

RESEARCH ARTICLE

WILEY

High-frequency longitudinal white matter diffusion- and myelin-based MRI database: Reliability and variability

Manon Edde^{1,2}  | Guillaume Theaud² | Matthieu Dumont² | Antoine Théberge^{1,3} | Alex Valcourt-Caron¹ | Guillaume Gilbert⁴ | Jean-Christophe Houde² | Loika Maltais² | François Rheault⁵  | Federico Spagnolo⁶ | Muhamed Barakovic⁶ | Stefano Magon⁶ | Maxime Descoteaux^{1,2}

¹Sherbrooke Connectivity Imaging Lab (SCIL), Université de Sherbrooke, Sherbrooke, QC, Canada

²Imeka Solutions, Inc., Sherbrooke, QC, Canada

³Videos & Images Theory and Analytics Laboratory (VITAL), Université de Sherbrooke, Sherbrooke, QC, Canada

⁴MR Clinical Science, Philips Healthcare Canada, Mississauga, Ontario, Canada

⁵Medical Imaging and Neuroinformatics (MINI) Lab, Université de Sherbrooke, Sherbrooke, QC, Canada

⁶Pharma Research and Early Development, Neuroscience and Rare Diseases Roche Innovation Center Basel, F. Hoffmann-La Roche Ltd., Basel, Switzerland

Correspondence

Manon Edde, Sherbrooke Connectivity Imaging Lab (SCIL), Université de Sherbrooke, Sherbrooke, QC, Canada.
Email: eddemanon@gmail.com

Funding information

Canadian Network for Research and Innovation in Machining Technology, Natural Sciences and Engineering Research Council of Canada, Grant/Award Number: RGPIN-2020-04818; Mitacs Accelerate program; Université de Sherbrooke Institutional Chair in Neuroinformatics

Abstract

Assessing the consistency of quantitative MRI measurements is critical for inclusion in longitudinal studies and clinical trials. Intraclass coefficient correlation and coefficient of variation were used to evaluate the different consistency aspects of diffusion- and myelin-based MRI measures. Multi-shell diffusion and inhomogeneous magnetization transfer data sets were collected from 20 healthy adults at a high-frequency of five MRI sessions. The consistency was evaluated across whole bundles and the track-profile along the bundles. The impact of the fiber populations on the consistency was also evaluated using the number of fiber orientations map. For whole and profile bundles, moderate to high reliability of diffusion and myelin measures were observed. We report higher reliability of measures for multiple fiber populations than single. The overall portrait of the most consistent measurements and bundles drawn from a wide range of MRI techniques presented here will be particularly useful for identifying reliable biomarkers capable of detecting, monitoring and predicting white matter changes in clinical applications and has the potential to inform patient-specific treatment strategies.

KEYWORDS

diffusion, healthy subject, MRI, myelin, reliability, variability, white matter

1 | INTRODUCTION

In recent decades, there has been a growing body of evidence that white matter (WM) plays a prominent role in various pathologies. In

this context, longitudinal studies of WM have become increasingly important. Currently, various magnetic resonance imaging (MRI) techniques are used to assess different properties of WM tissues such as axonal density, fiber organization, or myelin (Jones et al., 2013). The most common technique is diffusion-weighted imaging (DWI), which includes single-compartment models such as diffusion tensor imaging

Stefano Magon and Maxime Descoteaux should be considered co-senior authors.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Human Brain Mapping* published by Wiley Periodicals LLC.

(DTI, Basser et al., 1994) or more complex models using multi-compartment fitting such as high angular resolution diffusion imaging (HARDI, Dell'Acqua & Tournier, 2019; Jeurissen et al., 2013) and neurite orientation density and dispersion imaging (NODDI, Zhang et al., 2012) from multi-shell diffusion MRI. Other techniques like magnetization transfer imaging (MTI, Kim & Cercignani, 2014; Wolff & Balaban, 1989) are also increasingly used to examine changes in the myelin content of WM. Key measures of WM microstructure derived from these models are sensitive to changes in a healthy population (Alexander, 2017; Beck et al., 2021; Boukadi et al., 2019; Honnedeavasthana Arun et al., 2021; Koshiyama et al., 2020; Munsch et al., 2021; Uddin et al., 2019) or in pathological conditions (Beaudoin et al., 2021; Brown et al., 2017; Granziera et al., 2021; Laule & Moore, 2018; Lu et al., 2021; Rahmanzadeh et al., 2022; Schneider et al., 2017). In addition, studies have also shown the effect of a treatment or therapeutic intervention on WM measures in clinical trials (Arnold et al., 2022; Gurevich et al., 2018; Roy et al., 2021; Vavasour et al., 2019).

