Reliability on multiband diffusion NODDI models: A test retest study on children and adults

Martina Lucignani [a,1], Laura Breschi [a,1], Maria Camilla Rossi Espagnet [b,c], Daniela Longo [b], Lorenzo Figà Talamanca [b], Elisa Placidi [d], Antonio Napolitano [a,*]

[a] *Medical Physics Department, Bambino Gesù Children's Hospital IRCCS, Rome, Italy*
[b] *Neuroradiology Unit, Bambino Gesù Children's Hospital IRCCS, Rome, Italy*
[c] *Nesmos Department, Sapienza University, Rome, Italy*
[d] *Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Medical Physics UOC, Rome, Italy*

A R T I C L E   I N F O

A B S T R A C T

Neurite Orientation Dispersion and Density Imaging (NODDI) and Bingham-NODDI diffusion MRI models are nowadays very well-known models in the field of diffusion MRI as they represent powerful tools for the estimation of brain microstructure. In order to efficiently translate NODDI imaging findings into the diagnostic clinical practice, a test-retest approach would be useful to assess reproducibility and reliability of NODDI biomarkers, thus providing validation on precision of different fitting toolboxes. In this context, we conducted a test-retest study with the aim to assess the effects of different factors (i.e. fitting algorithms, multiband acceleration, shell configuration, age of subject and hemispheric side) on diffusion models reliability, assessed in terms of Intra-class Correlation Coefficient (ICC) and Variation Factor (VF). To this purpose, data from pediatric and adult subjects were acquired with Simultaneous-MultiSlice (SMS) imaging method with two different acceleration factor (AF) and four b-values, subsequently combined in seven shell configurations. Data were then fitted with two different GPU-based algorithms to speed up the analysis. Results show that each factor investigated had a significant effect on reliability of several diffusion parameters. Particularly, both datasets reveal very good ICC values for higher AF, suggesting that faster acquisitions do not jeopardize the reliability and are useful to decrease motion artifacts. Although very small reliability differences appear when comparing shell configurations, more extensive diffusion parameters variability results when considering shell configuration with lower b-values, especially for simple model like NODDI. Also fitting tools have a significant effect on reliability, but their difference occurs in both datasets and AF, so it appears to be independent from either misalignment and motion artifacts, or noise and SNR. The main achievement of the present study is to show how 10 min multi-shell diffusion MRI acquisition for NODDI acquisition can have reliable results in WM. More complex models do not appear to be more prone to less data acquisition as well as noisier data thus stressing the idea of Bingham-NODDI having greater sensitivity to true subject variability.

Data availability statement

Raw and preprocessed data supporting the findings of this study would be available upon reasonable request, made via email to the corresponding author. The more common software packages used come with a link in references.

1. Introduction

Neurite Orientation Dispersion and Density Imaging (NODDI) is a diffusion MRI (dMRI) technique that produces direct estimation of brain microstructure by modeling the tissue as a multi-compartmental system (Zhang et al., 2012), accounting for free, hindered and restricted water diffusion. Expressing the dMRI signal as a sum of contributions from each of the three compartments modeled, NODDI provides specific metrics (i.e. the Orientation Dispersion Index ODI, the volume fraction of intra-cellular v_{intra} and of cerebro-spinal fluid v_{iso} compartments) which are able to quantify the effects of density and orientation distribution of dendrites and axons. Based on NODDI, Bingham-NODDI model was then introduced (Tariq et al., 2016), in order to characterize anisotropic orientation dispersion of complex configurations like fanning and bending axons, overcoming the commonly used Diffusion Tensor Imaging (DTI).

* Corresponding author.
E-mail address: antonio.napolitano@opbg.net (A. Napolitano).
[1] Martina Lucignani and Laura Breschi equally contributed to this work

Since DTI metrics (e.g. the Fractional Anisotropy (FA) and Mean Diffusivity (MD)) were estimated by approximating a single fiber at each voxel and assuming mono-Gaussian model for water diffusion, they lack specificity in the White Matter (WM) tracts (Kodiweera et al., 2016; Wheeler-Kingshott and Cercignani, 2009).

Similarly to DTI, NODDI and Bingham-NODDI methods allow one fiber tract in each voxel and they represent powerful instruments, able to provide highly specific and biologically meaningful measures of tissue microstructure (Kunz et al., 2014). For this reason, these techniques recently gained increasing interest for applications in the clinical field, as they allowed to non-invasively investigate cognitive processes (Kamiya et al., 2020) and *in-vivo* cellular architecture of the developing human brain and to follow brain maturation (Lynch et al., 2020; Mah et al., 2017). Particularly, the ability to detect possible alterations in WM fiber organization (Dean et al., 2017) made the NODDI approach suitable to probe brain injury (Caverzasi et al., 2016; Palacios et al., 2020) and to investigate specific neurological disorders, such as Alzheimer's disease (Fu et al., 2020), Multiple Sclerosis (Mustafi et al., 2019; Schneider et al., 2017) and epilepsy disorders (Sone et al., 2018).

Despite their success in quantifying complex brain morphology, both NODDI and Bingham-NODDI present some clinical limitations, mainly related to the long acquisition and processing times. Multi-shell acquisition protocol with a large number of uniformly distributed diffusion weighted gradient directions for each b-value leads to long scan time, which limits its application in the clinical practice especially in the pediatric population. In this context, Multiband or Simultaneous Multi-Slice (SMS) imaging could represent a valid option for scan time reduction as it is able to simultaneously excite multiple slices (Setsompop et al., 2012), thus providing an acceleration in data acquisition which is proportional to the number of excited slices (i.e. acceleration factor AF). Moreover, fitting the data for diffusion parameter maps estimation is computationally expensive particularly when increasing dataset size and for highly complex biophysical models. The use of computational power offered by recent parallel computing architectures like Graphics Processing Units (GPUs) could represent a helpful solution to reduce processing time (Cheng et al., 2013). Among other algorithms for fast model fitting optimization like the Accelerated Microstructure Imaging via Convex Optimization (AMICO) ((Daducci et al., 2015)), two GPU-based toolboxes have been recently presented to fit both NODDI and Bingham-NODDI models. The first tool is *Microstructure Diffusion Toolbox* (MDT), a python-based toolbox freely available under an open source L-GPL license that is primarily meant for microstructure modeling and analysis of dMRI (Harms et al., 2017), while the second tool is *CUDA Diffusion Modeling Toolbox* (cuDIMOT), a GPU-based library for accelerated non-linear optimization (Hernandez-Fernandez et al., 2019). Compared to the widely used AMICO toolbox, MDT and cuDIMOT allow to fit NODDI and Bingham-NODDI models much faster, but it is not yet clear whether the significant fitting time reduction could sacrifice reliability of the metrics estimated. In order to efficiently translate NODDI imaging findings into the diagnostic clinical practice (Lerma-Usabiaga et al., 2019), a test-retest approach would be useful to assess reproducibility and reliability of NODDI biomarkers, thus providing validation on precision of the method, especially when considering GPU-based approach. In this context several recent studies focused on investigating the reliability and reproducibility of metrics from different dMRI techniques, such as DTI (Boekel et al., 2017; Duan et al., 2015), NODDI (Parvathaneni et al., 2018; Tariq et al., 2012) and 3-tissues Constrained Spherical Deconvolution (CSD) (Newman et al., 2019). Focusing on NODDI, previous studies demonstrated significant dependence of metrics on magnetic field strength (Chung et al., 2016) and reported excellent repeatability, especially for ODI and v_{intra} metrics (Andica et al., 2020; Granberg et al., 2017; Huber et al., 2019). Moreover, although a reproducibility study was recently performed to assess the effect of b-values and gradient directions on metrics (Parvathaneni et al., 2018), the effect of their interaction with other factors (e.g. SMS protocol and fitting tools) still need to be evaluated. .

The purpose of this work is therefore to perform a test-retest reliability assessment of NODDI and Bingham-NODDI models fitted with two different GPU-based algorithms (MDT, cuDIMOT) in terms of Intraclass Correlation Coefficient (ICC) and Variation Factor (VF). Focusing only on NODDI metrics, the reliability results from GPU-based algorithms were also compared to those from AMICO toolbox in order to provide quantifiable comparison between toolboxes with different performances. Particularly, the test-retest reliability was evaluated based on children and adult data for different SMS protocols and for seven different shell configurations. Along with it, a test on the brain asymmetry and its effect on reliability was performed.

2. Materials and methods

2.1. Subjects

This research was conducted in accordance with the Declaration of Helsinki. Ethical committee approval was obtained through Institutional Review Board (IRB) and written informed consent was obtained from every subject or next of kin for the execution of MRI acquisitions. Between September 2017 and September 2020 MRI data from healthy subjects were collected at the Bambino Gesù Children's Hospital (Rome, Italy). According to the recommendations on sample size for test-retest studies (Bujang and Baharum, 2017), assuming two observation per subject, a minimum sample size of 15 subjects is considered to be sufficient to detect ICC values of 0.6 for fixed values of alpha and power (respectively 0.05 and 80%). Consequently, we identified 31 subjects: 15 children (mean age = 8.9 y, range = 4.5–15.6 y, 8 males and 7 females) and 16 adult volunteers (mean age = 29.2y, range = 22.8–59.2 y, 8 males and 8 females).

2.2. Data acquisition

Images from both cohorts were acquired on a 3T scanner (Siemens Magnetom Skyra, Siemens Medical Systems, Erlangen, Germany) equipped with a 32 channels head-coil. Each participant received a test-retest protocol, including 2 diffusion MRI acquisitions with a Siemens prototype EPI-based sequence (SMS). Pediatric subjects underwent MRI for different clinical reasons and received a supra-tentorial acquisition, where 9 out of 15 were acquired with AF = 2 (TR: 4500 ms, TE: 114 ms, FA: 90°, ST: 2 mm, #Slice: 45, AF: 2), while the other 6 received AF=3 protocol (TR: 3000 ms, TE: 114 ms, FA: 90°, ST: 2 mm, #Slice: 45, AF: 3). Adults received full brain acquisition, where 8 out of 16 volunteers were acquired with AF = 2 protocol (TR: 6400 ms, TE: 114 ms, FA: 90°, ST: 2 mm, #Slice: 66, AF: 2), while the other 8 volunteers received AF = 3 (TR: 4500 ms, TE: 114 ms, FA: 90°, ST: 2 mm, #Slice: 69, AF: 3). Data were acquired at four different b-values to maximize signal-to-noise ratio (SNR) as shown before (Zhang et al., 2018): (1) $b = 300$ s/mm^2 along 10 gradient direction, (2) $b = 700$ s/mm^2 and (3) $b = 1000$s/mm^2, both with gradients along 30 directions and (4) $b = 2000$s/mm^2, with 64 directions.

Additional five brain volumes with no diffusion weighting ($b = 0$ s/mm^2) were collected. Only in adult volunteers one further acquisition using a reversed phase encoding direction (posterior to anterior) was also acquired to allow the estimation of susceptibility-induced distortion. All diffusion images were acquired with a spatial resolution of $2.0 \times 2.0 \times 2.0$ mm and acquisition matrix of 128×128. Total acquisition time for diffusion imaging was approximately 20 min for AF=2 and 15 min for AF=3. Children did not exit the scanner between test and retest but the two sessions were separated by about 20 mins. Conversely, adult volunteers were out of the scanner for about 20 min between sessions.

2.3. Data preprocessing

We assessed diffusion properties for seven different shell combination: three 2-shell (P_{21}: b = 300–2000s/mm^2, P_{22}: b = 700–2000s/mm^2 and P_{23}: b = 1000–2000s/mm^2), three 3-shell (P_{31}: b = 300–700–2000s/mm^2, P_{32}: b = 300–1000–2000s/mm^2 and P_{33}: b = 700–1000–2000) and one 4-shell (P_4: b = 300–700–1000–2000s/mm^2). They were all processed for denoising, eddy current and motion artifacts correction using MRTrix3 *dwidenoise* (Veraart et al., 2016b, 2016a) and *dwipreproc* functions (https://github.com/mrtrix3, version 3.0.2). Diffusion-Weighted Image (DWI) data were segmented with SPM12 toolbox (https://www.fil.ion.ucl.ac.uk/spm/software/spm12/), producing tissue segmentation maps, that were logically summed to produce brain masks required as input during the fitting step. Whenever available, a susceptibility-induced off-resonance field was estimated from the pairs of images acquired with reversed phase-encode directions (Andersson et al., 2003; Smith et al., 2004) using topup of FSL (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki, version 6.0). Moreover, from eddy output of FSL, we estimated the amount of total movement during the single acquisition for each subject, averaging the RMS movement relative to the first volume (see Supplementary Material).

