template by optimizing tensor reorientation on high-quality DTI data from a large sample of cognitively normal participants aged 6–12 years. With an age-balanced design, the SACT template represented the entire age range well by showing high similarity to the age-specific templates. Compared with the tensor template of adults, the SACT template revealed significantly higher spatial normalization accuracy and inter-subject coherence upon evaluation of subjects in two different datasets of school-age children. A practical application regarding the age associations with the normalized DTI-derived data was conducted to further compare the SACT template and the adult template. Although similar spatial patterns were found, the SACT template showed significant effects on the distributions of the statistical results, which may be related to the performance of spatial normalization. Looking forward, the SACT template could contribute to future studies of white matter development in both healthy and

Congying Chu and Haoran Guan these authors contributed equally to this work.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12264-022-00820-1>.

✉ Sha Tao
taosha@bnu.edu.cn

¹ National Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, China

² Beijing Key Laboratory of Brain Imaging and Connectomics, Beijing Normal University, Beijing 100875, China

³ Center for MRI Research, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China

⁴ Beijing City Key Laboratory for Medical Physics and Engineering, Institute of Heavy Ion Physics, School of Physics, Peking University, Beijing 100871, China

⁵ McGovern Institute for Brain Research, Peking University, Beijing 100871, China

⁶ Psychiatry Research Center, Beijing HuiLongGuan Hospital, Peking University, Beijing 100096, China

⁷ Division of Molecular Neuroimaging, Institute of Neuroscience and Medicine-2, Research Center Jülich, 52425 Jülich, Germany

⁸ Institute of Biomedical Engineering and Instrumentation, School of Automation, Hangzhou Dianzi University, Hangzhou 310018, China

⁹ Brainnetome Center and National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

clinical populations. The SACT template is publicly available now (https://figshare.com/articles/dataset/SACT_template/14071283).

Keywords School-age children · Diffusion-weighted MRI · Diffusion tensor template · Spatial normalization

Introduction

School-age children, typically between the ages of 6 and 12, are in the age period referred to as middle childhood, which represents a distinctive period of human juvenility before the developmental transition to adolescence [1]. Different from children in infancy and early childhood, school-age children experience cognitive development by interacting with systematic education in elementary schools [2]. Meanwhile, the continued white matter maturation during this period heralds remarkable microstructural changes underlying the structural development [3]. Diffusion-weighted magnetic resonance imaging (DWI) techniques enable non-invasive approaches to quantitatively assess white matter microstructure based on water molecule diffusion signals [4]. It has been gaining popularity to conduct group-wise analyses based on modeled DWI data regarding developmental or clinical topics [5–7].

As a commonly used model with DWI data, diffusion tensor imaging (DTI) has been extensively applied to developmental studies because it requires a relatively shorter scan time, making it feasible for children, and has accessible protocols for MRI scanners [3, 6, 8]. Particularly, spatial normalization is an essential step for inter-subject analyses or linking the native space with the standard space regarding DTI data processing [9]. The registration algorithm is the key element for spatial normalization. There are two main kinds of registration algorithms feasible for DTI data: the scalar-based algorithm and the tensor-based algorithm. Unlike scalar-based registration using scalar features such as fractional anisotropy (FA) or b_0 intensity, tensor-based registration leverages full tensor information to achieve more accurate alignments in white matter [10–13]. Moreover, tensor-based registration is also useful for the spatial normalization of other types of DWI data, such as high angular resolution diffusion imaging [11] and neurite orientation dispersion and density imaging [14], by applying the tensor-based deformation field to the corresponding data. To take advantage of tensor-based registration algorithms for developmental and clinical studies in school-age children, a diffusion tensor (DT) template, serving as the optimization target of the registration, is required.

A few DT templates for adults have been developed to date [15–20]. However, several concerns still exist regarding

the scenario of using a DT template in school-age children. First, with the development of the brain, a number of anatomical properties of white matter differ significantly between school-age children and adults [3]. As found with T1-weighted templates, it is not optimal to directly apply adult templates to pediatric data [21, 22]. Therefore, it would be worth constructing a DT template specifically based on the data of school-age children. Second, regarding the construction of the template, a large sample of subjects would be expected to reduce template variability [23]. Third, it would be informative and convenient for practical applications to provide a DT template in standard space. This is related to the difference between the standardized template and a study-specific template. It has been demonstrated that a DT template carefully established in standard space outperforms a study-specific template in the spatial normalization of adult data [19, 24]. Meanwhile, an available standardized template not only saves the cost of time for constructing a study-specific template each time but also provides a reference space to boost correspondence across MRI modalities or studies. Finally, to sustain the anatomical consistency of white matter, a DT template should be constructed in terms of the optimization of tensor reorientation rather than using a surrogate deformation field [25, 26].

By considering the concerns described above, we aimed to develop a SACT template dedicated to school-age children by using a large sample of DTI data (380 participants). After data quality control and preprocessing, we constructed the SACT template with tensor reorientation optimization by using Diffusion Tensor Imaging Toolkit (DTI-TK) [27, 28] in an age-balanced way. Moreover, the SACT template was mapped to the standard space represented by the Chinese pediatric (CHN-PD) atlases [29], which were previously established based on the same dataset as the current study. We further demonstrated the performance of the SACT template based on two different datasets of school-age children regarding the accuracy of spatial normalization and the statistical analyses of age associations with DT-derived metrics. To the best of our knowledge, this is the first attempt to establish a standardized DT template dedicated to school-age children. Combined with the CHN-PD atlases, the SACT template contributes to studies of white matter properties in middle childhood.

Materials and Methods

Participants

In this study, we included three datasets covering the full age range of school-age children (6–12 years) to construct

and evaluate the SACT template. Dataset 1 was the exploratory data pool for constructing the template. Both datasets 2 and 3 were the test data pools for validating the newly-constructed template, especially the spatial normalization accuracy. All three datasets focused on typically-developing school-age children who did not meet any of the following criteria: (1) intellectual or developmental disabilities, (2) history of neurological disorders, (3) psychoactive drug use, (4) having a brain injury, or (5) ever losing consciousness due to head injury. No sedatives were used during the MRI scans of the three datasets. A mock session, simulating an actual MRI scan by using a decommissioned MRI scanner and recordings of the noise of MRI operation, was used in the recruitment of datasets 1 and 2 to acclimate each child to the MRI environment. Specifically, after quality control (26 children were excluded according to the criteria described below), dataset 1 consisted of data from 380 participants aged 6–12 years (mean age \pm SD, 9.09 ± 1.36 years, 201 boys/179 girls) with neuroimaging data collected at Peking University (PKU dataset). After quality control (14 children were excluded), dataset 2 contained data from 115 children aged 6–12 years (9.09 ± 1.34 years, 69 boys/46 girls), whose neuroimaging data were acquired at Beijing HuiLongGuan Hospital (HLG dataset). Both the PKU dataset and the HLG dataset were obtained from the Children School Functions and Brain Development project (CBD, Beijing Cohort) [30], which aims to establish a large longitudinal cohort of school-age children with comprehensive assessments, such as multimodal neuroimaging data, cognitive measurements, and academic achievements. All the recruited children were cognitively normal based on the assessment battery of the CBD project [31]. All the children's legal guardians provided written informed consent approved by the Ethics Committee of Beijing Normal University (approval number, IRB_A_0004_2019001). Dataset 3 had the data of 61 participants aged 6–12 years (9.29 ± 1.38 years, 31 boys/30 girls), who were selected from the baseline data of the Enhanced Nathan Kline Institute Rockland Sample (NKI dataset) [32]. Only subjects with MRI data passing quality control were selected. Written informed consent was given by the participants and their legal guardians following procedures approved by the Nathan Kline Institute Review Board. The demographic information of each dataset is listed in Table 1.

MRI Data Acquisition

For the PKU and HLG datasets, the same type of scanner (3T Siemens MRI Scanner Magnetom Prisma; Erlangen, Germany), was used at both sites with the same scanning parameters for T1-weighted (T1w) and diffusion-weighted

images (DWI). Specifically, the T1w data were acquired using an MPRAGE sequence with the following parameters: repetition time (TR) = 2,530 ms, echo time (TE) = 2.98 ms, flip angle = 7° , field of view (FOV) = 256 mm \times 224 mm, in-plane resolution = 1 mm \times 1 mm, sagittal slices = 192, and slice thickness = 1 mm. The DWI data were acquired using a single-shot 2-dimensional echo-planar imaging (EPI) sequence with the following parameters: TR = 7,500 ms, TE = 64 ms, flip angle = 90° , FOV = 224 mm \times 224 mm, in-plane resolution = 2 mm \times 2 mm, axial slices = 70, slice thickness = 2 mm, phase-encoding direction = posterior $>$ anterior, 64 diffusion volumes weighted with a b -value of 1,000 s/mm² and 10 $b = 0$ s/mm² volumes. Prior to DWI scanning, a corresponding field map with the same spatial resolution was acquired for correction of susceptibility-induced distortion by using a double-echo gradient-echo (GRE) field map sequence: TR = 400 ms, TE1/TE2 = 4.92 ms/7.38 ms and flip angle = 60° . For the NKI datasets, we only used the DWI data, which were collected on a 3T Siemens MRI Scanner Magnetom Trio (Erlangen, Germany). A multi-band EPI sequence [33] was used with the following parameters: TR = 2,400 ms, TE = 85 ms, flip angle = 90° , FOV = 212 mm \times 180 mm, in-plane resolution = 2 mm \times 2 mm, axial slices = 64, slice thickness = 2 mm, phase-encoding direction = anterior $>$ posterior, 128 $b = 1,500$ s/mm² volumes corresponding to diffusion gradient directions, 9 $b = 0$ s/mm² volumes and a multi-band acceleration factor of 4.