However, reliably evaluating, monitoring or predicting any changes in WM microstructure requires data with high consistency and enough statistical power to detect these changes (Poldrack et al., 2017; Zuo et al., 2019). MRI measurements can be influenced by random effects introducing measurement errors such as image noise or MRI signal variation (Wang et al., 2012). In addition, repeated measure analyses also introduce additional sources of variation (e.g., a change of technologist during data acquisition or subject positioning). Together, these errors affect data consistency which is an important factor in the sensitivity and specificity of the analysis (Tofts et al., 2018; Wang et al., 2012; Zuo et al., 2019). Therefore, it is essential to evaluate the different aspects of the consistency of the measurements including reliability, reproducibility and variability of measurements derived from MR images, especially so for the more novel and more complex quantitative MRI protocols. Here, reliability refers to the overall consistency of the measurements across subjects, i.e., it reflects both the degree of correlation and agreement between measures (Bruton et al., 2000; Koo & Li, 2016). Variability can be separated into within-subject variability and between-subject variability. Within-subject variability can be used to assess the ability to obtain similar values across sessions of the same subject, i.e., an index of measurement reproducibility. Finally, between-subject variability represents the sample heterogeneity, i.e., how much one subject differs from another.

To date, consistency of MRI-based WM measurements has been evaluated through numerous studies, especially for the DWI technique (Boekel et al., 2017; Grech-Sollars et al., 2015; Hakulinen et al., 2021; Magnotta et al., 2012; Teipel et al., 2011; Thieleking et al., 2021; Veenith et al., 2013). These studies reported moderate to high reliability in the WM using Intraclass correlation coefficient (ICC) or Pearson's correlation ranging from 0.5 to >0.8 as well as within- and between-subject coefficients of variation (CV) ranging from 1% to 8% and 1% to 15%, respectively. Among DTI-derived measures, Fractional anisotropy (FA) and Mean Diffusivity (MD) generally show the highest reliability across different WM regions (Acheson et al., 2017;

Hakulinen et al., 2021; Luque Laguna et al., 2020; Palacios et al., 2017; Shahim et al., 2017; Thieleking et al., 2021; Zhou et al., 2018). For NODDI-derived measures, studies reported similar (intracellular volume fraction, ICvf) or higher (orientation dispersion, OD) reliability compared to DTI measures, while isotropic volume fraction (ISOvf) showed the poorest reliability (ICC <0.6) (Andica et al., 2020; Chung et al., 2016; Granberg et al., 2017; Lucignani et al., 2021; Tariq, 2013). In contrast, to the best of our knowledge, the reliability of HARDI-derived measurements such as apparent fiber density (AFD, Raffelt et al., 2012) and the number of fiber orientations (NuFO; Dell'Acqua et al., 2013) has not been yet examined in healthy subjects or clinical population. Regarding MTI, MTR (Henkelman et al., 2001; Vavasour et al., 2011)—the most common measure—has been shown to have good reliability (ICC >0.7) (Filippi et al., 2000; Hickman et al., 2004; Schwartz et al., 2019; van der Weijden et al., 2021; Weiskopf et al., 2013). More recently, Inhomogeneous magnetization transfer (ihMT, Varma et al., 2015)—a novel development of MT—has been shown to be more specific to myelin content compared to MTR (Duhamel et al., 2019; Girard et al., 2015; Manning et al., 2017; Varma et al., 2015; Zhang et al., 2020) and sensitive to multiple sclerosis-related (MS) processes in transversal studies (Obberghen et al., 2018; Rasoanandrianina et al., 2020; Zhang et al., 2020). To date, two studies reported good reliability of ihMT measuring with ICC ranging from 0.6 to 0.95 in different regions of the WM (Mchinda et al., 2018; Zhang et al., 2019), whereas the only longitudinal study suggests that ihMT may not have enough statistical power to detect changes during brain development (Geeraert et al., 2019). Hence, that reinforces the need to examine the reliability and variability of this recent, but promising, technique.

Nevertheless, several important issues remain to be addressed: (1) although the recent studies include a reasonably large sample size $n \geq 20$ (Boekel et al., 2017; Hakulinen et al., 2021; Lehmann et al., 2021; Thieleking et al., 2021), most of them are based on limited data with sample sizes of $n \leq 10$ (Andica et al., 2020; Chung et al., 2016; Granberg et al., 2017; Koller et al., 2020; Tariq, 2013; Zhang et al., 2019); (2) most of the previous studies are focused on a short-period (scan-rescan within a week or with 2–4 weeks intervals) rather than longer time intervals. Indeed, longitudinal neuroimaging studies or clinical trials are typically separated by a few weeks (>3 weeks) to several months; (3) only a few studies include more than two or three-time points (Cai et al., 2021; Schwartz et al., 2019), thus not generating enough data per subject to assess relevant reliabilities; and (4) few studies have directly compared multiple WM microstructural measures from several MRI techniques in the same population (Koller et al., 2020; Schwartz et al., 2019).