2.4. Data fitting

NODDI and Bingham-NODDI were applied to the pre-processed data from all the shell configurations. Model fitting was performed with two different toolboxes: MDT (https://github.com/cbclab/MDT, version 0.19.1), a python-based GPU accelerated toolbox freely available under an open source L-GPL licence (Harms et al., 2017) and cuDIMOT (https://users.fmrib.ox.ac.uk/~moisesf/cudimot/), a GPU-base library for accelerated nonlinear optimization (Hernandez-Fernandez et al., 2019). We used Powell's and Levenberg-Marquardt optimization algorithms respectively for MDT and cuDIMOT. Additionally, the NODDI model was also implemented using the python package of the AMICO toolbox (https://github.com/daducci/AMICO). The fitting was achieved with an Oracle cloud server setup with two NVIDIA® Tesla® P100, four Intel® Xeon® Gold 5120 CPUs @ 2.20 GHz, 187 GB RAM. Each toolbox produced three maps (ODI, v_{intra} and v_{iso}) for NODDI model and six maps (ODI along primary and secondary direction ODI_p and ODI_s, total ODI ODI_{tot}, Dispersion Anisotropy DA, v_{intra} and v_{iso}) for Bingham-NODDI model. An analysis based on Region of Interest (ROI) was then performed to investigate test-retest reliability. To this purpose, individual FA maps were first computed by fitting Diffusion Tensor Model on a multi-shell configuration and then aligned with Tract-Based Spatial Statistics (TBSS) FSL pipeline (Smith et al., 2006) to construct WM mean skeleton. ROIs were automatically extracted from JHU-ICBM-DTI-81 WM atlas and overlapped (*fslmath* function) to the skeleton in order to remove any cerebro-spinal fluid (CSF) and gray matter (GM) voxels and to ensure reproducibility to be evaluated on the same projections. Finally, the ROIs were back-projected from Montreal Neurological Institute (MNI) space into each subject space with inverse TBSS transformations. During the fitting, each tool approximates the noise standard deviation, whose estimate could be influenced by changes in the brain mask and misalignment in the b0 images. For this reason we investigated the effect of noise standard deviation estimates on the test-retest variability. To this purpose we fitted both NODDI and Bingham NODDI models on data from all shell configurations with MDT toolbox, comparing reliability results obtained in standard setting with those obtained with specific noise standard deviation estimates, previously computed with specific approaches, i.e. Moments (https://github.com/samuelstjean/autodmri) and MPPCA (https://github.com/mrtrix3) (St-Jean et al., 2020). A MANOVA analysis was performed between data with different noise standard deviation estimates both on ICC and VF (see Tables 20,21 of Supplementary Material).

2.5. Test-retest reliability

Diffusion parameter maps were computed for each shell configuration for both NODDI (Fig. 1A) and Bingham-NODDI (Fig. 1B) models. We evaluated ICC and VF distributions within multi-shell configurations for the two fitting toolboxes. In particular, we compared results from seven different shell combinations (P_{21}, P_{22}, P_{23}, P_{31}, P_{32}, P_{33} and P_4) for MDT and cuDIMOT. This approach was used for SMS sequences acquired with different AF both for pediatric subjects (we referred to children acquired with AF=2/3 as PED_SMS2/PED_SMS3), and adult volunteers (AD_SMS2/AD_SMS3). Since pediatric subjects were acquired with limited acquisition stack to reduce exam time, we focused their reliability assessment within 36 supra-tentorial WM ROIs. Out of 36 ROIs, 32 were investigated for test-retest reliability in order to assess effect of hemispheric side (16 ROIs for each hemispheric side). For adults, ROI-based analysis was performed both within 32 supra- and 8 infra-tentorial ROIs. Both ICC and VF were obtained on all diffusion scalar metrics (ODI, v_{intra}, v_{iso} for NODDI and ODI_s, ODI_p, ODI_{tot}, DA, v_{intra}, v_{iso} for Bingham-NODDI) and for each ROI. The VF gives a measure of within-subject variability (Albi et al., 2018; Jovicich et al., 2014; Papinutto et al., 2013), where lower scores indicate better reproducibility. It is computed as the ratio between the absolute test-retest error and their mean. The ICC is a widely used statistical approach (McGraw and Wong, 1996; Shrout and Fleiss, 1979) which combine both within- and across-subject variability to assess agreement in test-retest reliability (Albi et al., 2018; Buchanan et al., 2014; Duan et al., 2015; Hodkinson et al., 2013). Particularly, according to the practical guidelines reported in literature (Koo and Li, 2016), we identified the two-way mixed effects model for single measurement as best ICC measurement for our study (McGraw and Wong, 1996), ranging from 0 (poor reliability) to 1 (excellent reliability) and defined as follows:

$$ICC(A,1) = \frac{MS_R - MS_E}{MS_R - (k-1)MS_E + \frac{k}{n}(MS_C - MS_E)}$$

where MS_R is the mean square for rows (i.e. between subject), MS_C is the mean square of columns(i.e. within subjects), MS_E is the mean square error, k is the number of measurements and n is the number of subjects.

Negative ICC values appear when the between-subject variation is relatively small compared to the within-subject variation and are theoretically difficult to interpret (Giraudeau, 1996; Rousson et al., 2002). Therefore, negative ICCs were set to zero (i.e. completely non-reliable), as suggested in other test–retest studies using the ICC (Braun et al., 2012; Duan et al., 2015). Both VF and ICC (https://www.mathworks.com/matlabcentral/fileexchange/22099-intraclass-correlation-coefficient-icc) estimates were performed on a MATLAB R2019 environment (Salarian, 2020).

2.6. Statistical analysis

MANOVA analysis was performed on ICC and VF values for each NODDI parameter as extracted from the different ROIs. The model included main effects of acceleration factor (SMS), adult or children group (Subj), shell configuration, software choice (Tools) and hemispheric side. For each factor (Tool, SMS, Configuration, Subj and Side), tables of statistical results reported mean ICC (see Table 1 and Supplementary Material Tables 1,3) and VF (see Table 2 and Supplementary Material Tables 2,4) values and standard deviation, values of F statistic, significance and effect size (η) for each diffusion parameter.

In order to assess significant differences between diffusion models (NODDI and NODDI-Bingham), we run a *t*-test analysis on the ICC and VF of common metrics (i.e. ODI, ODI_{tot}, v_{intra} and v_{iso}). Moreover, the effect of different noise standard deviation estimates on reliability was also investigated performing a MANOVA analysis (see Tables 20,21 of Supplementary Material). The statistics were carried out in Statistical Package for Social Science (SPSS) software. We also computed mean

Table 1

Mean values, standard deviation, *F*-statistics, *p*-value and size effect of ICC for NODDI and Bingham-NODDI metrics to assess effect of different factors (tool, SMS, configuration and side) within 32 supra-tentorial ROIs. N.S. means Not Significant.

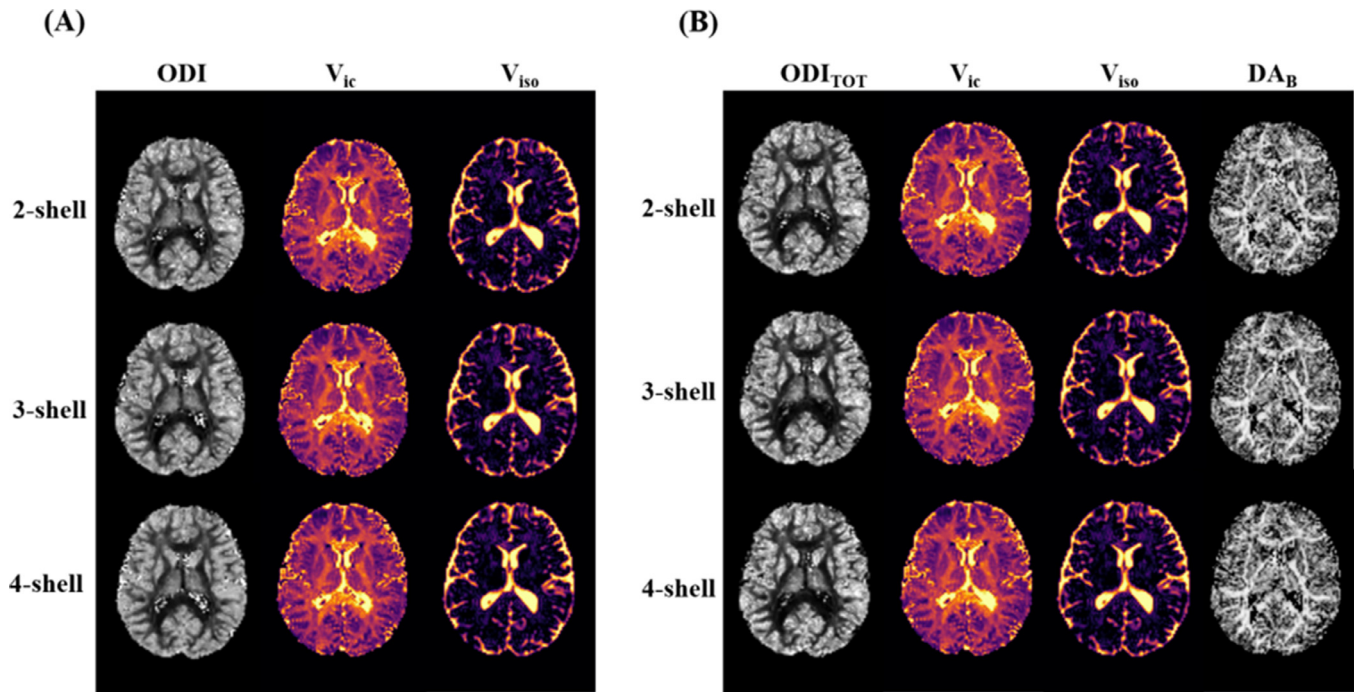| INTRA-CLASS CORRELATION COEFFICIENT ICC WITHIN SUPRA-TENTORIAL ROIS				
	Parameter	**Type of parameter**	**Mean (std dev)**	**F (pval) (η)**
EFFECT OF TOOL	ODI	MDT	.847 (0.005)	$F(1,1781) = 7.657$
		cuDIMOT	.827 (0.005)	$(5.71\,e^{-03})$ (4.28
	Vintra	MDT	.924 (0.004)	$F(1,1781) = 85.879$
				$(5.34e^{-20})$ (0.046)
		cuDIMOT	.876 (0.004)	
	Viso	MDT	.812 (0.006)	$F(1,1781) = 25.376$
		cuDIMOT	.851 (0.006)	$(5.19e^{-07})$ (0.014)
	DA	MDT	.822 (0.005)	$F(1,1781) = 8.453$
		cuDIMOT	.842 (0.005)	$(3.69\,e^{-03})$ (4.72
	ODIp	MDT	.828 (0.005)	N.S.^{03})
		cuDIMOT	.825 (0.005)	
	ODIs	MDT	.832 (0.005)	N.S.
		cuDIMOT	.834 (0.005)	
	ODItot	MDT	.829 (0.005)	N.S.
		cuDIMOT	.831 (0.005)	
	Vintra	MDT	.898 (0.004)	$F(1,1781) = 18.678$
		cuDIMOT	.876 (0.004)	$(1.63e^{-05})$ (0.01)
	Viso	MDT	.845 (0.005)	N.S.
		cuDIMOT	.850 (0.005)	
EFFECT OF SMS	ODI	SMS2	.782 (0.005)	$F(1,1781) = 237.794$
		SMS3	.892 (0.005)	$(1.87e^{-50})$ (0.118)
	Vintra	SMS2	.856 (0.004)	$F(1,1781) = 294.457$
		SMS3	.944 (0.004)	$(3.35e^{-61})$ (0.142)
	Viso	SMS2	.814 (0.006)	$F(1,1781) = 20.268$
		SMS3	.849 (0.006)	$(7.17e^{-06})$ (0.011)
	DA	SMS2	.794 (0.005)	$F(1,1781) = 120.255$
		SMS3	.870 (0.005)	$(4.01e^{-27})$ (0.063)
	ODIp	SMS2	.785 (0.005)	$F(1,1781) = 126.873$
		SMS3	.869 (0.005)	$(1.77e^{-28})$ (0.066)
	ODIs	SMS2	.776 (0.005)	$F(1,1781) = 219.357$
		SMS3	.889 (0.005)	$(6.83e^{-47})$ (0.109)
	ODItot	SMS2	.771 (0.005)	$F(1,1781) = 251.088$
		SMS3	.890 (0.005)	$(5.27e^{-53})$ (0.123)
	Vintra	SMS2	.825 (0.004)	$F(1,1781) = 575.393$
		SMS3	.949 (0.004)	$(2.05e^{-110})$ (0.244)
	Viso	SMS2	.838 (0.005)	$F(1,1781) = 7.998$
		SMS3	.857 (0.005)	(4.73^{-03}) (4.47^{-03})
EFFECT OF CONFIGURATIONS	ODI	P_{21}	.843 (0.010)	N.S.
		P_{22}	.841 (0.010)	
		P_{23}	.832 (0.010)	
		P_{31}	.837 (0.010)	
		P_{32}	.837 (0.010)	
		P_{33}	.835 (0.010)	
		P_4	.834 (0.010)	
	Vintra	P_{21}	.869 (0.007)	$F(6,1781) = 8.793$
		P_{22}	.893 (0.007)	(1.82^{-09}) $(9.00e^{-03})$
		P_{23}	.915 (0.007)	
		P_{31}	.886 (0.007)	
		P_{32}	.894 (0.007)	
		P_{33}	.923 (0.007)	
		P_4	.920 (0.007)	
	Viso	P_{21}	.820 (0.010)	$F(6,1781) = 2.750$
		P_{22}	.853 (0.010)	(0.012) (0.029)
		P_{23}	.845 (0.010)	
		P_{31}	.835 (0.010)	
		P_{32}	.842 (0.010)	
		P_{33}	.805 (0.010)	
		P_4	.818 (0.010)	
	DA	P_{21}	.813 (0.009)	N.S.
		P_{22}	.841 (0.009)	
		P_{23}	.847 (0.009)	
		P_{31}	.844 (0.009)	
		P_{32}	.819 (0.009)	
		P_{33}	.828 (0.009)	
		P_4	.830 (0.009)	
	ODIp	P_{21}	.832 (0.010)	N.S.
		P_{22}	.828 (0.010)	
		P_{23}	.822 (0.010)	
		P_{31}	.831 (0.010)	
		P_{32}	.826 (0.010)	
		P_{33}	.823 (0.010)	
		P_4	.825 (0.010)	