Quality Control of MRI Data

All the included T1w and DWI data were evaluated for quality. For the T1w data, four well-trained raters (H.G., Y.W., J.L., and Z.P.) independently performed visual examinations of individual images to determine serious data problems, including obvious brain lesions, lack of slices, and excessive noise. As the visual check was only conducted to detect obvious problems, the data identified as problematic by any one of the raters were excluded. We then applied the Computational Anatomy Toolbox (CAT12, <http://dbm.neuro.uni-jena.de/cat>) to derive the weighted overall image quality considering the inhomogeneities and noise of the T1w data. T1w data with percentage rating points $<75\%$ were excluded. For the DWI data, we applied DTIPrep v1.2.8 [34] for quality control and visual checks. First, the raw Dicom data were converted to Nearly Raw Raster Data file format using the DWIConvert tool of DTIPrep. Second, the converted DWI data were loaded to visually check the uniformity of the diffusion gradient direction on a unit sphere as well as to check for issues including missing slices and heavy distortions. Third, we selected the default protocol of DTIPrep

Table 1 Demographic information of the three datasets.

		Age range (years)	Number	Sex (female/male)
Dataset 1 (PKU) Age groups	Group 1	6–7	25	15/10
	Group 2	7–8	56	28/28
	Group 3	8–9	103	43/60
	Group 4	9–10	98	48/50
	Group 5	10–11	58	26/32
	Group 6	11–12	40	19/21
Dataset 1 Overall		6–12 (9.09 ± 1.36)	380	179/201
Dataset 2 (HLG)		6–12 (9.09 ± 1.34)	115	46/69
Dataset 3 (NKI)		6–12 (9.29 ± 1.38)	61	30/31

The mean and the SD of the age for the whole dataset are included in parentheses (mean ± SD).

to run an automatic quality check, except for setting the tolerance of the percentage of problematic gradients as 0.1. Fourth, a visual check was conducted to confirm the automatic exclusion of problematic directions for each set of DWI data passing the DTIPrep check. The excluded volumes of each set of DWI data were also recorded. In addition, for the PKU and HLG datasets, we considered the participants that had both T1w and DWI data in the following analyses. For the NKI dataset, we only considered the quality of the DWI data.

Data Preprocessing

We mainly used FSL version 6.0 [35] to preprocess the DWI data. First, head motion across DWI volumes and the effects of eddy current were corrected using `eddy_cuda` [36]. In particular, outlier replacement [37] and correction of slice-to-volume movement [38] were considered to compensate for the noise. Second, we corrected distortions in the eddy-corrected DWI data with GRE field maps and corresponding T1w data. Specifically, `optiBET` [39] was applied for brain extraction based on T1w data. After segmenting the extracted brain data with `FAST` [40], we leveraged the tool `epi_reg` of FSL to correct the geometric distortions of the DWI data with the information of the corresponding field map. The corrected data were mapped back to the native space by applying an inversed rigid spatial transformation between the native space and the of T1w data space. As no field maps were available for the NKI dataset, distortion correction was only performed in the PKU and HLG datasets. Finally, after removing the problematic directions recorded from the quality check for each distortion-corrected set of DWI data, the `DTIFIT` tool of FSL was used to calculate voxel-wise DT. All the DT data were finally converted into the format (e.g., diffusivity units) required by DTI-TK.

Constructing the SACT Template

Given the large sample size of the PKU dataset, we used it to establish the SACT template. In the PKU dataset, there were six age groups at one-year intervals, such as the 6–7-year-old group and the 7–8-year-old group. The demographic information of each age group is listed in Tables 1 and S1. Considering the variation in the number of participants across age groups, we randomly sampled the data pool to provide age-balanced datasets, avoiding bias related to specific age groups with more participants. Under the sampling scheme without replacement, we randomly sampled 20 participants of each age group 100 times. After grouping the re-sampled data each time, we subsequently had 100 datasets with 120 participants per dataset.

For each re-sampled dataset, we established the DT template in three sequential steps. Briefly, we first established a population-specific DT template using the DTI-TK software package [25, 27, 28], which uses the full tensor information for registration across subjects. The displacement field from each subject's native space to the population space ($T_{\text{nat} \rightarrow \text{pop}}$) was derived accordingly. Second, after applying the alignment from the T1w data to the DWI data derived from the distortion correction, the T1w data (skull-stripped) were realigned with the DT data in native space. The T1w data were further transformed into population space by using the corresponding $T_{\text{nat} \rightarrow \text{pop}}$. The T1w data were averaged to represent the population space, which were subsequently warped to standard space. Third, we used the CHN-PD atlas [29] as the standard space reference that was situated in Montreal Neurological Institute (MNI) space with high-quality T1w templates established from the same CBD dataset. The $T_{\text{pop} \rightarrow \text{CHN-PD}}$ mapping from the population space to the standard space was derived by using Advanced Normalization Tools (ANTs, version 2.2.0). We combined $T_{\text{nat} \rightarrow \text{pop}}$ and $T_{\text{pop} \rightarrow \text{CHN-PD}}$ to derive the displacement field from native

space to standard space ($T_{nat \rightarrow CHN-PD}$) for each subject. Using the combined displacement field, we directly re-oriented the subject’s DT to CHN-PD space with a single interpolation, avoiding the smoothing effect of multiple interpolations. The subjects’ DTs in the space of the CHN-PD atlas were averaged to derive a DT template in the standard space. A schematic demonstration of the above procedures is shown in Fig. 1 (a complete description of the methodological details is available in the supplementary materials).

In addition to the re-sampled full-age (6–12 years old) dataset, the above procedures were applied to each age group to derive the corresponding age-specific DT template using the full sample of the group. Notably, the age-specific T1w template of the CHN-PD atlas was selected as the standard space reference according to the age group.

Assessment of the Similarity Across DT Templates

To determine how to represent the derived DT templates from the resampled datasets, we assessed the similarity across these DT templates based on tensor and tensor-derived scalar quantities. To measure the tensor-based similarity, we applied Euclidean distance squared (EDS) to measure the dissimilarity between two diffusion tensors.

This was defined as $EDS = \sqrt{\text{trace}((D_1 - D_2)^2)}$, where D_1 and D_2 are diffusion tensors. EDS was then calculated

for the corresponding voxels (i.e., the same coordinates) over each pair of DT templates. For the tensor-derived scalar quantities, we selected two commonly-used quantities, fractional anisotropy (FA) and mean diffusivity (MD), to calculate the voxel-wise absolute difference and the spatial correlation between each pair of FA (or MD) maps derived from the DT templates. If strong similarities across the DT templates were found, the SACT template then were represented by averaging the DT templates of the re-sampled datasets.

Evaluation of the SACT Template for DTI Spatial Normalization

To demonstrate the effect of different DT templates on the DTI spatial normalization of school-age children, we compared the normalization performance between the SACT template and the DT template of the IIT Human Brain Atlas v.5.0, which has been shown to exhibit top performance in DTI spatial normalization [19]. Applying the high-dimensional diffeomorphic registration of DTI-TK to both the HLG and NKI datasets in terms of the two DT templates, several metrics were used to quantitatively assess their performance from three aspects. First, we evaluated the tensor-based voxel-wise similarity between each set of normalized DT data and the template by using the metrics of EDS and deviatoric distance squared (DDS). In particular, DDS is the default similarity metric used by DTI-TK when conducting diffeomorphic registration [27]. DDS is similar to the definition of EDS, except for replacing the tensor as the deviatoric tensor that is defined as $D_{dev} = D - [\text{trace}(D)/3]*I$ (I being the identity tensor). Second, we assessed the voxel-wise volumetric deformation from the template to individual DT data when conducting registration. The deformation was represented by the Jacobian determinant (JD) of the deformation matrix in terms of local deformations (no affine transformations) and global deformations (including affine transformations). Third, we evaluated the inter-subject alignment by calculating the normalized FA SD ($\bar{\sigma}_{FA}$) [25] and the dyadic coherence (dCOH) [41] based on the normalized DT data of the entire group. Both higher values of dCOH and lower values of $\bar{\sigma}_{FA}$ indicate better inter-subject alignment. In addition to the voxel-wise metric maps, we also used the ICBM-DTI-81 white matter labels atlas [16] for comparisons at the regional level. To align the labels atlas with the SACT and the IIT template, we used the `fsl_reg` of FSL with the configuration optimized for FA data to register the corresponding FA map of the labels atlas to the FA maps derived from the SACT template and the IIT template. We further applied an FA-based mask ($FA > 0.25$) to the aligned labels atlas; this was based on the FA maps of the two DT templates. Then, the mean value within each

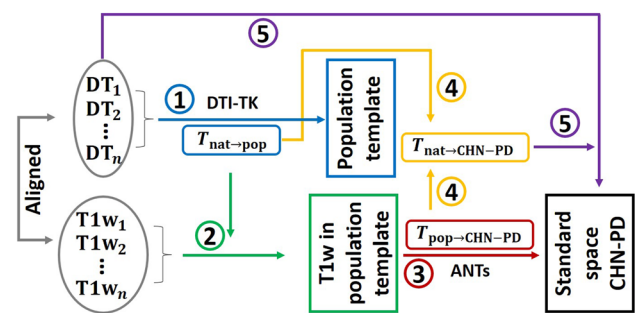


Fig. 1 Schematic of establishment of the diffusion tensor (DT) template in the standard space corresponding to the CHN-PD atlas. All arrows indicate the application of a registration or deformation field. Numbers with circles (1–5) indicate sequential steps. n = number of participants. Step 1 (blue) constructs the population DT template for deducing the deformation $T_{nat \rightarrow pop}$. Step 2 (green) warps the corresponding skull-stripped T1w data to the space of the population DT template. Then, the warped T1w data are averaged. Step 3 (red) spatially normalizes the averaged T1w data to the corresponding T1w template of the CHN-PD atlas. The deformation $T_{pop \rightarrow CHN-PD}$ is deduced accordingly. Step 4 (yellow) calculates the deformation $T_{nat \rightarrow CHN-PD}$ by combining $T_{nat \rightarrow pop}$ and $T_{pop \rightarrow CHN-PD}$. Step 5 (purple) applies the $T_{nat \rightarrow CHN-PD}$ to the corresponding DT data. DTI-TK, Diffusion Tensor Imaging Toolkit; nat, native space; pop, population space; CHN-PD, Chinese pediatric atlases; ANTs, the Advanced Normalization Tools.

region was used for comparisons. In addition, to guarantee a fair comparison, we kept all the parameters of the registration the same (for affine registration: optimization-stop threshold = 0.01, (x, y, z) sep = 4 mm, similarity metric = EDS; for diffeomorphic registration: optimization-stop threshold = 0.002, number of iterations = 6), except the target template was changed accordingly.

Evaluation of the Age Associations of DTI Metrics Based on the SACT Template

To evaluate a practical application of the SACT template, we took the associations between chronological ages and DTI metrics (FA and MD) as examples regarding the use of different DT templates, i.e., the SACT template and the IIT template. Specifically, we applied voxel-based analysis (VBA) [42] and tract-based spatial statistics (TBSS) analysis [43] to the DT data spatially normalized by using DTI-TK. To increase statistical power, we considered the HLG and NKI data as one group. With a general linear model (GLM) to test the associations, we added gender and site information (both dummy coded) of participants as covariates. For the VBA, we further included the deformation measured by the corresponding JD as another metric to evaluate the age association. For the TBSS analysis, we skeletonized the mean FA maps with a threshold of 0.2 and further applied randomization [44] with the TFCE statistic [45] to determine the significance with 5,000 permutations. In addition, for the TBSS analysis, we aligned the lower cingulum mask of the FMRIB58_FA template with the DT templates by non-linear registration of the corresponding FA maps using `fs_l_reg`, and these were used to properly project the data onto the extracted skeleton of the lower cingulum [43].

Statistical Approaches for the Comparisons

For both the voxel-wise metrics (measuring the similarity, deformation, and coherence between individual data and template, such as DDS, JD, and dCOH) and the voxel-wise statistical output from the VBA and the TBSS assessment, we compared the spatial distributions of the voxel-wise values based on the two templates using a two-sample Kolmogorov–Smirnov (KS) test to determine whether they followed the same continuous distribution. Given the spatial difference between the two templates, a direct voxel-wise comparison was not feasible, so we considered the white matter regions (ICBM-DTI-81 white matter labels atlas) instead. Specifically, regarding the metrics of each white matter region (such as the EDS, DDS, JD, and dCOH), we averaged the region-level values to deduce a whole-brain mean value for each subject. Based on the type of template used for the registration, we defined two groups

regarding the same subjects. The paired Wilcoxon signed-rank test was used to determine the difference between the two groups in terms of the deduced mean value of each metric.

Results

Assessment of Similarity Across the DT Templates of the Re-sampled Datasets

The first set of analyses examined the similarity across the DT templates, aiming to evaluate the sensitivity of the SACT template to the underlying samples. We found high similarity across the DT templates derived from different re-sampled datasets (Fig. 2A). Specifically, the voxel-wise mean EDS between each pair of DT templates ranged from 1.1×10^{-6} to 2.1×10^{-6} mm²/s which was averaged within the entire brain. The spatial correlations across the DT templates in terms of FA and MD were also high (the minimal Pearson correlation coefficient for FA was 0.9969 and for MD was 0.9967; Fig. 2A). Similarly, the mean of voxel-wise absolute differences in terms of FA and MD were trivial between pairs of DT templates (for FA: from 6.1×10^{-3} to 7.7×10^{-3} ; for MD: from 2.7×10^{-6} to 4.1×10^{-6} mm²/s; Fig. S2). We also showed that the number of overlapping subjects between any two re-sampled datasets (Fig. 2A) ranged from 33 to 62 with a mean value of 48. The gender ratio (boys/girls) of each resampled dataset ranged from 0.76 to 1.50, with a mean value of 1.03 (Fig. S2). Based on the high consistency found across the re-sampled datasets, we derived the SACT template as the averaged DT template, which was well-aligned with the corresponding CHN-PD T1w template in the MNI space (Fig. 2B). No visible artifacts, such as eddy-current artifacts, were found. Fine white matter details were observed throughout the brain in the SACT template (Fig. 2B). The age-specific SACT templates are shown in Fig. S3, aligning well with the corresponding age-specific T1w templates of the CHN-PD atlases.

Using the SACT Template for Spatial Normalization

Turning now to the experimental evidence on the performance gain of using the SACT template for spatial normalization in school-age children, we organized the experimental results into three parts regarding the evaluation metrics. Meanwhile, we separately assessed the spatial normalization for the HLG and NKI data.

First, we found that the mean value of regional DDS between the normalized DT data and the template was significantly lower for the SACT template than for the IIT template in both the HLG and the NKI data (Fig. 3A, B).