Fig. 1. Diffusion parameter maps obtained for different shell configurations. Maps within mid-axial slice for NODDI (A) and Bingham-NODDI (B) computed on adult dataset.

and standard deviation of VF and ICC values (see Tables 5–4 of Supplementary Material). For every model parameter, toolbox performances and shell configurations of each group were compared. We analyzed results with Bland-Altman plots, where parameter difference between acquisitions was plotted against its mean value, averaged over ROIs for each subject. We exhibited outcomes from the two toolboxes, shell configurations and datasets (see Supplementary Material).

3. Results

We reported MANOVA analysis results obtained for both ICC and VF when considering supra-tentorial and infra-tentorial brain areas separately. The supra-tentorial MANOVA multivariate analysis for VF reported a significant Trace of Pillai (ToP) for all factors (SMS Acceleration factor: ToP = 0.111, F(9, 1773) = 24.475, $p < 10^{-40}$; Software: ToP = 0.396, F(9, 1773) = 129.092, $p < 10^{-187}$; Subject: ToP= 0.491, F(9, 1773) = 190.128, $p < 10^{-252}$; Configuration: ToP= 0.243, F(54, 10,668) = 8.329, $p < 10^{-62}$; Side: ToP=0.052, F(9.1773) = 10.712, $p < 10^{-16}$). The supra-tentorial MANOVA multivariate analysis for ICC reported a significant ToP for all factors (SMS Acceleration factor: ToP = 0.295, F(9, 1773) = 82.288, $p < 10^{-127}$; Software: ToP = 0.111, F(9, 1773) = 24.547, $p < 10^{-40}$; Subject: ToP = 0.421, F(9, 1773) = 142.985, $p < 10^{-203}$; Configuration: ToP = 0.153, F(54, 10,668) = 5.157, $p < 10^{-31}$; Side: ToP = 0.052, F(9,1773) = 10.813, $p < 10^{-16}$).

Focusing on adult dataset, the infra-tentorial MANOVA multivariate analysis for VF reported a significant ToP for all factors (SMS Acceleration factor: ToP= 0.245, F(9, 262)= 9.462, $p < 10^{-12}$; Software: ToP = 0.505, F(54, 262) = 29.717, $p < 10^{-35}$; Configuration: ToP= 0.485, F(54, 1602)= 2.610, $p < 10^{-9}$; Side: ToP = 0.067, F(9, 262) = 2.107, $p < 0.03$). The infra-tentorial MANOVA multivariate analysis for ICC reported a significant ToP for all factors excluding the configuration factor. (SMS Acceleration factor: ToP = 0.643, F(9, 262) = 52.357, $p < 10^{-53}$; Software: ToP= 0.099, F(9, 262) = 3.20, $p < 0.001$; Configuration: ToP = 0.240, F(54, 1602) = 1.236, ns; Side: ToP = 0.122, F(9262) = 4.027, $p < 10^{-05}$).

We also performed statistical analysis on a single dataset, obtained merging data from pediatric and adult populations. The supra-tentorial MANOVA multivariate analysis for ICC reported a significant ToP for all factors (SMS Acceleration factor: ToP= 0.486, F(9, 878)= 92.069, $p < 10^{-120}$; Software: ToP = 0.398, F(9, 878)= 64.627, $p < 10^{-91}$; Configuration: ToP = 0.321, F(54, 5298) = 5.538, $p < 10^{-34}$; Side: ToP = 0.070, F(9, 878) = 7.317, $p < 10^{-10}$). Multivariate VF analysis for complete dataset reported significant ToP for all factors (SMS Acceleration factor: ToP = 0.217, F(9878) = 27.074, $p < 10^{-41}$; Software: ToP = 0.638, F(9878) = 171.843, $p < 10^{-187}$; Configuration: ToP = 0.517, F(54, 5298) = 9.258, $p < 10^{-69}$; Side: ToP = 0.118, F(9878) = 13.049, $p < 10^{-19}$).

Moreover, statistics on models (NODDI and Bingham-NODDI) revealed no significant differences for ICC values of ODI, v_{intra} and v_{iso} between the 2 NODDI models under investigation, while VF produced NODDI ODI values significantly reduced ($p < 10^{-07}$) with respect to Bingham ODI (ODI_{tot}).

3.1. Dataset results

Due to the lack of infra-tentorial acquisition in children, the comparison between the 2 datasets was performed on the 32 supra-tentorial ROIs. Children exhibited ICC values of roughly 0.9 for both NODDI and Bingham-NODDI models (Table 1) and a VF values below 5.5% (Table 2). Higher VF values were observed only for v_{iso} (18% for NODDI v_{iso} and 15% for Bingham-NODDI v_{iso}). For adult supra-tentorial areas we found ICC results for ODI and DA with value of around 0.75 whereas v_{intra} and v_{iso} were around 0.8 (Table 1). VF in adult dataset was in the range of 2–7% in all estimated parameters but v_{iso} that had over 31% variation (Table 2). The 2 datasets show significant differences in all metrics both for ICC and VF (Fig. 2).

3.2. Shell configurations results

A comparison among the seven different shell combinations (P_{21}, P_{22}, P_{23}, P_{31}, P_{32}, P_{33} and P_4) was performedFor supra-tentorial ROIs

Table 1 (continued)

	Parameter	Type of parameter	Mean (std dev)	F (pval) (η)
INTRA-CLASS CORRELATION COEFFICIENT ICC WITHIN SUPRA-TENTORIAL ROIS				
	ODIs	P_{21}	.840 (0.010)	N.S.
		P_{22}	.842 (0.010)	
		P_{23}	.832 (0.010)	
		P_{31}	.839 (0.010)	
		P_{32}	.826 (0.010)	
		P_{33}	.823 (0.010)	
		P_4	.826 (0.010)	
	ODItot	P_{21}	.839 (0.010)	N.S.
		P_{22}	.838 (0.010)	
		P_{23}	.826 (0.010)	
		P_{31}	.834 (0.010)	
		P_{32}	.825 (0.010)	
		P_{33}	.826 (0.010)	
		P_4	.824 (0.010)	
	Vintra	P_{21}	.876 (0.007)	N.S.
		P_{22}	.889 (0.007)	
		P_{23}	.880 (0.007)	
		P_{31}	.896 (0.007)	
		P_{32}	.888 (0.007)	
		P_{33}	.891 (0.007)	
		P_4	.890 (0.007)	
	Viso	P_{21}	.823 (0.009)	$F(6,1781) = 3.664$
		P_{22}	.853 (0.009)	(1.27^{-03}) (0.012)
		P_{23}	.829 (0.009)	
		P_{31}	.870 (0.009)	
		P_{32}	.843 (0.009)	
		P_{33}	.853 (0.009)	
		P_4	.862 (0.009)	
EFFECT OF SUBJ	ODI	PED	.917 (0.005)	$F(1,1781) = 499.197$
		AD	.757 (0.005)	$(1.11e^{-97})$ (0.218)
	Vintra	PED	.962 (0.004)	$F(1,1781) = 578.017$
		AD	.838 (0.004)	$(7.59e^{-111})$ (0.245)
	Viso	PED	.885 (0.006)	$F(1,1781) = 188.927$
		AD	.778 (0.006)	$(6.19e^{-41})$ (0.095)
	DA	PED	.883 (0.005)	$F(1,1781) = 221.812$
		AD	.781 (0.005)	$(2.28e^{-47})$ (0.111)
	ODIp	PED	.897 (0.005)	$F(1,1781) = 350.597$
		AD	.757 (0.005)	$(1.48e^{-71})$ (0.164)
	ODIs	PED	.915 (0.005)	$F(1,1781) = 468.202$
		AD	.750 (0.005)	$(2.24e^{-92})$ (0.208)
	ODItot	PED	.911 (0.005)	$F(1,1781) = 460.088$
		AD	.750 (0.005)	$(5.64e^{-91})$ (0.205)
	Vintra	PED	.911 (0.004)	$F(1,1781) = 912.383$
		AD	.750 (0.004)	$(3.52e^{-162})$ (0.339)
	Viso	PED	.965 (0.005)	$F(1,1781) = 145.324$
		AD	.809 (0.005)	$(3.16e^{-32})$ (0.075)
EFFECT OF SIDE	ODI	Left	.846 (0.005)	$F(1,1781) = 6.222$
		Right	.828 (0.005)	(0.013) $(3.48e^{-03})$
	Vintra	Left	.896 (0.004)	N.S.
		Right	.904 (0.004)	
	Viso	Left	.820 (0.006)	$F(1,1781) = 8.956$
		Right	.843 (0.006)	$(2.8e^{-03})$ $(5e^{-03})$
	DA	Left	.832 (0.005)	N.S.
		Right	.832 (0.005)	
	ODIp	Left	.848 (0.005)	$F(1,1781) = 31.754$
		Right	.806 (0.005)	$(2.03e^{-08})$ (0.017)
	ODIs	Left	.835 (0.005)	N.S.
		Right	.830 (0.005)	
	ODItot	Left	.847 (0.005)	$F(1,1781) = 20.40$
		Right	.813 (0.005)	$(6.56e^{-06})$ (0.011)
	Vintra	Left	.883 (0.004)	N.S.
		Right	.892 (0.004)	
	Viso	Left	.835 (0.005)	$F(1,1781) = 13.791$
		Right	.860 (0.005)	$(2.11e^{-04})$ $(7.68e^{-03})$

we observed that shell configurations had a statistical significant effect on VF on the values of v_{intra} and v_{iso} for both NODDI models and DA. As reported in Fig. 3A, the lowest VF values for v_{intra} were obtained in the configuration P_{33} whereas the highest VF values were always obtained in the P_{21} configuration (Table 2). We also observed that shell configurations had a statistical significant effect on ICC on the values of v_{iso} for both NODDI models and v_{intra} of NODDI. Particularly, the highest ICC values were obtained in v_{intra} of NODDI in the P_{33} configuration whereas the lowest performances were obtained in P_{21} (Table 1 and Fig. 3B), that seems to be in accordance with VF results. The infra-tentorial results did not show any significant effect of shell configuration on ICC, while slightly significant results were obtained for VF on the values of DA and Bingham v_{iso}, both showing worst VF values for P_{22} configuration (Tables 1,2 of Supplementary Material).

Table 2

Mean values, standard deviation, *F*-statistics, *p*-value and size effect of VF for NODDI and Bingham-NODDI metrics to assess effect of different factors (tool, SMS, configuration and side) within 32 supra-tentorial ROIs. N.S. means not significant.

VARIATION FACTOR VF WITHIN SUPRA-TENTORIAL ROIS				
	Parameter	Type of parameter	Mean (std dev)	F (pval) (η)
EFFECT OF TOOL	ODI	MDT	4.190 (0.070)	N.S.
		cuDIMOT	4.266 (0.070)	
	Vintra	MDT	2.291 (0.044)	$F(1,1781) = 411.447$
		cuDIMOT	3.541 (0.044)	($1.81e^{-82}$) (0.188)
	Viso	MDT	34.510 (0.513)	$F(1, 1781) = 281.300$
		cuDIMOT	22.342 (0.513)	($9.85e^{-59}$) (0.136)
	DA	MDT	2.836 (0.060)	$F(1,1781) = 8.420$
		cuDIMOT	2.588 (0.060)	($3.76e^{-03}$) ($5.00e^{-03}$)
	ODIp	MDT	3.786 (0.064)	$F(1,1781) = 17.999$
		cuDIMOT	3.399 (0.064)	($2.32e^{-05}$) (0.01)
	ODIs	MDT	6.647 (0.112)	$F(1,1781) = 5.713$
		cuDIMOT	6.267 (0.112)	(0.017) ($3.00e^{-03}$)
	ODItot	MDT	5.241 (0.087)	$F(1,1781) = 12.729$
		cuDIMOT	4.800 (0.087)	($3.70e^{-04}$) ($7.00e^{-03}$)
	Vintra	MDT	2.365 (0.041)	$F(1,1781) = 119.268$
		cuDIMOT	2.997 (0.041)	($6.39e^{-27}$) (0.063)
	Viso	MDT	28.459 (0.483)	$F(1,1781) = 236.280$
		cuDIMOT	17.969 (0.483)	($3.65e^{-50}$) (0.117)
EFFECT OF SMS	ODI	SMS2	4.454 (0.070)	$F(1,1781) = 20.737$
		SMS3	4.002(0.070)	($5.63e^{-06}$) ($5.75e^{-05}$)
	Vintra	SMS2	2.906 (0.044)	N.S.
		SMS3	2.926 (0.044)	
	Viso	SMS2	25.767 (0.513)	$F(1,1781) = 53.747$
		SMS3	31.085 (0.513)	($3.44e^{-13}$) (0.029)
	DA	SMS2	2.454 (0.060)	$F(1,1781) = 36.319(2.03e^{-09})$
		SMS3	2.970 (0.060)	(0.020)
	ODIp	SMS2	3.714 (0.064)	$F(1,1781) = 7.147$
		SMS3	3.471 (0.064)	($7.58e^{-03}$) ($4.00e^{-03}$)
	ODIs	SMS2	6.917 (0.112)	$F(1,1781) = 33.582$
		SMS3	5.997 (0.112)	($8.06e^{-09}$) (0.019)
	ODItot	SMS2	5.343 (0.087)	$F(1,1781) = 27.254$
		SMS3	4.698 (0.087)	($1.99e^{-06}$) (0.015)
	Vintra	SMS2	2.563 (0.041)	$F(1,1781) = 16.564$
		SMS3	2.799 (0.041)	($4.91e^{-05}$) ($9.00e^{-03}$)
	Viso	SMS2	20.948 (0.483)	$F(1,1781) = 44.090$
		SMS3	25.480 (0.483)	($4.15e^{-11}$) (0.024)
EFFECT OF CONFIGURATIONS	ODI	P_{21}	4.285 (0.131)	N.S.
		P_{22}	4.209 (0.131)	
		P_{23}	4.262 (0.131)	
		P_{31}	3.989 (0.131)	
		P_{32}	4.263 (0.131)	
		P_{33}	4.342 (0.131)	
		P_4	4.244 (0.131)	
	Vintra	P_{21}	3.478 (0.082)	$F(6,1781) = 10.931$
		P_{22}	2.785 (0.082)	($5.51e^{-12}$) (0.036)
		P_{23}	2.808 (0.082)	
		P_{31}	2.894 (0.082)	
		P_{32}	2.990 (0.082)	
		P_{33}	2.640 (0.082)	
		P_4	2.818 (0.082)	
	Viso	P_{21}	29.797 (0.960)	$F(6,1781) = 5.739$
		P_{22}	26.770 (0.960)	($6.31e^{-06}$) (0.019)
		P_{23}	33.048 (0.960)	
		P_{31}	27.404 (0.960)	
		P_{32}	27.335 (0.960)	
		P_{33}	28.027 (0.960)	
		P_4	26.600 (0.960)	
	DA	P_{21}	2.948 (0.113)	$F(6,1781) = 3.258$
		P_{22}	2.806 (0.113)	($3.45e^{-03}$) (0.011)
		P_{23}	3.000 (0.113)	
		P_{31}	2.507 (0.113)	
		P_{32}	2.623 (0.113)	
		P_{33}	2.545 (0.113)	
		P_4	2.555 (0.113)	
	ODIp	P_{21}	3.513 (0.120)	N.S.
		P_{22}	3.817 (0.120)	
		P_{23}	3.827 (0.120)	
		P_{31}	3.369 (0.120)	
		P_{32}	3.511 (0.120)	
		P_{33}	3.594 (0.120)	
		P_4	3.517 (0.120)	

(continued on next page)

Table 2 (*continued*)

VARIATION FACTOR VF WITHIN SUPRA-TENTORIAL ROIS	Parameter	Type of parameter	Mean (std dev)	F (pval) (η)
	ODIs	P_{21}	6.266 (0.210)	N.S.
		P_{22}	6.745 (0.210)	
		P_{23}	6.895 (0.210)	
		P_{31}	6.080 (0.210)	
		P_{32}	6.402 (0.210)	
		P_{33}	6.462 (0.210)	
		P_4	6.350 (0.210)	
	ODItot	P_{21}	4.892 (0.163)	N.S.
		P_{22}	5.281 (0.163)	
		P_{23}	5.305 (0.163)	
		P_{31}	4.735 (0.163)	
		P_{32}	4.964 (0.163)	
		P_{33}	5.013 (0.163)	
		P_4	4.955 (0.163)	
	Vintra	P_{21}	3.092 (0.077)	$F(6,1781) = 9.358$
		P_{22}	2.758 (0.077)	$(3.95e^{-10})$ (0.031)
		P_{23}	2.806 (0.077)	
		P_{31}	2.525 (0.077)	
		P_{32}	2.662 (0.077)	
		P_{33}	2.366 (0.077)	
		P_4	2.560 (0.077)	
	Viso	P_{21}	25.576 (0.903)	$F(6,1781) = 10.569$
		P_{22}	26.218 (0.903)	$(1.47e^{-11})$ (0.034)
		P_{23}	24.908 (0.903)	
		P_{31}	17.567 (0.903)	
		P_{32}	23.533 (0.903)	
		P_{33}	22.949 (0.903)	
		P_4	21.748 (0.903)	
EFFECT OF SUBJ	ODI	PED	3.339 (0.070)	$F(1,1781) = 321.3020$
		AD	5.116 (0.070)	$(3.46e^{-66})$ (0.153)
	Vintra	PED	2.560 (0.044)	$F(1,1781) = 133.617$
		AD	3.272 (0.044)	$(7.48e^{-30})$ (0.070)
	Viso	PED	17.843 (0.513)	$F(1,1781) = 851.230$
		AD	39.009 (0.513)	$(2.75e^{-153})$ (0.323)
	DA	PED	2.812 (0.060)	$F(1,1781) = 5.472$
		AD	2.612 (0.060)	(0.019) $(3.00e^{-03})$
	ODIp	PED	3.151 (0.064)	$F(1,1781) = 94.198$
		AD	4.034 (0.064)	$(9.76e^{-22})$ (0.050)
	ODIs	PED	5.505 (0.112)	$F(1,1781) = 144.057$
		AD	7.409 (0.112)	$(5.70e^{-32})$ (0.075)
	ODItot	PED	4.267 (0.087)	$F(1,1781) = 149.121$
		AD	5.775 (0.087)	$(5.41e^{-33})$ (0.077)
	Vintra	PED	2.477 (0.041)	$F(1,1781) = 49.595$
		AD	2.885 (0.041)	$(2.69e^{-12})$ (0.027)
	Viso	PED	15.266 (0.483)	$F(1,1781) = 542.618$
		AD	31.162 (0.483)	$(5.47e^{-105})$ (0.234)
EFFECT OF SIDE	ODI	Left	4.351 (0.070)	$F(1,1781) = 6.173$
		Right	4.105 (0.070)	(0.013) $(3.00e^{-03})$
	Vintra	Left	3.089 (0.044)	$F(1,1781) = 31.443$
		Right	2.743 (0.044)	$(2.38e^{-08})$ (0.017)
	Viso	Left	29.664 (0.513)	$F(1,1781) = 11.646$
		Right	27.188 (0.513)	$(6.58e^{-04})$ $(6.00e^{-03})$
	DA	Left	2.787 (0.060)	N.S.
		Right	2.637 (0.060)	
	ODIp	Left	3.637 (0.064)	N.S.
		Right	3.547 (0.064)	
	ODIs	Left	6.488 (0.112)	N.S.
		Right	6.426 (0.112)	
	ODItot	Left	5.018 (0.087)	N.S.
		Right	5.023 (0.087)	
	Vintra	Left	2.819 (0.041)	$F(1,1781) = 22.743$
		Right	2.543 (0.041)	$(2.00e^{-06})$ (0.013)
	Viso	Left	24.383 (0.483)	$F(1,1781) = 11.733$
		Right	22.045 (0.483)	$(6.28e^{-04})$ $(7.00e^{-03})$

3.3. Acceleration factor results

We then assessed effects of AF on reliability results for both adult and pediatric datasets (Fig. 4). Focusing on supra-tentorial analysis, we found that SMS3 ICC values were significantly higher compared to those in SMS2 for all parameters (Table 1). Concerning the VF, we observed that SMS2 VF values were significantly lower compared to those in SMS3