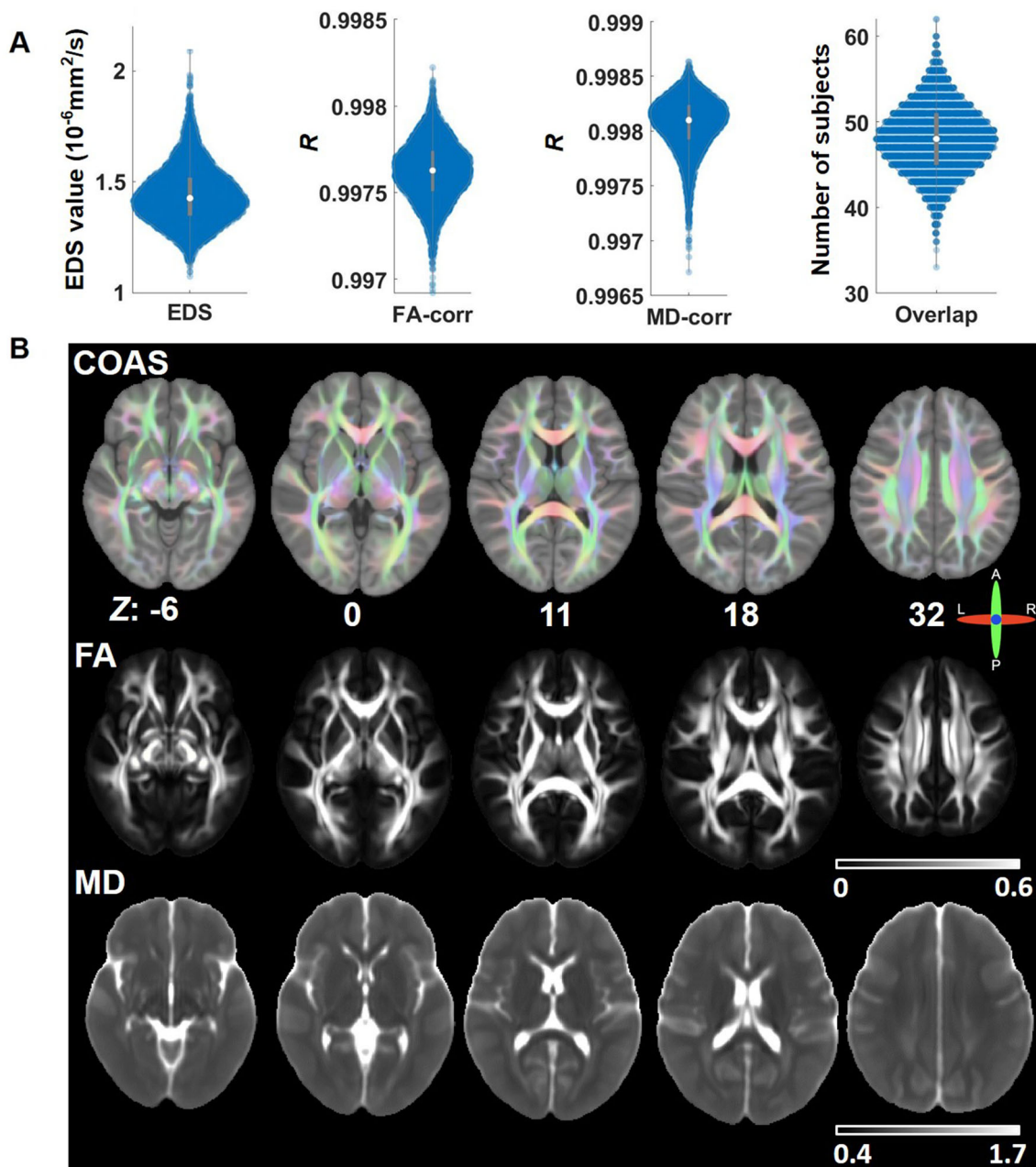


Fig. 2 Evaluation of the SACT template. **A** Evaluation of the similarity across DT templates derived from re-sampled datasets. The violin plot shows the distribution of the data (white circle, median value; grey vertical bar, range from the first to the third quartile). For EDS, each data point is the averaged difference of voxel-wise difference between two DT templates. For FA-corr, each data point is the Pearson correlation coefficient (R) between two FA maps derived from the corresponding DT templates. For MD-corr, each data point

is the Pearson correlation coefficient (R) between two MD maps. For overlap, each data point is the number of overlapped subjects between two resampled datasets. **B** Demonstration of the SACT template. The color orientation and anisotropy scaling (COAS) map is overlapped on the T1w template of the CHN-PD atlas. The numbers below are the MNI axial coordinates (Z) corresponding to the CHN-PD atlas. The diffusivity unit of MD is $10^{-3} \text{mm}^2/\text{s}$. The radiological convention is applied here.

Large effect sizes of the paired Wilcoxon rank tests (0.6137 for the HLG data, 0.6149 for the NKI data) were also found, which were measured by r [46]. A value of $r > 0.5$ represents a large effect size [46]. A lower DDS indicates greater closeness between the normalized data and the template. For the voxel-wise DDS, we checked the

cumulative distributions of the intra-group (HLG and NKI) averaged DDS maps of white matter ($\text{FA} > 0.25$ based on the FA map of the corresponding template). For both the HLG and NKI data, the cumulative distributions significantly differed between the use of the SACT template and the IIT template (for the HLG data: $P < 0.001$, the effect

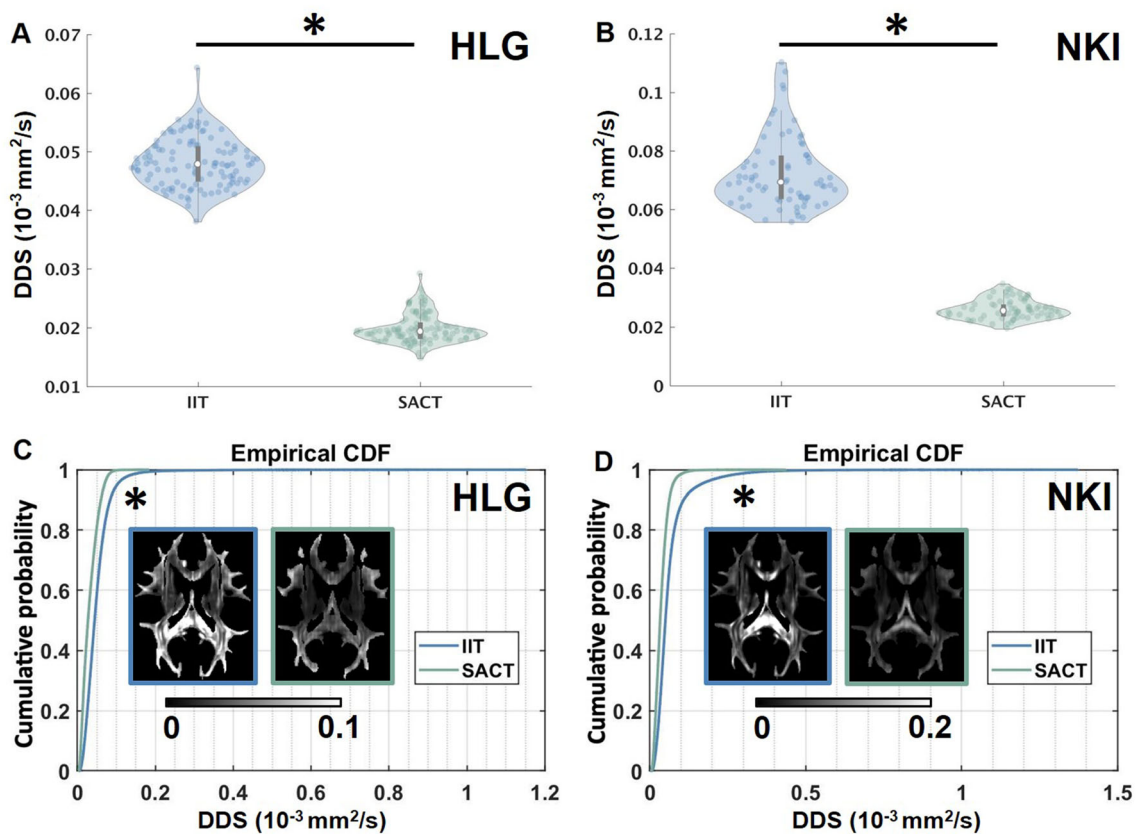


Fig. 3 Comparison of the DDS metrics between the spatial normalization using the SACT template (green) and the IIT template (blue). **A** Each data point in the violin map is the averaged region-wise DDS value for a subject corresponding to the use of different templates (white circle, median value; grey vertical bar, range from the first to the third quartile). $*P < 0.001$, paired Wilcoxon signed-rank. The result here is for the HLG dataset. **B** Results for the NKI dataset.

C Empirical cumulative distributions of the voxel-wise DDS values with different templates. The distribution is based on the DDS map averaged across the HLG datasets. Two slices are shown as examples of the distribution of voxel-wise DDS. The radiological convention is applied here. $*P < 0.001$, two-sample Kolmogorov–Smirnov test. CDF, the cumulative distribution function. **D** Results for the NKI dataset.

size referred to as the non-negative KS statistic value = 0.2777; for the NKI data: $P < 0.001$, KS statistic value = 0.2969). The mean DDS map of the SACT template showed a higher percentage of small DDS values than the IIT template (Fig. 3C, D). In particular, the splenium of corpus callosum showed high DDS values when applying the IIT template. In addition to the DDS metric, the results were corroborated by using the EDS (Fig. S4).

Second, regarding the deformation caused by registration, the mean value of regional deformations (voxel-wise mean of JDs within the white matter region) was significantly higher in the use of the SACT template for both the HLG and NKI data (Fig. 4A, B). Large effect sizes of the paired Wilcoxon rank tests were found (r -values: 0.6137 for the HLG data and 0.6149 for the NKI data). The mean JDs based on the SACT template were closer to 1 compared with the use of the IIT template, indicating less compression of the template when deformed into individual DT data. For the voxel-wise deformation, we checked the absolute value of the difference between the voxel-wise

JD and 1 (simply referred to as the JD difference). For both the HLG and the NKI data, the cumulative distributions of the mean JD difference maps were significantly different between the SACT template and the IIT template (for the HLG data: $P < 0.001$, KS statistic value = 0.4900; for the NKI data: $P < 0.001$, KS statistic value = 0.3329). The mean JD difference map of the SACT template showed a higher percentage of small JD difference compared with the IIT template (Fig. 4C, D). The above results were based on the global deformations, which were highly similar to the results based on the local deformations (Fig. S5).