for v_{iso} in both models, v_{intra} of Bingham as well as DA (Table 2). Conversely, ODI VF of both models were significantly higher when comparing SMS2 vs SMS3. The infra-tentorial ICC confirmed evidences of supra-tentorial results, showing that ICC values were significantly higher when comparing SMS3 vs SMS2 for all diffusion parameters (Table 1 of Supplementary Material). The infra-tentorial results showed that VF values
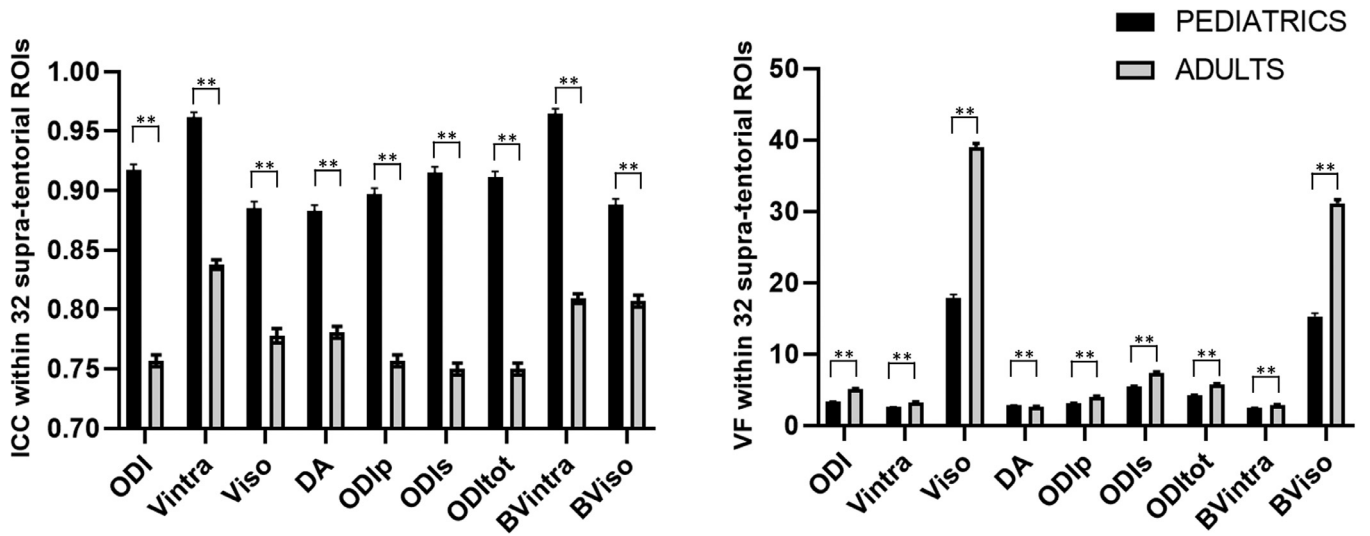
Fig. 2. Boxplots of ICC (left) and VF (right) values within each dataset (pediatrics vs adults) for all diffusion parameter. Asterisks highlight significant differences between datasets.
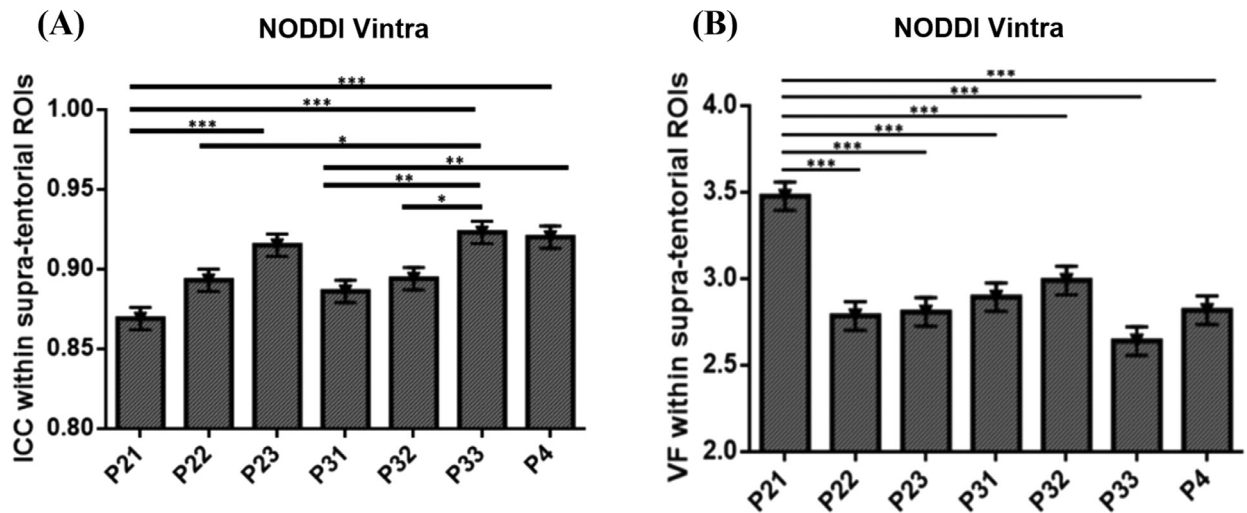

Fig. 3. Boxplots of ICC (left) and VF (right) values within each shell configuration (P_{21}, P_{22}, P_{23}, P_{31}, P_{32}, P_{33} and P_4) for NODDI v_{intra} parameter. Asterisks highlight significant differences between shell configurations.
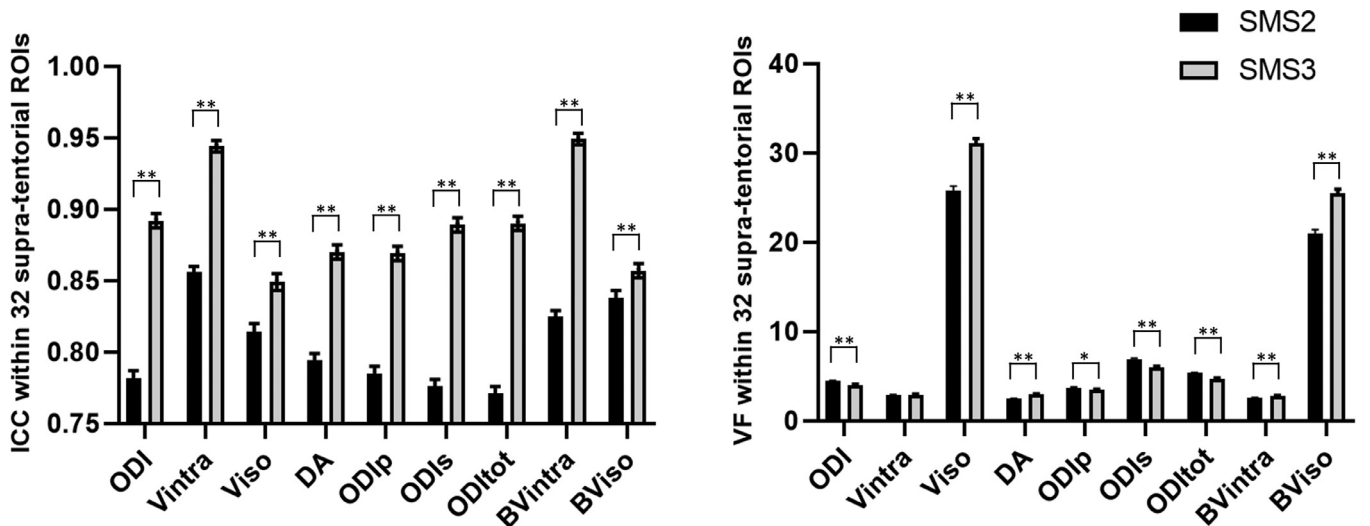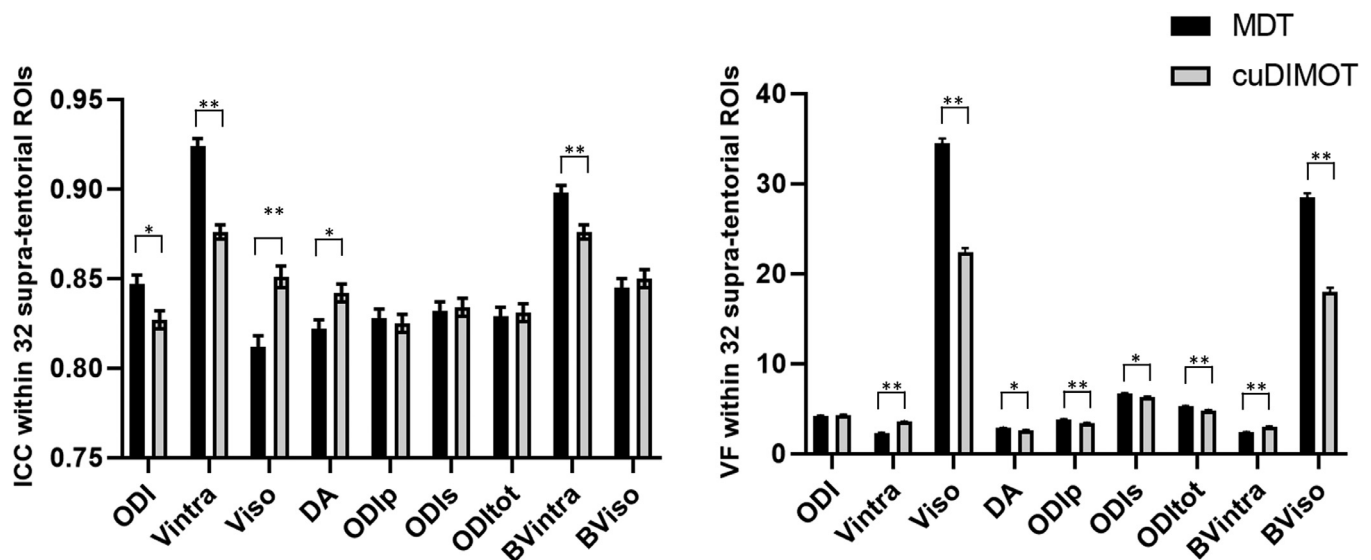

Fig. 4. Boxplots of ICC (left) and VF (right) values within each acceleration factor (SMS2 vs SMS3) for all diffusion parameter. Asterisks highlight significant differences between acceleration factors.
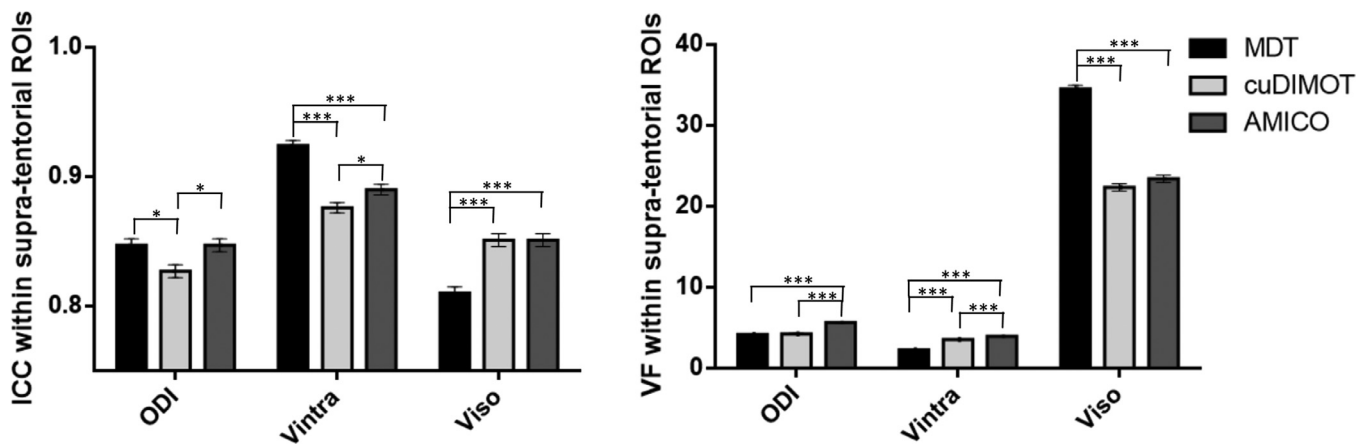
Fig. 5. Boxplots of ICC (left) and VF (right) values within each fitting tool (MDT vs cuDIMOT) for all diffusion parameter. Asterisks highlight significant differences between fitting tools.

Table 3
Average runtimes for NODDI and bingham-NODDI model fitting on oracle cloud server.

| | cuDIMOT | | MDT | |
	NODDI	Bingham-NODDI	NODDI	Bingham-NODDI
P_{21}	2 min 5s	8 min 57s	35s	3 min 44s
P_{22}	2 min 28s	12 min 22s	41s	4 min 14s
P_{23}	2 min 35s	12 min 32s	45 s	4 min 27s
P_{31}	2 min 48s	12 min 47s	35s	4 min 25s
P_{32}	3 min 2s	12 min 50s	37s	4 min 32s
P_{33}	3 min 40s	16 min 10s	44s	4 min 42s
P_4	4 min 20s	16 min 25s	42s	4 min 50s

were significantly different for ODI and v_{intra} of both models, with lower values for SMS3 (Table 2 of Supplementary Material).

3.4. Fitting toolboxes results

Since the fitting was performed with two different GPU-based tools, we compared their reliability results (Fig. 5). Regarding the supra-tentorial data, a significant VF difference resulted for all the metrics, but not for NODDI ODI. In particular, MDT revealed significantly higher VF values than cuDIMOT toolbox in all metrics, except for v_{intra} of both NODDI models (Table 2). ICC values were significantly higher in v_{intra} and ODI and lower in v_{iso} and DA when comparing MDT to cuDIMOT (Table 1). The infra-tentorial data results showed a strongly significant difference in VF values for v_{intra} and v_{iso} of both models, as well as DA, ODI_p and ODI_{tot}, where MDT revealed higher VF values than cuDIMOT (Table 2 of Supplementary Material). The effect of tool revealed no significant ICC results for all diffusion parameters excluding Bingham v_{iso}, where MDT produced higher ICC than cuDIMOT (Table 1 of Supplementary Material). Beyond that, an interesting difference between cuDIMOT and MDT was the fitting runtime (Table 3). On Oracle Cloud server setup (two NVIDIA® Tesla® P100, four Intel® Xeon® Gold 5120 CPUs @ 2.20 GHz, 187 GB RAM), MDT performed fitting 400% faster than cuDIMOT for Bingham-NODDI and up to 600% faster for NODDI. NODDI models was also performed on AMICO toolbox (Table 4). Assessing the effect of tool, we found significantly increased ICC values for AMICO when compared to cuDIMOT in ODI and v_{intra} metrics, while no differences were found in v_{iso}. Compared to MDT, AMICO revealed better ICC values in v_{iso}, while significantly reduced values appeared in v_{intra} (Fig. 6). VF assessment reported significantly higher values in AMICO

when compared to both MDT and cuDIMOT in ODI and v_{intra} metrics (Fig. 6).

3.5. Hemispheric side results

Based on recent evidences of hemispheric asymmetries in the NODDI signal (Schmitz et al., 2019), we computed diffusion metrics within left and right ROIs, comparing their reliability results (Fig. 7). For supra-tentorial ROIs, we observed that hemispheric side had a significant effect on ICC on the values of ODI and v_{iso} of both models, with higher ICC values respectively on left and right sides (Table 1). Concerning the VF, significant differences resulted for v_{intra} and v_{iso} of both models and NODDI ODI, always with higher VF values for left ROIs (Table 2). The infra-tentorial data confirmed VF significant results only for ODI, again with higher values for left ROIs (Table 2 of Supplementary Material), while ICC results showed significant differences for Bingham ODI_p and v_{intra}, where ICC was higher respectively within right and left ROIs.

4. Discussion

The study investigated test-retest reliability of diffusion MRI metrics computed for different acquisition setting, with the aim to assess whether fast acquisition and processing setting could help to translate NODDI within the clinical setting.. To this purpose, we assessed reliability for 7 different shell configurations and 2 AF. We also analyzed the data with 2 different GPU-based tools and compared the results. The study made use of 2 metrics to assess reliability of the measurements, ICC and VF.

Excellent reliability in terms of ICC was observed for all diffusion parameters, especially for v_{intra} (see Tables 1 and boxplots in the Sup-
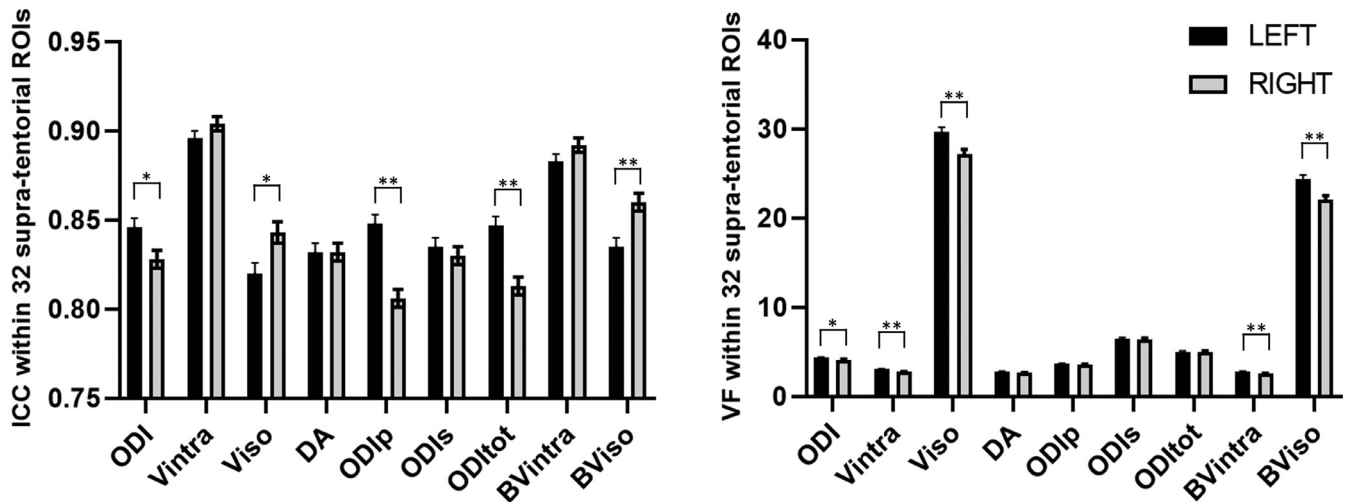
Fig. 7. Boxplots of ICC (left) and VF (right) values within each brain side (left vs right) for all diffusion parameter. Asterisks highlight significant differences between fitting tools.

consequently they had better data alignment when compared to adults, which led to an improved data stability. Moreover, in order to reduce patient discomfort and to collect images as fast as possible, acquisition was limited to the supra-tentorial areas and no additional volumes with reversed phase-encoding direction were acquired. The lack of these acquisitions prevented children to be preprocessed with the standard correction for susceptibility-induced distortion correction. Conversely, adult volunteers received a longer acquisition for a whole-brain coverage further corrected for susceptibility-induced distortion. During ROI-based analysis, supra-tentorial and infra-tentorial areas were considered separately for reliability assessment in adult dataset. Average reliability was usually better in supra-tentorial areas (Table 1) than in infra-tentorial areas (Table 1 of Supplementary Material). In particular, we recorded global ICC reduction for all the metrics when analysis was performed within the 8 infra-tentorial ROIs. The infra-tentorial data are indeed more affected by noise and more prone to fitting errors compared to the supra-tentorial ROIs.

4.2. Shell configurations

Comparing shell configurations, very small reliability differences appeared and the only parameters that were affected by the configuration were v_{iso} of both models and NODDI v_{intra}. Referring to the Tables 1 and Table 2, v_{intra} is better reproduced between the 2 sessions in the configuration P_{33} whereas the performance dropped in the case of P_{21} configuration, both in terms of ICC and VF (Fig. 3). This trend suggests more extensive diffusion parameters variability when considering shell configuration with lower b-values (e.g. 300 s/mm^2), especially for simple model like NODDI. The ODI from both models were not affected by the configurations and that finding is in line with previous results showing that orientation dispersion index is also well estimated with a single shell whereas it is more sensitive to the number of gradient directions. In our case, fewer shell ODI results were comparable in term of reliability to those with up to 4 shells as it was already confirmed in a study (Timmers et al., 2016) showing that single shell ODI analysis was able to reproduce the group differences from multi-shell analysis. Furthermore, our results seems to be in line with ODI value behavior seen in a recent study Parvathaneni et al., 2018. The authors showed how in WM the acquisitions with a single shell with $b = 1000/2000s/mm^2$ outperform in term of root mean square error (RMSE) the sequences acquired with 2 shells. Although we did not perform single shell acquisition, our results revealed a trend of increased ICC for lower number of shells (i.e.

two-shell configuration) in ODI (see Table 1). Moreover, similar results were obtained for DA, an evidence that might suggest the ability to reliably fit the data even for more complex models such as Bingham model. Conversely, our finding confirmed the sensitivity of v_{intra} to the choice of outer shell b-value (Zhang et al., 2012; Parvathaneni et al., 2018), since for equal number of shells, the ICC appeared increased when considering higher b-values.

4.3. Acceleration factors

The study observed that SMS3 ICC values were significantly higher compared to those in SMS2 for all diffusion parameters, while there was not a clear winner when looking at VF, that exhibited very high values for v_{iso} both in SMS2 and SMS3 (Fig. 4). This parameter was already shown to be unreliably fitted elsewhere and we might hypothesize how the deviance from Gaussian noise towards a more non-central chi distribution in case of multiband acquisition (St-Jean et al., 2020) could mostly affect the model fitting in case of higher acceleration factors as due to their lower SNR. Better alignment of the data along different directions and different shells, as in the case of shorter acquisition time might also be responsible for more reliable performances. The higher acceleration factor protocol was indeed 5 min shorter than the one with lower acceleration and then the data were more prone to having motion misalignment. Given the dependence on the number of directions for diffusion parameters, the lower SMS2 ICC values are likely due to this motion artifact only partially corrected via post-processing. In this context, head motion estimates corroborates the idea of an overall slightly larger motion artefact in case of lower acceleration factor (see Fig. 11 in Supplementary Material). The effect of the head motion can be disruptive of the reliability and it seems to be the main factor as due to slice acceleration when looking ICC and VF (see Figs. 10–12 of Supplementary Material)

VF values for v_{intra}, v_{iso} and DA were higher in SMS3 compared to SMS2, probably due to the higher acquisition noise of faster acquisition. In this context, we hypothesized a possible competitive effect between acquisition noise and motion artifacts, which could be evaluated when looking at the two datasets separately. Pediatric subjects were less prone to motion artifacts since they did not exit the scanner, thus their reliability values mainly account for acquisition noise. As we can see from Supplementary Material (Tables 5 and 7), although ICC values were always higher for SMS3 in both cohorts, the gap between different AF was prominent in adults rather than children, reflecting the beneficial effect
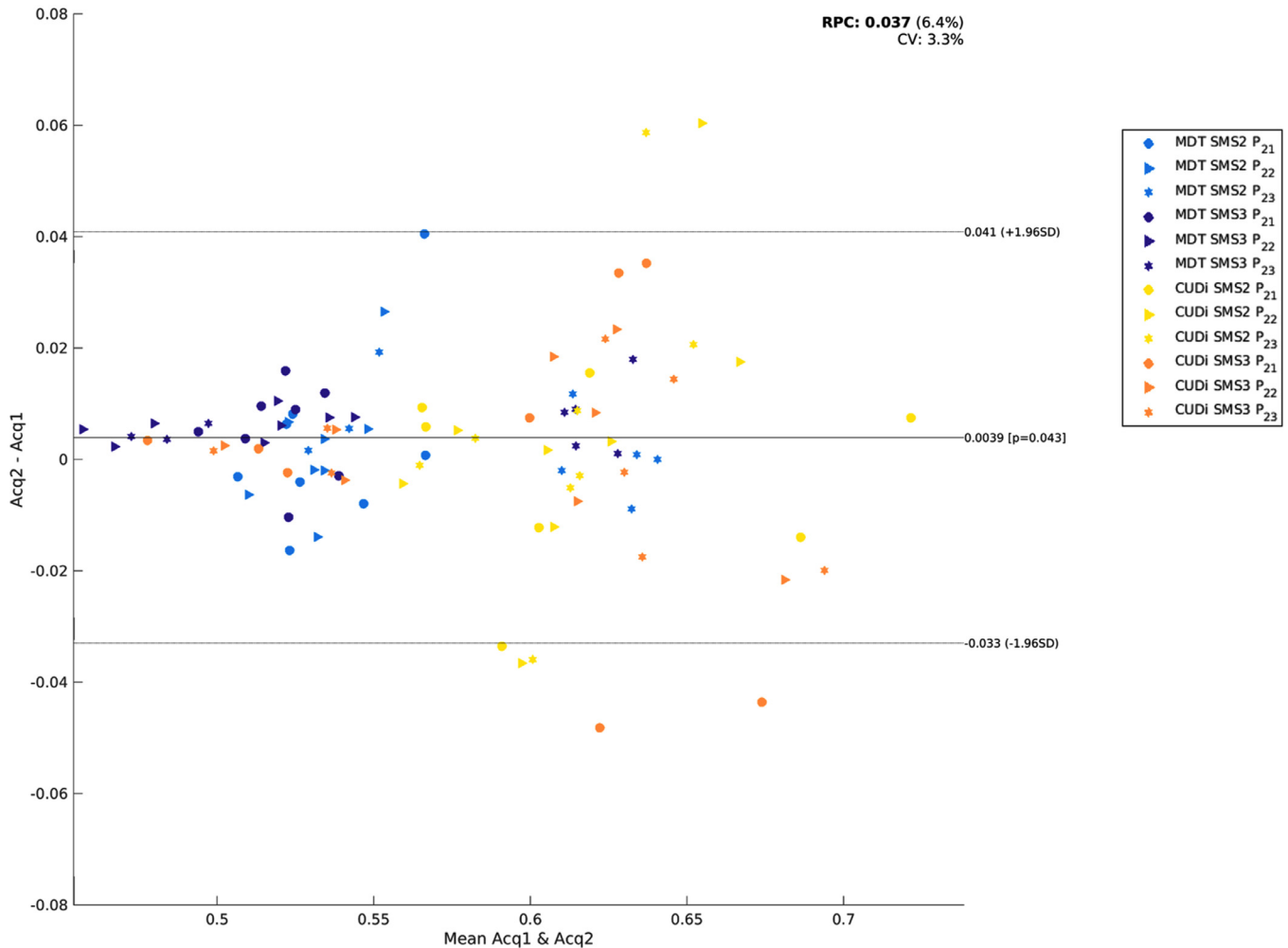
Fig. 8. NODDI v_{intra} Bland-Altman plot for two-shell configurations adult dataset. CuDIMOT values are more spread than the ones from MDT. Each point in the plot represents one subject, averaged over 16 supra-tentorial ROIs of right hemisphere.

of faster acquisition over motion artifacts which indeed was strongly reduced in children. Focusing on diffusion parameter significantly affected by AF, we found that SMS3 produced significantly higher VF values in all parameter for children and in v_{intra}, v_{iso} and DA of adults (Tables 6 and 8 of Supplementary Material), probably related to higher inter-subject variability caused by acquisition noise of faster protocol.