Third, the quality of voxel-wise inter-subject normalization was assessed using both the dCOH and the $\bar{\sigma}_{FA}$. For the dCOH measuring the coherence in the dominant direction of diffusion across participants, we found that the cumulative distributions of the dCOH map showed a higher percentage of high dCOH values when the SACT template was used for both the HLG and NKI datasets (Fig. 5A, B; for the HLG data: $P < 0.001$, KS statistic value = 0.0493; for the NKI data: $P < 0.001$, KS statistic value =

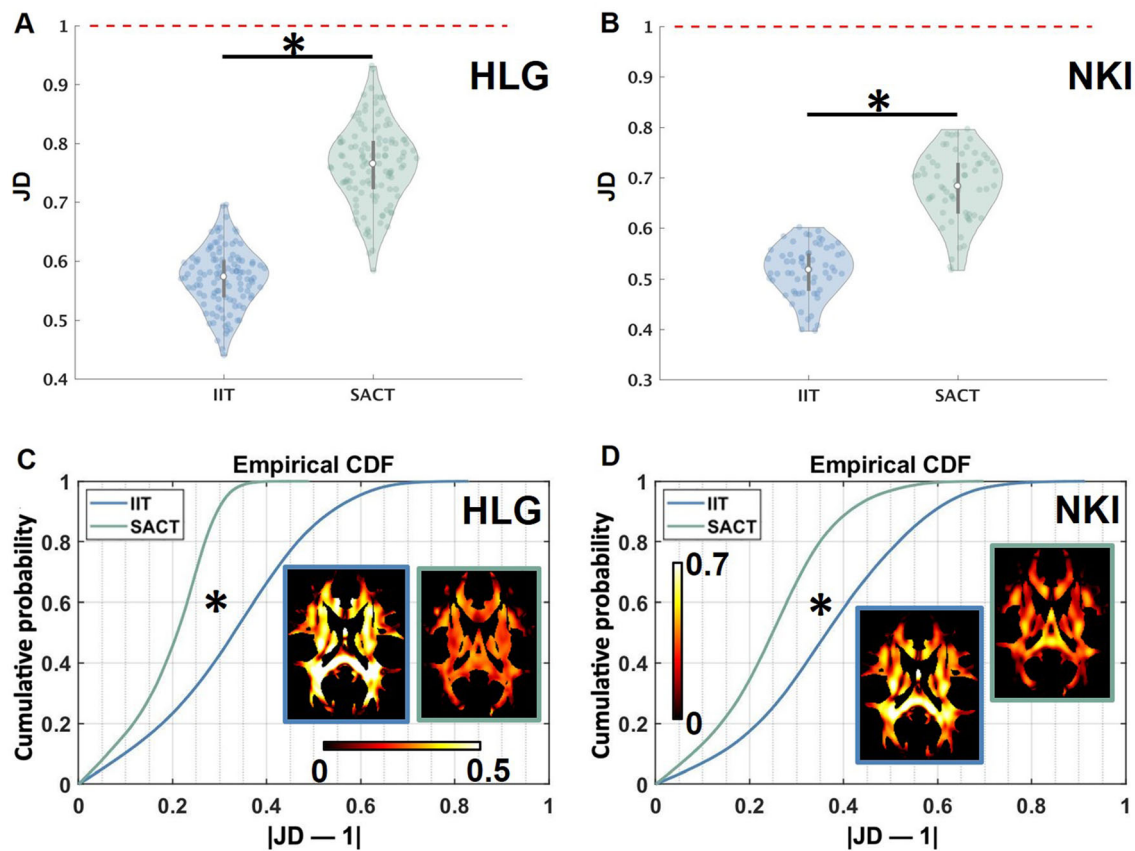


Fig. 4 Comparison of the global JD between the spatial normalization using the SACT template (green) and the IIT template (blue). **A** Each data point in the violin map is the averaged region-wise JD value for a subject corresponding to the use of different templates (white circle, median value; grey vertical bar, range from the first to the third quartile; red dashed line, JD value of 1 corresponding to neither compression nor expansion). $*P < 0.001$, paired Wilcoxon signed-rank. Results for the HLG dataset. **B** Results for the NKI

dataset. **C** Empirical cumulative distributions of the voxel-wise absolute difference between JD and 1 with different templates. The distributions are based on the averaged $|JD - 1|$ maps across the HLG datasets. Two slices are shown as examples of the distribution of voxel-wise JD. The radiological convention is applied. $*P < 0.001$, the two-sample Kolmogorov–Smirnov test. CDF, the cumulative distribution function. **D** Results for the NKI dataset.

0.0473). Similarly, for the $\bar{\sigma}_{FA}$ capturing the variability in diffusion anisotropy, a lower percentage of $\bar{\sigma}_{FA}$ was found with the SACT template for both datasets (Fig. 5C, D; for the HLG data: $P < 0.001$, KS statistic value = 0.0498; for the NKI data: $P < 0.001$, KS statistic value = 0.0443). Taken together, these results suggest that the quality of spatial normalization can be increased by using the SACT template in the two different datasets of school-age children.

Age Associations with White Matter Properties Based on the SACT Template

Regarding the DT-derived metrics including FA and MD, we found a highly similar pattern of age-related associations in the VBA analyses based on the IIT template and the SACT template (Fig. S6). The cumulative distributions of the z -statistics measuring the voxel-wise association between FA (or MD) and age were close to each other

regarding the two templates (Fig. S6A, B). Although the P -values of the two-sample KS tests between the cumulative distributions of the z -statistics were significant for both the FA-based comparison of the two templates and the MD-based comparison ($P < 0.05$), the KS statistic values were small, which were 0.0061 for the FA-based comparison and 0.0156 for the MD-based comparison. For the TBSS analyses based on FA or MD, the cumulative distributions of the GLM-based t -statistics on the skeleton were also close to each other when using the two templates (Fig. S7). The KS statistic values were small (0.0087 for the FA-based comparison, 0.0179 for the MD-based comparison). In particular, by performing permutation tests on the TBSS results, significant associations ($P < 0.05$, corrected for the familywise error rate) between FA and age passing the correction of multiple comparisons were only found in the positive values based on either the IIT template or the SACT template. Therefore, we compared the cumulative distributions of the positive t -statistics between the FA-

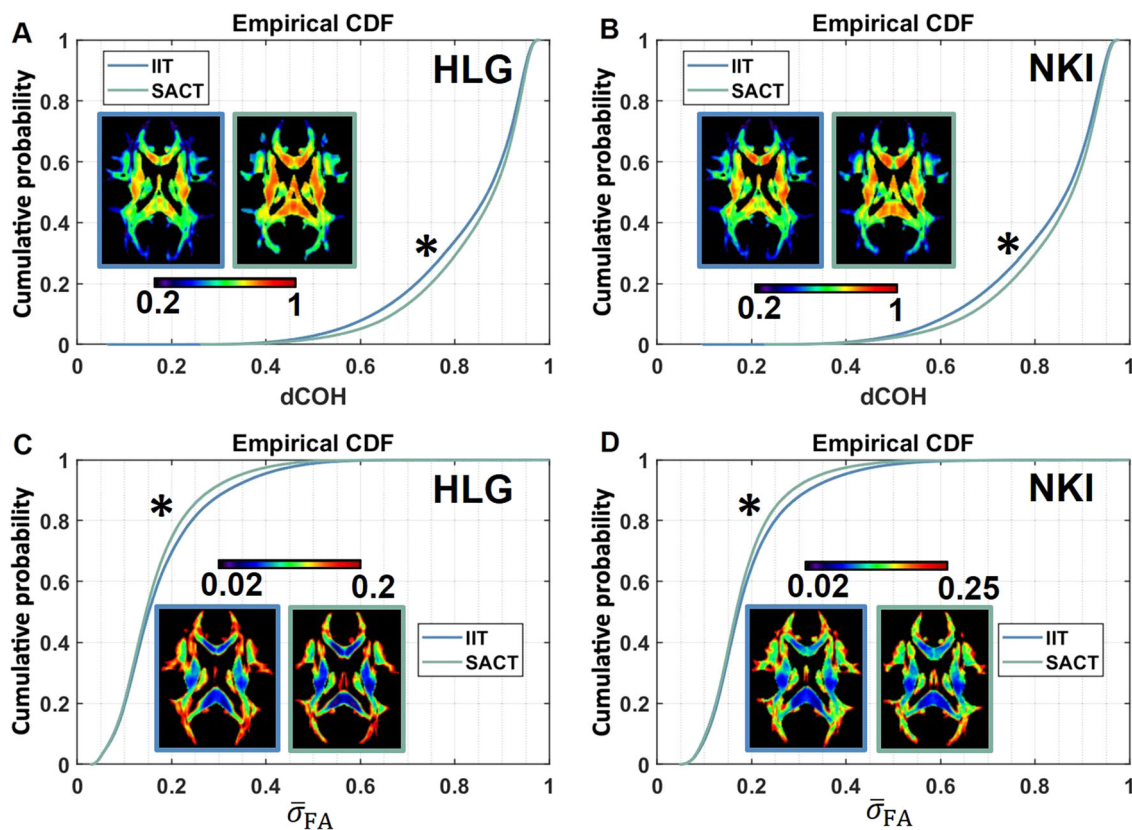


Fig. 5 Evaluation of the inter-subject coherence of normalized data based on the SACT template (green) and the IIT template (blue). The radiological convention is applied. **A** For the HLG dataset, the empirical cumulative distributions of the voxel-wise dCOH values regarding different templates. Two slices are shown as examples of the distribution of voxel-wise dCOH. * $P < 0.001$, the two-sample

Kolmogorov–Smirnov test. CDF, the cumulative distribution function. **B** The dCOH results for the NKI dataset. **C** For the HLG dataset, the empirical cumulative distributions of the voxel-wise $\bar{\sigma}_{FA}$ values regarding different templates. Two slices are shown as examples of the distribution of voxel-wise $\bar{\sigma}_{FA}$. **D** The $\bar{\sigma}_{FA}$ results for the NKI dataset.

based TBSS using the IIT template and that using the SACT template but found no significant difference between the two distributions ($P = 0.2296$, KS statistic value = 0.0047, Fig. 6A). However, a significant difference with a higher effect size ($P < 0.001$, KS statistic value = 0.0929) was found between the cumulative distributions of the corrected P -values (Fig. 6B). FA-based TBSS using the IIT template tended to show a higher percentage of small P -values. In addition, the significant results using either the IIT template or the SACT template were spatially similar (Fig. 6C). For the MD-based TBSS, significant results were found in negative associations only. Moreover, MD-based TBSS using the SACT template tended to show a higher percentage of small, corrected P -values (Fig. S8A), although the spatial distributions of the significant results were similar (Fig. S8B).

Regarding the local deformation assessed by the JD of the non-linear deformation field (local JD), we found highly similar statistical maps in the VBA analyses based on the IIT template and the SACT template, although higher positive age associations were found on part of the

right posterior thalamic radiation when using the SACT template (Fig. S9A, B). Notably, the cumulative distributions of the corresponding z -statistics were significantly different regarding the two templates ($P = 1.04 \times 10^{-296}$, KS statistic value = 0.0407). There was a higher percentage of positive age associations with the local deformation when using the SACT template, where the positive association indicated a local expansion of white matter with age. When using the global JD, we still found similar results (Fig. S10).

Discussion

In this study, we established a diffusion tensor template for school-age children, the SACT template, based on a large sample of high-quality DWI data. Compared with the IIT template, we found higher spatial normalization accuracy when conducting registration with the SACT template in two DWI datasets of school-age children. In addition, as an example of practical application, we demonstrated that the

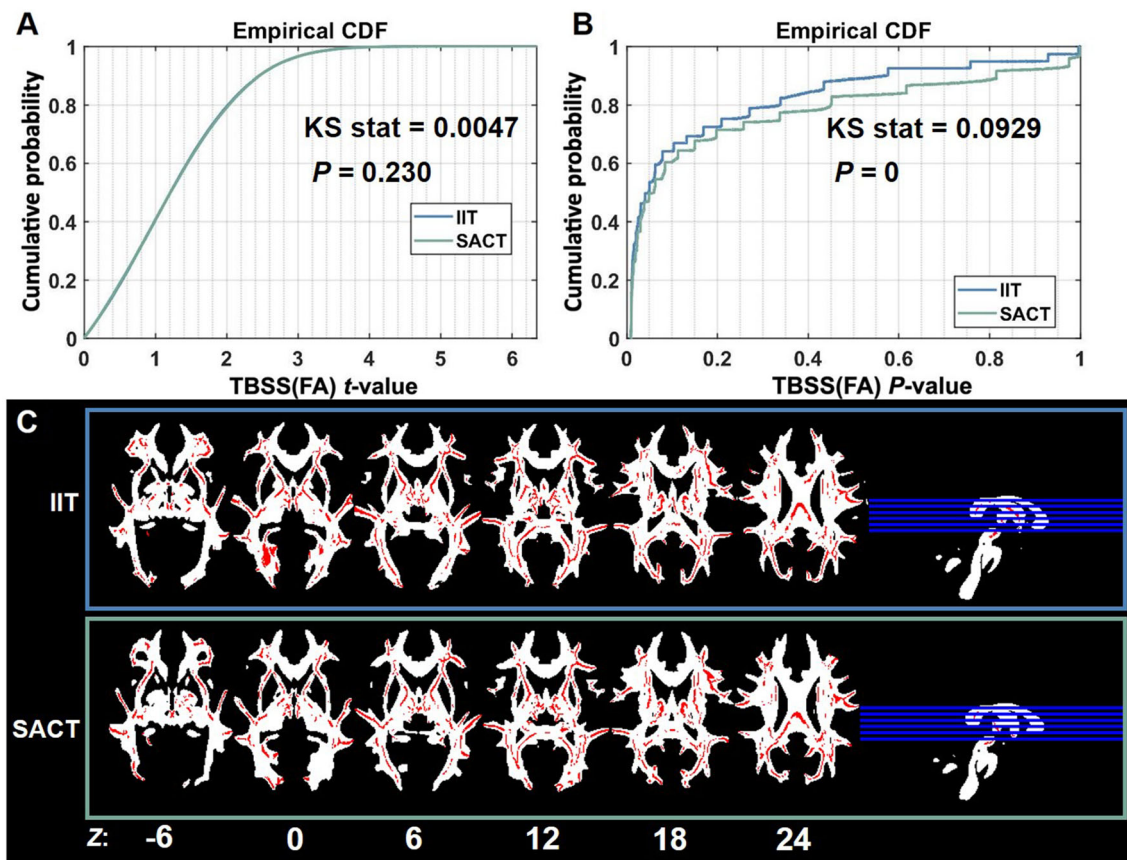


Fig. 6 TBSS analysis of FA and age for the SACT template (green) and the IIT template (blue). **A** Empirical cumulative distributions of GLM-based t -values on the skeleton. Only positive t -values are included. KS stat, two-sample Kolmogorov–Smirnov test. CDF, the cumulative distribution function. **B** Empirical cumulative distributions of corrected P -values for positive associations. **C** Examples of

axial slices showing the spatial distribution of significant positive association ($P < 0.05$, corrected for the familywise error rate). The significant results (red) are overlaid on the FA mask (white) derived from the mean FA maps across subjects with a threshold of 0.2. Z, MNI axial coordinates. The radiological convention is applied.

use of different DT templates influenced the distributions of statistics regarding age associations with the DT-derived metrics. Taken together, the SACT template provides a standard space reference for the spatial normalization of DWI data in school-age children.

Establishment of the SACT Template

Sample size and data quality are basic but crucial factors that affect the reliability of the established template [23]. To construct the SACT template, we used a large sample dataset comprising >300 participants that fully cover the school-age period. Compared with previous efforts to establish an adult DT template [17, 19, 47], a larger sample size would cover more population variation. Moreover, with the same protocol, all of the data were acquired from the same scanner operating by a fixed technical team at Peking University, avoiding the harmonization of multi-site DWI data [48]. Regarding the data quality control, several independent raters carefully evaluated the data

quality to remove outliers. We further used DTIPrep to double-check the quality of DWI data, especially for corrupted gradients. DTIPrep has been shown to exhibit suitable performance for this purpose [49].

Regarding the methodological consideration of the construction of the SACT template, we considered the distribution of the subjects over the age range. Using a dataset with subjects not uniformly distributed over the age range may bias the established template to some extent. This is more apparent for a pediatric template [21, 50, 51], given the rapid development of the underlying brain structures. Moreover, age has been recognized as the dominant factor explaining the variance of brain structures when constructing a pediatric template [51]. In our scenario, an extreme example would be if we used a dataset comprised only of children of the same age; the template derived from this dataset would only represent the DT features of this specific age, which could not be directly generalized to cover a broader age range. Therefore, we assigned priority to balance the distribution of age by re-

sampling the datasets. This strategy contributed to not only the balance of age but also the test of the dependence between the SACT template and the underpinning samples. Specifically, given the possibility that different levels of variance might exist among age groups and samplings, if the derived templates from different re-sampled datasets of balanced age distribution were significantly inconsistent, there would be strong dependence between the derived template and the adopted samples, which might be accounted for by the inhomogeneity of the datasets. This would imply that it is not appropriate to derive a representative template for a specific population. In contrast, if we could derive highly consistent templates, it would demonstrate the existence of a representative template regardless of the re-sampled datasets. Based on these considerations, we found highly consistent templates derived from the 100 re-sampled datasets with small mean EDS values across the voxel-wise tensors of the entire brain and highly correlated FA/MD maps (Fig. 2A). Notably, EDS is a metric used to directly measure the difference between two tensors, providing a straightforward evaluation of the similarity between DT templates. Given the high consistency across the templates derived from the re-sampled datasets, we further averaged these templates to derive the SACT template, reducing the limited variations across them.

In addition, the SACT template and the other age-specific templates were highly similar in terms of EDS and FA values (Tables S2–4). In particular, the SACT template showed the highest FA-based correlation and the lowest EDS with the other age-specific templates. This suggests that the SACT template represents the DT features of the age range well.