4.4. Fitting toolboxes

Bland-Altman plot in Fig. 8 showed two clusters of results in adult datasets for v_{intra}. The data fitted with cuDIMOT had higher values and were more spread in the plot. Most of the spread difference between SMS2 and SMS3 was actually observed in cuDIMOT estimations whereas MDT fitted parameters showed to be more clustered. Interestingly, CuDIMOT's VFs result in being generally lower than the MDT counterpart, whereas mean values for each metric is actually higher (see Bland-Altman plots in Supplementary Material). The latter might cause the lower ICC seen thus leading to a less reliability in terms of how well the method can discriminate between individuals. We do not have a clear idea regarding the reasons why there is this discrepancy in the results for v_{iso} between the 2 toolboxes. Both the toolboxes use iter-

ative optimization algorithms for the solution of non-linear least square problems. Even though they do not use the same way to reach the minimum (Powell for MDT and Levenberg-Marquardt for cuDIMOT), their approaches are very similar as they both interpolates between Gaussian-Newton and gradient descendent and this may not be a reasonable cause of this discrepancy. Eventually, we observed that boundary parameters are tighter in MDT than cuDIMOT and this could give rise to more variability and overestimation of metrics for the latter. The use of different optimization methods might likely influence the analysis time, which in MDT resulted in being faster. Furthermore, both toolboxes have a cascade method to fit complex models starting from simpler ones, and the way this is applied is similar in the two toolboxes. Furthermore, this difference occurs in both datasets and acceleration factors, so it appears to be independent from either misalignment and motion artifacts, or noise and SNR. Only MDT has the option to model noise distribution (Gaussian, Offset-Gaussian and Rician) and As far as we know, offset Gaussian provides stable results (Panagiotaki et al., 2012) when compared to other noise distribution and that could explain better reliability in terms of ICC for MDT compared to cuDIMOT for ODI and v_{intra} parameters. Conversely, we did observe that MDT also showed to have a higher VF and lower ICC compared to cuDIMOT for v_{iso} and DA (Fig. 5).

Focusing on v_{iso}, the effect of noise standard deviation estimates on VF, we found that autodMRI noise standard deviation produced better results, almost comparable with those from cuDIMOT. Since the poorer reliability performances seen in MDT for v_{iso} are not well understood, these results could suggest that more precise noise standard deviation estimate could be a possible solution for that. Despite the use of OpenCL in MDT and CUDA in cuDIMOT better performances have been seen in terms of run time for MDT vs cuDIMOT. This is not in line with what is reported in the manuscript of the author of cuDIMOT who concludes that cuDIMOT would potentially achieve better perfomances for Nvidia GPU cards.

4.5. Hemispheric side

Recent studies performed NODDI to investigate possible microstructural hemispheric asymmetries across the whole brain of adult (Schmitz et al., 2019), children (Dimond et al., 2020) and infants (Dean et al., 2017). Particularly, they observed both leftward and rightward microstructural asymmetries over 22.78%, 12.78% and 6.11% of the overall brain area respectively for v_{intra}, ODI and v_{iso} (Schmitz et al., 2019). Multivariate analysis on diffusion data from right-handed subjects revealed significant effect of hemispheric side in ODI and v_{iso} reliability (Fig. 7), thus suggesting the role of brain microstructure in reliability estimation.

Having said that, we also want to point out that the present study suffers from a few limitations. First of all, even though the study has an overall larger number of subjects compared to few other studies on the topic (Andica et al., 2020; Tariq, 2018), relatively few subjects per acquisition were actually attained thus making the inference less robust and the results less interpretable. In this context, we provided reliability information assessing ICC and VF statistics on a single dataset (see Tables 3,4 of Supplementary Material), where pediatric and adult data were merged, but future studies should address it to strengthen the study, recruiting a large number of adult subjects. Moreover, in this study we used the default mode of NODDI fitting, thus neglecting that optimised intrinsic diffusivity in NODDI might be different in children and adults (Guerrero et al., 2019). This aspect need to be further investigated in future studies. Second limitation of this study is the heterogeneity of the acquired data, with differences between children and adult datasets. Since pediatric data were much more difficult to acquire, a few sacrifices have been made in terms of protocol: only supra-tentorial brain part was actually sampled and no additional volumes with reversed phase-encoding direction were available for children, thus preventing them to be pre-processed for susceptibility-induced distortions corrections. Moreover, children did not exit the scan between test and retest acquisitions, while adults were taken out of the scanner for about 20 min.

5. Conclusion

In summary, we have applied the NODDI and Bingham NODDI models to a cohort of pediatric and adult subjects and showed the effects of different acquisitions and methods on the reliability of NODDI metrics. We observed how different metrics show different patterns when looking at the effects of multiband factors and shell configurations. The main achievement of the present study is to show how 10 min NODDI acquisition with 3 shells can have reliable results in WM. More complex models do not appear to be more prone to less data acquisition as well as noisier data thus stressing the idea of Bingham-NODDI having greater sensitivity to true subject variability. Multiband acquisition did not result in worsening the reliability; conversely, shortening the acquisition might be beneficial for attenuating the motion artifacts. Lastly, we also performed a test to study GPU analysis by comparing two GPU-boosted analysis toolboxes mainly showing that faster fittings does not jeopardize the reliability when compared to more stable methods.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2021.118234.

References

A. Salarian (2020). Intraclass correlation coefficient (ICC) (https://www.mathworks.com/matlabcentral/fileexchange/22099-intraclass-correlation-coefficient-icc), MATLAB Central File Exchange.

A. Albi, O. Pasternak, L. Minati, M. Marizzoni, D. Bartrés-faz, N. Bargalló, B. Bosch, P.M. Rossini, C. Marra, B. Müller, 2018. Multisite Study of Healthy Elderly Subjects 38, 12–26. 10.1002/hbm.23350.Free

Andersson, J.L.R., Skare, S., Ashburner, J., 2003. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. Neuroimage 20, 870–888. doi:10.1016/S1053-8119(03)00336-7.

Andica, C., Kamagata, K., Hayashi, T., Hagiwara, A., Uchida, W., Saito, Y., Kamiya, K., Fujita, S., Akashi, T., Wada, A., Abe, M., Kusahara, H., Hori, M., Aoki, S., 2020. Scan–rescan and inter-vendor reproducibility of neurite orientation dispersion and density imaging metrics. Neuroradiology 62, 483–494. doi:10.1007/s00234-019-02350-6.

Boekel, W., Forstmann, B.U., Keuken, M.C., 2017. A test-retest reliability analysis of diffusion measures of white matter tracts relevant for cognitive control. Psychophysiology 54, 24–33. doi:10.1111/psyp.12769.

Braun, U., Plichta, M.M., Esslinger, C., Sauer, C., Haddad, L., Grimm, O., Mier, D., Mohnke, S., Heinz, A., Erk, S., Walter, H., Seiferth, N., Kirsch, P., Meyer-Lindenberg, A., 2012. Test-retest reliability of resting-state connectivity network characteristics using fMRI and graph theoretical measures. Neuroimage 59, 1404–1412. doi:10.1016/j.neuroimage.2011.08.044.

Buchanan, C.R., Pernet, C.R., Gorgolewski, K.J., Storkey, A.J., Bastin, M.E., 2014. Test-retest reliability of structural brain networks from diffusion MRI. Neuroimage 86, 231–243. doi:10.1016/j.neuroimage.2013.09.054.

Bujang, M.A., Baharum, N., 2017. A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. Arch. Orofac. Sci. 12, 1–11.

By, S., Xu, J., Box, B.A., Bagnato, F.R., Smith, S.A., 2017. Application and evaluation of NODDI in the cervical spinal cord of multiple sclerosis patients. NeuroImage Clin. 15, 333–342. doi:10.1016/j.nicl.2017.05.010.

Caverzasi, E., Papinutto, N., Castellano, A., Zhu, A.H., Scifo, P., Riva, M., Bello, L., Falini, A., Bharatha, A., Henry, R.G., 2016. Neurite orientation dispersion and density imaging color maps to characterize brain diffusion in neurologic disorders. J. Neuroimaging 26, 494–498. doi:10.1111/jon.12359.

Chang, Y.S., Owen, J.P., Pojman, N.J., Thieu, T., Bukshpun, P., Wakahiro, M.L.J., Berman, J.I., Roberts, T.P.L., Nagarajan, S.S., Sherr, E.H., Mukherjee, P., 2015. White matter changes of neurite density and fiber orientation dispersion during human brain maturation. PLoS ONE 10. doi:10.1371/journal.pone.0123656.

Cheng, J., Grossman, M., McKercher, T., 2013. Professional CUDA C Programming. J. Chem. Inf. Model doi:10.1017/CBO9781107415324.004.

Chung, A.W., Seunarine, K.K., Clark, C.A., 2016. NODDI reproducibility and variability with magnetic field strength: a comparison between 1.5 T and 3 T. Hum. Br. Mapp. 37, 4550–4565. doi:10.1002/hbm.23328.

Daducci, A., Canales-Rodríguez, E.J., Zhang, H., Dyrby, T.B., Alexander, D.C., Thiran, J.P., 2015. Accelerated microstructure imaging via convex optimization (AMICO) from diffusion MRI data. Neuroimage 105, 32–44. doi:10.1016/j.neuroimage.2014.10.026.

Dean, D.C., Planalp, E.M., Wooten, W., Adluru, N., Kecskemeti, S.R., Frye, C., Schmidt, C.K., Schmidt, N.L., Styner, M.A., Goldsmith, H.H., Davidson, R.J., Alexander, A.L., 2017. Mapping white matter microstructure in the one month human brain. Sci. Rep. 7, 1–14. doi:10.1038/s41598-017-09915-6.

Dimond, D., Heo, S., Ip, A., Rohr, C.S., Tansey, R., Graff, K., Dhollander, T., Smith, R.E., Lebel, C., Dewey, D., Connelly, A., Bray, S., 2020. Maturation and interhemispheric asymmetry in neurite density and orientation dispersion in early childhood. Neuroimage 221. doi:10.1016/j.neuroimage.2020.117168.

Duan, F., Zhao, T., He, Y., Shu, N., 2015. Test-retest reliability of diffusion measures in cerebral white matter: a multiband diffusion MRI study. J. Magn. Reson. Imaging 42, 1106–1116. doi:10.1002/jmri.24859.

Fu, X., Shrestha, S., Sun, M., Wu, Q., Luo, Y., Zhang, X., Yin, J., Ni, H., 2020. Microstructural white matter alterations in mild cognitive impairment and alzheimer's disease: study based on neurite orientation dispersion and density imaging (NODDI). Clin. Neuroradiol. 30, 569–579. doi:10.1007/s00062-019-00805-0.

Giraudeau, B., 1996. Negative values of the intraclass correlation coefficient are not theoretically possible. J. Clin. Epidemiol. 49, 1205–1206. doi:10.1016/0895-4356(96)00053-4.