Spatial Normalization Using the SACT Template

Improving the quality of spatial normalization is one of the crucial purposes to construct a template. Regarding the registration of DTI data, it has been demonstrated that the use of tensor-based registration algorithms combined with tensor templates presents higher normalization accuracy than scalar-based algorithms [11, 19]. Therefore, we evaluated the SACT template in terms of comparison with another DT template by using a tensor-based registration method from DTI-TK [25, 27]. As there are no publicly available DT templates of school-age children (to the best of our knowledge), we compared the SACT template with the DT template of the IIT Human Brain Atlas (v.5.0), which has demonstrated function in the spatial normalization of adult data [19]. More importantly, we conducted the evaluation based on two datasets (HLG and NKI) with different demographic properties and parameters of DWI data acquisition to offer more general comparisons. The

comparison covered three aspects to provide a multi-view evaluation of the template-based effect on spatial normalization. First, we found a closer tensor-based distance (DDS and EDS) between the normalized individual DT data and the SACT template (Figs 3 and S4), indicating that the data are close to the template. As noted before, the tensor-based distance provides an intuitive metric to measure the relationship between two tensors, which is often used for the evaluation of DTI spatial normalization [20, 24, 52]. Second, we found less volumetric deformation (expansion or compression) measured by JD when registering the individual DT data to the SACT template compared with the IIT template. We compared both the global deformation and the local deformation to confirm that results were similar regarding the lower deformation accompanied by the SACT template (Figs 4 and S5). Given the continued development of brain volume during middle childhood [53], we used global volumetric deformation to consider the effect of the different brain sizes between the individual data and the template. When including the affine transformation for calculation of the JD, the effect size of the comparison of the global deformation was larger than the local deformation, indicating the effect of brain size. Third, we found higher inter-subject consistency across the normalized DT data when using the SACT template (Fig. 5). High inter-subject consistency is beneficial to group-wise analyses [25, 54, 55]. Notably, the above findings were consistent between the two test datasets (the HLG one and the NKI one). Especially, participants of the NKI dataset are primarily Caucasian and African-American [32, 56], indicating the general feasibility of the SACT template in non-Chinese pediatric samples. A potential explanation for the performance of the SACT template might be its closeness to the brains of school-age children regarding brain size and white matter microstructures.

Effects of the SACT Template on the Detection of Age Associations

We assessed the association between the properties of white matter and chronological age, which was a straightforward but practical example to evaluate the statistical results based on the SACT template and the IIT template. Rather than interpreting the age associations, we were more interested in the comparison of the statistical results in terms of the two templates.

In the comparison of the statistical results of VBA based on the two templates, the statistical results of the voxel-wise associations between FA (or MD) and age were highly similar in their cumulative distributions of z-statistics and the corresponding spatial maps (Fig. S6). Although the cumulative distributions here were significantly different, the highly-overlapped patterns of the distributions and

limited effect sizes of the difference still indicate a negligible effect of the templates on these statistical results. Our findings are consistent with a previous study that found highly similar results of age associations with FA using different DT templates in adults [57]. We further explained the high similarity by showing the highly-correlated FA (MD) values across participants when using the two templates (Fig. S11). The high accuracy of spatial normalization using DTI-TK might also contribute to the high similarity. Interestingly, we found an effect of the templates on the statistical results of the VBA-based associations between the deformations and ages, using either the global JD or the local JD to measure the deformations. The cumulative distributions of z -statistics were separate from each other, showing a relatively large effect size (Figs S9, S10). The difference might be accounted for by the variation of the correlation between JDs derived from the two templates. Compared with FA, the correlation coefficient of the JD between the two templates had a wider range from 0.81 to 0.99 (Fig. S12), although it still indicated high similarity. As the JD was directly derived from the deformation field of registration, it might be more sensitive to the difference between the templates that served as the references for the optimization of the registration algorithm.

In addition to VBA, the quality of spatial normalization contributes to the reliability of TBSS analysis [47]. Consistent with the results of VBA, the GLM-based t -statistics showed negligible differences regarding the age associations with the skeletonized FA or MD when comparing the TBSS results based on the two templates (Fig. S7). However, the statistical results of the TFCE-based permutation tests presented a significant difference in the cumulative distributions of the voxel-wise corrected P -values, especially in the analyses using FA (Fig. 6B). That is, although there was no significant difference between the cumulative distributions of the t -statistics (Fig. 6A), a higher percentage of voxels survived multiple comparison corrections based on spatial normalization using the IIT template. As the TFCE score of each voxel depends on the supporting sections underneath it [45, 58], the difference might be caused by the different spatial layout of the skeleton from the mean FA maps (in our case: 1.2×10^5 voxels for the IIT-based skeleton and 1.0×10^5 voxels for the SACT-based skeleton). Therefore, for TBSS on the DTI data of the school-age children, the difference between the SACT template and the IIT template may affect the distribution of the permutation-based P -values more than the GLM-based statistics, although the statistical maps are still spatially similar. In addition, our TBSS findings with both FA and MD are spatially consistent with previous studies [59, 60].

Limitations and Open Problems

It is worth discussing several concerns that may be addressed in the future. First, as the adaptive kernel regression approach was used for the construction of the T1w template [50], a similar idea could also be developed regarding the DTI data to establish the DT templates with a continuous age range and a balanced sample size. Second, the SACT template is based on a Chinese pediatric dataset, although we demonstrated that non-Chinese pediatric samples (the NKI dataset) also had higher accuracy of spatial normalization using the SACT template. It may be worth constructing another SACT template to explore the population-level difference based on non-Chinese samples. Third, although we found high and close similarities among age-specific DT templates derived from age groups with different sample sizes, we should be cautious with unbalanced sample sizes of age groups. Further efforts will be required to validate the age-specific templates when more data become available. Fourth, the evaluation of the inter-subject spatial normalization accuracy was based on data from healthy subjects. It would be interesting to examine the performance of the SACT template on data from subjects with atypical development.

Acknowledgments This work was supported by the Beijing Brain Initiative of Beijing Municipal Science & Technology Commission (Z181100001518003) and the National Natural Science Foundation of China (31521063). All the authors thank the National Center for Protein Sciences at Peking University for assistance with brain scanning. The authors thank the “2019 Helmholtz – OCPC – Program for the involvement of postdocs in bilateral collaboration projects” for financial support that enabled this important study. All the authors are grateful for the contribution of the participants to the CBD project and the open access to the NKI-Rockland sample.

Conflict of interest The authors declare no conflicts of interest.

References

1. DelGiudice M. Middle Childhood: An Evolutionary-developmental Synthesis: Handbook of Life Course Health Development, Cham: Springer, 2018, pp 95–107.
2. Ghetti S, Bunge SA. Neural changes underlying the development of episodic memory during middle childhood. *Dev Cogn Neurosci* 2012, 2: 381–395.
3. Lebel C, Deoni S. The development of brain white matter microstructure. *Neuroimage* 2018, 182: 207–218.
4. Pierpaoli C, Basser PJ. Toward a quantitative assessment of diffusion anisotropy. *Magn Reson Med* 1996, 36: 893–906.
5. Kochunov P, Hong LE, Dennis EL, Morey RA, Tate DF, Wilde EA, *et al.* ENIGMA-DTI: Translating reproducible white matter deficits into personalized vulnerability metrics in cross-diagnostic psychiatric research. *Hum Brain Mapp* 2020: hbm.24998.
6. Lebel C, Treit S, Beaulieu C. A review of diffusion MRI of typical white matter development from early childhood to young adulthood. *NMR Biomed* 2019, 32: e3778. <https://doi.org/10.1002/nbm.3778>.