Granberg, T., Fan, Q., Treaba, C.A., Ouellette, R., Herranz, E., Mangeat, G., Louapre, C., Cohen-Adad, J., Klawiter, E.C., Sloane, J.A., Mainero, C., 2017. In vivo characterization of cortical and white matter neuroaxonal pathology in early multiple sclerosis. Brain 140, 2912–2926. doi:10.1093/brain/awx247.

Grussu, F., Schneider, T., Zhang, H., Alexander, D.C., Wheeler-Kingshott, C.A.M., 2015. Neurite orientation dispersion and density imaging of the healthy cervical spinal cord in vivo. Neuroimage 111, 590–601. doi:10.1016/j.neuroimage.2015.01.045.

Guerrero, J.M., Adluru, N., Bendlin, B.B., Goldsmith, H.H., Schaefer, S.M., Davidson, R.J., Kecskemeti, S.R., Zhang, H., Alexander, A.L., 2019. Optimizing the intrinsic parallel diffusivity in NODDI: an extensive empirical evaluation. PLoS ONE 14, 1–17. doi:10.1371/journal.pone.0217118.

Harms, R.L., Fritz, F.J., Tobisch, A., Goebel, R., Roebroeck, A., 2017. Robust and fast nonlinear optimization of diffusion MRI microstructure models. Neuroimage 155, 82–96. doi:10.1016/j.neuroimage.2017.04.064.

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. doi:10.1016/j.neuroimage.2018.12.015.

Hodkinson, D.J., Krause, K., Khawaja, N., Renton, T.F., Huggins, J.P., Vennart, W., Thacker, M.A., Mehta, M.A., Zelaya, F.O., Williams, S.C.R., Howard, M.A., 2013. Quantifying the test-retest reliability of cerebral blood flow measurements in a clinical model of on-going post-surgical pain: a study using pseudo-continuous arterial spin labelling. NeuroImage Clin. 3, 301–310. doi:10.1016/j.nicl.2013.09.004.

Huber, E., Henriques, R.N., Owen, J.P., Rokem, A., Yeatman, J.D., 2019. Applying microstructural models to understand the role of white matter in cognitive development. Dev. Cogn. Neurosci. 36, 100624. doi:10.1016/j.dcn.2019.100624.

Jovicich, J., Marizzoni, M., Bosch, B., Bartrés-Faz, D., Arnold, J., Benninghoff, J., Wiltfang, J., Roccatagliata, L., Picco, A., Nobili, F., Blin, O., Bombois, S., Lopes, R., Bordet, R., Chanoine, V., Ranjeva, J.P., Didic, M., Gros-Dagnac, H., Payoux, P., Zoccatelli, G., Alessandrini, F., Beltramello, A., Bargalló, N., Ferretti, A., Caulo, M., Aiello, M., Ragucci, M., Soricelli, A., Salvadori, N., Tarducci, R., Floridi, P., Tsolaki, M., Constantinidis, M., Drevelegas, A., Rossini, P., Marra, C., Otto, J., Reiss-Zimmermann, M., Hoffmann, K.T., Galluzzi, S., Frisoni, G.B., 2014. Multisite longitudinal reliability of tract-based spatial statistics in diffusion tensor imaging of healthy elderly subjects. Neuroimage 101, 390–403. doi:10.1016/j.neuroimage.2014.06.075.

Kamiya, K., Hori, M., Aoki, S., 2020. NODDI in clinical research. J. Neurosci. Methods 346, 108908. doi:10.1016/j.jneumeth.2020.108908.

Kodiweera, C., Alexander, A.L., Harezlak, J., McAllister, T.W., Wu, Y.C., 2016. Age effects and sex differences in human brain white matter of young to middle-aged adults: a DTI, NODDI, and q-space study. Neuroimage 128, 180–192. doi:10.1016/j.neuroimage.2015.12.033.

Koo, T.K., Li, M.Y., 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J. Chiropr. Med. 15, 155–163. doi:10.1016/j.jcm.2016.02.012.

Kunz, N., Zhang, H., Vasung, L., O'Brien, K.R., Assaf, Y., Lazeyras, F., Alexander, D.C., Hüppi, P.S., 2014. Assessing white matter microstructure of the newborn with multi-shell diffusion MRI and biophysical compartment models. Neuroimage 96, 288–299. doi:10.1016/j.neuroimage.2014.03.057.

Lerma-Usabiaga, G., Mukherjee, P., Ren, Z., Perry, M.L., Wandell, B.A., 2019. Replication and generalization in applied neuroimaging. Neuroimage 202, 116048. doi:10.1016/j.neuroimage.2019.116048.

Lynch, K.M., Cabeen, R.P., Toga, A.W., Clark, K.A., 2020. Magnitude and timing of major white matter tract maturation from infancy through adolescence with NODDI. Neuroimage 212, 116672. doi:10.1016/j.neuroimage.2020.116672.

Mah, A., Geeraert, B., Lebel, C., 2017. Detailing neuroanatomical development in late childhood and early adolescence using NODDI. PLoS ONE 12, 1–16. doi:10.1371/journal.pone.0182340.

McCunn, P., Gilbert, K.M., Zeman, P., Li, A.X., Strong, M.J., Khan, A.R., Bartha, R., 2019. Reproducibility of neurite orientation dispersion and density imaging (NODDI) in rats at 9.4 tesla. PLoS ONE 14, 1–14. doi:10.1371/journal.pone.0215974.

McGraw, K.O., Wong, S.P., 1996. Forming inferences about some intraclass correlation coefficients. Psychol. Methods 1, 30–46. doi:10.1037/1082-989X.1.1.30.

Mustafi, S., Harezlak, J., Kodiweera, C., Randolph, J., Ford, J., Wishart, H., Wu, Y.C., 2019. Detecting white matter alterations in multiple sclerosis using advanced diffusion magnetic resonance imaging. Neural. Regen. Res. 14, 114–123. doi:10.4103/1673-5374.243716.

Newman, B.T., Dhollander, T., Reynier, K.A., Panzer, M.B., Druzgal, T.J., 2019. Test-retest reliability and long-term stability of 3-tissue constrained spherical deconvolution methods for analyzing diffusion MRI data. Biorxiv, 764506 doi:10.1101/764506.

Palacios, E.M., Owen, J.P., Yuh, E.L., Wang, M.B., Vassar, M.J., Ferguson, A.R., Diaz-Arrastia, R., Giacino, J.T., Okonkwo, D.O., Robertson, C.S., Stein, M.B., Temkin, N., Jain, S., McCrea, M., MacDonald, C.L., Levin, H.S., Manley, G.T., Mukherjee, P., 2020. The evolution of white matter microstructural changes after mild traumatic brain injury: a longitudinal DTI and NODDI study. Sci. Adv. 6, 1–12. doi:10.1126/sciadv.aaz6892.

Panagiotaki, E., Schneider, T., Siow, B., Hall, M.G., Lythgoe, M.F., Alexander, D.C., 2012. Compartment models of the diffusion MR signal in brain white matter: a taxonomy and comparison. Neuroimage 59, 2241–2254. doi:10.1016/j.neuroimage.2011.09.081.

Papinutto, N.D., Maule, F., Jovicich, J., 2013. Reproducibility and biases in high field brain diffusion MRI: an evaluation of acquisition and analysis variables. Magn. Reson. Imaging 31, 827–839. doi:10.1016/j.mri.2013.03.004.

Parvathaneni, P., Nath, V., Blaber, J.A., Schilling, K.G., Hainline, A.E., Mojahed, E., Anderson, A.W., Landman, B.A., 2018. Empirical reproducibility, sensitivity, and optimization of acquisition protocol, for neurite orientation dispersion and density imaging using AMICO. Magn. Reson. Imaging 50, 96–109. doi:10.1016/j.mri.2018.03.004.

Rousson, V., Gasser, T., Seifert, B., 2002. Assessing intrarater, interrater and test-retest reliability of continuous measurements. Stat. Med. 21, 3431–3446. doi:10.1002/sim.1253.

Schmitz, J., Fraenz, C., Schlüter, C., Friedrich, P., Jung, R.E., Güntürkün, O., Genç, E., Ocklenburg, S., 2019. Hemispheric asymmetries in cortical gray matter microstructure identified by neurite orientation dispersion and density imaging. Neuroimage 189, 667–675. doi:10.1016/j.neuroimage.2019.01.079.

Schneider, T., Brownlee, W., Zhang, H., Ciccarelli, O., Miller, D.H., Wheeler-Kingshott, C.G., 2017. Sensitivity of multi-shell NODDI to multiple sclerosis white matter changes: a pilot study. Funct. Neurol. 32, 97–101. doi:10.11138/FNeur/2017.32.2.097.

Setsompop, K., Cohen-Adad, J., Gagoski, B.A., Raij, T., Yendiki, A., Keil, B., Wedeen, V.J., Wald, L.L., 2012. Improving diffusion MRI using simultaneous multi-slice echo planar imaging. Neuroimage 23, 1–7. doi:10.1038/jid.2014.371.

Shrout, P.E., Fleiss, J.L., 1979. Intraclass correlations: uses in assessing rater reliability. Psychol. Bull. 86, 420–428. doi:10.1037/0033-2909.86.2.420.

Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E., Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M.Z., Matthews, P.M., Behrens, T.E.J., 2006. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. Neuroimage 31, 1487–1505. doi:10.1016/j.neuroimage.2006.02.024.

Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J.M., Matthews, P.M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23, S208–S219. doi:10.1016/j.neuroimage.2004.07.051.

Sone, D., Sato, N., Ota, M., Maikusa, N., Kimura, Y., Matsuda, H., 2018. Abnormal neurite density and orientation dispersion in unilateral temporal lobe epilepsy detected by advanced diffusion imaging. NeuroImage Clin. 20, 772–782. doi:10.1016/j.nicl.2018.09.017.

St-Jean, S., De Luca, A., Tax, C.M.W., Viergever, M.A., Leemans, A., 2020. Automated characterization of noise distributions in diffusion MRI data. Med. Image Anal. 65, 101758. doi:10.1016/j.media.2020.101758.

M. Tariq, 2018. In vivo quantification of complex neurite configurations using magntic resonance imaging.

Tariq, M., Schneider, T., Alexander, D.C., Gandini Wheeler-Kingshott, C.A., Zhang, H., 2016. Bingham-NODDI: mapping anisotropic orientation dispersion of neurites using diffusion MRI. Neuroimage 133, 207–223. doi:10.1016/j.neuroimage.2016.01.046.

Tariq, M., Schneider, T., Alexander, D.C., Wheeler-Kingshott, C.a.M., Zhang, H., 2012. Scan-rescan reproducibility of neurite microstructure estimates using NODDI. Med. Image Underst. Anal. 255–261.

Tariq, M., Schneider, T., Alexander, D.C., Wheeler-Kingshott, C.A., Zhang, H., 2013. Assessing scan-rescan reproducibility of the parameter estimates from NODDI. Proc. Intl. Soc. Mag. Reson. Med. 21 (21), 3187.

Timmers, I., Roebroeck, A., Bastiani, M., Jansma, B., Rubio-Gozalbo, E., Zhang, H., 2016. Assessing microstructural substrates of white matter abnormalities: a Comparative study using DTI and NODDI. PLoS ONE 11, 1–15. doi:10.1371/journal.pone.0167884.

Veraart, J., Fieremans, E., Novikov, D.S., 2016a. Diffusion MRI noise mapping using random matrix theory. Magn. Reson. Med. 76, 1582–1593. doi:10.1002/mrm.26059.

Veraart, J., Novikov, D.S., Christiaens, D., Ades-aron, B., Sijbers, J., Fieremans, E., 2016b. Denoising of diffusion MRI using random matrix theory. Neuroimage 142, 394–406. doi:10.1016/j.neuroimage.2016.08.016.

Wheeler-Kingshott, C.A.M., Cercignani, M., 2009. About "axial" and "radial" diffusivities. Magn. Reson. Med. 61, 1255–1260. doi:10.1002/mrm.21965.

Zhang, H., Schneider, T., Wheeler-Kingshott, C.A., Alexander, D.C., 2012. NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. Neuroimage 61, 1000–1016. doi:10.1016/j.neuroimage.2012.03.072.

Zhang, J., Gregory, S., Scahill, R.I., Durr, A., Thomas, D.L., Lehericy, S., Rees, G., Tabrizi, S.J., Zhang, H., 2018. In vivo characterization of white matter pathology in premanifest huntington's disease. Ann. Neurol. 84, 497–504. doi:10.1002/ana.25309.