7. Liang SG, Wang Q, Kong XZ, Deng W, Yang X, Li XJ. White matter abnormalities in major depression biotypes identified by diffusion tensor imaging. *Neurosci Bull* 2019, 35: 867–876.
8. Qiu AQ, Mori S, Miller MI. Diffusion tensor imaging for understanding brain development in early life. *Annu Rev Psychol* 2015, 66: 853–876.
9. Mori S, Oishi K, Faria AV. White matter atlases based on diffusion tensor imaging. *Curr Opin Neurol* 2009, 22: 362–369.
10. Bach M, Laun FB, Leemans A, Tax CM, Biessels GJ, Stieltjes B, *et al.* Methodological considerations on tract-based spatial statistics (TBSS). *Neuroimage* 2014, 100: 358–369.
11. Irfanoglu MO, Nayak A, Jenkins J, Hutchinson EB, Sadeghi N, Thomas CP, *et al.* DR-TAMAS: Diffeomorphic registration for tensor accurate alignment of anatomical structures. *Neuroimage* 2016, 132: 439–454.
12. Wang Y, Gupta A, Liu ZX, Zhang H, Escolar ML, Gilmore JH, *et al.* DTI registration in atlas based fiber analysis of infantile Krabbe disease. *Neuroimage* 2011, 55: 1577–1586.
13. Wang Y, Yu Q, Liu ZX, Lei T, Guo Z, Qi M, *et al.* Evaluation on diffusion tensor image registration algorithms. *Multimed Tools Appl* 2016, 75: 8105–8122.
14. Timmers I, Roebroek A, Bastiani M, Jansma B, Rubio-Gozalbo E, Zhang H. Assessing microstructural substrates of white matter abnormalities: a comparative study using DTI and NODDI. *PLoS One* 2016, 11: e0167884.
15. Hsu YC, Lo YC, Chen YJ, Wedeen VJ, Isaac Tseng WY. NTU-DSI-122: a diffusion spectrum imaging template with high anatomical matching to the ICBM-152 space. *Hum Brain Mapp* 2015, 36: 3528–3541.
16. Mori S, Oishi K, Jiang HY, Jiang L, Li X, Akhter K, *et al.* Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. *Neuroimage* 2008, 40: 570–582.
17. Peng HL, Orlichenko A, Dawe RJ, Agam G, Zhang SW, Arfanakis K. Development of a human brain diffusion tensor template. *Neuroimage* 2009, 46: 967–980.
18. Zhang H, Yushkevich PA, Rueckert D, Gee JC. A computational white matter atlas for aging with surface-based representation of fasciculi. *Biomed Image Regist* 2010, https://doi.org/10.1007/978-3-642-14366-3_8.
19. Zhang S, Arfanakis K. Evaluation of standardized and study-specific diffusion tensor imaging templates of the adult human brain: template characteristics, spatial normalization accuracy, and detection of small inter-group FA differences. *Neuroimage* 2018, 172: 40–50.
20. Zhang SW, Peng HL, Dawe RJ, Arfanakis K. Enhanced ICBM diffusion tensor template of the human brain. *Neuroimage* 2011, 54: 974–984.
21. Fonov V, Evans AC, Botteron K, Almli CR, McKinstry RC, Collins DL, *et al.* Unbiased average age-appropriate atlases for pediatric studies. *Neuroimage* 2011, 54: 313–327.
22. Yoon U, Fonov VS, Perusse D, Evans AC, Brain Development Cooperative Group. The effect of template choice on morphometric analysis of pediatric brain data. *Neuroimage* 2009, 45: 769–777.
23. Yang GY, Zhou SZ, Bozek J, Dong HM, Han MZ, Zuo XN, *et al.* Sample sizes and population differences in brain template construction. *Neuroimage* 2020, 206: 116318.
24. Zhang SW, Arfanakis K. Role of standardized and study-specific human brain diffusion tensor templates in inter-subject spatial normalization. *J Magn Reson Imaging* 2013, 37: 372–381.
25. Zhang H, Avants BB, Yushkevich PA, Woo JH, Wang SM, McCluskey LF, *et al.* High-dimensional spatial normalization of diffusion tensor images improves the detection of white matter differences: an example study using amyotrophic lateral sclerosis. *IEEE Trans Med Imaging* 2007, 26: 1585–1597.
26. Wang Y, Shen Y, Liu DY, Li GQ, Guo Z, Fan YY, *et al.* Evaluations of diffusion tensor image registration based on fiber tractography. *Biomed Eng Online* 2017, 16: 9.
27. Zhang H, Yushkevich PA, Alexander DC, Gee JC. Deformable registration of diffusion tensor MR images with explicit orientation optimization. *Med Image Anal* 2006, 10: 764–785.
28. Zhang H, Yushkevich PA, Rueckert D, Gee JC. Unbiased white matter atlas construction using diffusion tensor images. *Med Image Comput Assist Interv* 2007, 10: 211–218.
29. Zhao TD, Liao XH, Fonov VS, Wang QS, Men WW, Wang YP, *et al.* Unbiased age-specific structural brain atlases for Chinese pediatric population. *Neuroimage* 2019, 189: 55–70.
30. Sha Tao. Intelligence development and school adjustment of school-age children and adolescents: a follow-up cohort study. *Psychol Commun* 2019, 2: 88–90.
31. Dong Q, Lin C. Standardized Tests of the National Children's Study of China, Beijing, Science Press, 2011, pp 1–14.
32. Nooner KB, Colcombe SJ, Tobe RH, Mennes M, Benedict MM, Moreno AL, *et al.* The NKI-rockland sample: a model for accelerating the pace of discovery science in psychiatry. *Front Neurosci* 2012, 6: 152.
33. Xu JQ, Moeller S, Auerbach EJ, Strupp J, Smith SM, Feinberg DA, *et al.* Evaluation of slice accelerations using multiband echo planar imaging at 3 T. *Neuroimage* 2013, 83: 991–1001.
34. Oguz I, Farzinfar M, Matsui J, Budin F, Liu ZX, Gerig G, *et al.* DTIPrep: quality control of diffusion-weighted images. *Front Neuroinform* 2014, 8: 4.
35. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. *Fsl*. *Neuroimage* 2012, 62: 782–790.
36. Andersson JLR, Sotiropoulos SN. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* 2016, 125: 1063–1078.
37. Andersson JLR, Graham MS, Zsoldos E, Sotiropoulos SN. Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images. *Neuroimage* 2016, 141: 556–572.
38. Andersson JLR, Graham MS, Drobjak I, Zhang H, Filippini N, Bastiani M. Towards a comprehensive framework for movement and distortion correction of diffusion MR images: Within volume movement. *Neuroimage* 2017, 152: 450–466.
39. Lutkenhoff ES, Rosenberg M, Chiang J, Zhang KY, Pickard JD, Owen AM, *et al.* Optimized brain extraction for pathological brains (optiBET). *PLoS One* 2014, 9: e115551. <https://doi.org/10.1371/journal.pone.0115551>.
40. Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging* 2001, 20: 45–57.
41. Jones DK, Griffin LD, Alexander DC, Catani M, Horsfield MA, Howard R, *et al.* Spatial normalization and averaging of diffusion tensor MRI data sets. *Neuroimage* 2002, 17: 592–617.
42. van Hecke W, Leemans A, Emsell L. DTI analysis methods: Voxel-based analysis. *Diffusion Tensor Imaging*. New York: Springer, 2016: 183–203.
43. Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, MacKay CE, *et al.* Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage* 2006, 31: 1487–1505.
44. Winkler AM, Ridgway GR, Webster MA, Smith SM, Nichols TE. Permutation inference for the general linear model. *Neuroimage* 2014, 92: 381–397.
45. Smith SM, Nichols TE. Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 2009, 44: 83–98.

46. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. *J Exp Psychol Gen* 2012, 141: 2–18.
47. Zalesky A. Moderating registration misalignment in voxelwise comparisons of DTI data: a performance evaluation of skeleton projection. *Magn Reson Imaging* 2011, 29: 111–125.
48. Cetin Karayumak S, Bouix S, Ning LP, James A, Crow T, Shenton M, *et al.* Retrospective harmonization of multi-site diffusion MRI data acquired with different acquisition parameters. *Neuroimage* 2019, 184: 180–200.
49. Liu BL, Zhu T, Zhong JH. Comparison of quality control software tools for diffusion tensor imaging. *Magn Reson Imaging* 2015, 33: 276–285.
50. Serag A, Aljabar P, Ball G, Counsell SJ, Boardman JP, Rutherford MA, *et al.* Construction of a consistent high-definition spatio-temporal atlas of the developing brain using adaptive kernel regression. *Neuroimage* 2012, 59: 2255–2265.
51. Wilke M, Holland SK, Altaye M, Gaser C. Template-O-Matic: a toolbox for creating customized pediatric templates. *Neuroimage* 2008, 41: 903–913.
52. Adluru N, Zhang H, Fox AS, Shelton SE, Ennis CM, Bartosic AM, *et al.* A diffusion tensor brain template for rhesus macaques. *Neuroimage* 2012, 59: 306–318.
53. Peterson M, Warf BC, Schiff SJ. Normative human brain volume growth. *J Neurosurg Pediatr* 2018, 21: 478–485.
54. Evans AC, Janke AL, Collins DL, Baillet S. Brain templates and atlases. *Neuroimage* 2012, 62: 911–922.
55. Jones DK, Cercignani M. Twenty-five pitfalls in the analysis of diffusion MRI data. *NMR Biomed* 2010, 23: 803–820.
56. Park BY, Byeon K, Lee MJ, Chung CS, Kim SH, Morys F, *et al.* Whole-brain functional connectivity correlates of obesity phenotypes. *Hum Brain Mapp* 2020, 41: 4912–4924.
57. Cabeen RP, Bastin ME, Laidlaw DH. A Comparative evaluation of voxel-based spatial mapping in diffusion tensor imaging. *Neuroimage* 2017, 146: 100–112.
58. Spisák T, Spisák Z, Zunhammer M, Bingel U, Smith S, Nichols T, *et al.* Probabilistic TFCE: a generalized combination of cluster size and voxel intensity to increase statistical power. *Neuroimage* 2019, 185: 12–26.
59. Krogsrud SK, Fjell AM, Tamnes CK, Grydeland H, Mork L, Due-Tønnessen P, *et al.* Changes in white matter microstructure in the developing brain—a longitudinal diffusion tensor imaging study of children from 4 to 11 years of age. *Neuroimage* 2016, 124: 473–486.
60. Tamnes CK, Roalf DR, Goddings AL, Lebel C. Diffusion MRI of white matter microstructure development in childhood and adolescence: methods, challenges and progress. *Dev Cogn Neurosci* 2018, 33: 161–175.