On the other hand, none of these studies evaluated the impact of local WM complexity on consistency, particularly the number of fiber populations. Indeed, these studies consider each voxel as a single entity with a homogeneous fiber population. However, it has been described that voxels contain multiple fiber populations, i.e., between 66% and 90% of white matter voxels cannot be assumed to contain a single coherently oriented axon bundle (Jeurissen et al., 2013; Volz et al., 2018). In addition, Volz and colleagues have recently shown that

the value of FA depends on the number of fibers considered in the voxel, with a greater FA value for the single fiber population and smaller when the multiple fiber population is considered (Volz et al., 2018). Thus, measurements derived from different models—and by extension, their consistencies—may vary depending on the underlying WM organization, especially within bundle or track-profiles.

To address these problems, we designed a repeated-measure study to collect multiple microstructural (anatomical, multi-shell diffusion and inhomogeneous MT) MRI data sets in “high-frequency”—i.e., a high number of MRI acquisitions over a short period of time (6 months) for 20 healthy subjects. All subjects were scanned five times with an average interval of 4 weeks for a total of 100 MRIs. This high-frequency data set thus generates enough data per subject to allow a relevant assessment of the consistency of different brain MRI measurements. The reliability and variability were evaluated using the intraclass coefficient correlation (ICC) value and within- and between-subject coefficient of variation (CVw and CVb, respectively). Then, the consistency of MRI measurements was evaluated across the bundles as a tracts-of-interest analysis approach—i.e., averaging voxels within each WM bundle. To go further, the consistency of each WM measure was also evaluated as a profile along the bundle using *tractometry* (Cousineau et al., 2017; Yeatman et al., 2012). Finally, the same analyses were carried out by splitting each white matter bundle mask according to the number of fiber orientations using the NuFO map, a useful index of the number of fiber populations.

2 | METHODS

Method and results are documented and available at <https://high-frequency-mri-database-supplementary.rtfid.io>.

2.1 | Participants

Twenty healthy adults (mean age 36 years, age range 29–46, undefined 3 [SD = 4.7], 4 men and 16 women) were recruited from the environment of the University of Sherbrooke and the Centre Hospitalier Universitaire de Sherbrooke (CHUS). The study was designed with this proportion of male and female subjects to match a future MS group. Participants were screened for eligibility to undergo MRI, no history of brain disease or injury, left-handedness and received financial compensation for their participation. The study was approved by the ethics committee of the CHUS (Comité d'éthique de la recherche du CIUSSS de l'Estrie) in Sherbrooke, Canada and all participants gave prior informed written consent.

2.2 | MRI data acquisition

Whole-brain MRI data were acquired using a clinical 3 T MRI scanner (Ingenia, Philips Healthcare) using a 32-channel head coil. Each MRI session was repeated five times over 6 months and a 4-week interval

(+/- 1 week). For each participant, images were acquired at approximately the same time of day to avoid potential diurnal effects (i.e., a morning participant had all sessions in the morning, with a tolerated 2–3-h variation). All MRI data acquisitions were aligned on the anterior commissure-posterior commissure plan (AC-PC) and included (a) anatomical 3D T1-weighted, (b) multi-shell diffusion-weighted images (DWI), (c) inhomogeneous magnetization transfer (ihMT) for a total time of acquisition of 20 min (Table S1).

(a) 3D T1-weighted MPRAGE image was acquired axially at 1.0 mm isotropic resolution, repetition time (TR)/echo time (TE)/inversion time (TI) = 7.9/3.5/950 ms, field-of-view (FOV) = 224 × 224 mm² yielding 150 slices, flip angle = 8° for an acquisition time of 4 min 20 s.

(b) Multi-shell DWI images were acquired with a single-shot EPI spin-echo sequence at 2.0 mm isotropic resolution, TR/TE = 4800/92 ms, SENSE factor = 1.9, Multiband-SENSE factor = 2, flip angle of 90°, FOV = 224 × 224 mm², 66 slices for an acquisition time of 9 min 19 s. The data included 100 unique directions uniformly distributed over each shell and across the three shells at $b = 300 \text{ mm}^2/\text{s}$ ($n = 8$ directions), $b = 1000 \text{ mm}^2/\text{s}$ ($n = 32$ directions), $b = 2000 \text{ mm}^2/\text{s}$ ($n = 60$ directions), with non-diffusion-weighted images $b = 0 \text{ mm}^2/\text{s}$ ($n = 7$), for a total of 107 total diffusion volumes (Caruyer et al., 2013). To correct EPI distortions, a reversed phase-encoded $b = 0$ image was acquired right after the DWI acquisition, with the same geometry (Andersson et al., 2003).

(c) Inhomogeneous MT images were acquired using a 3D segmented-EPI gradient-echo sequence with different MT preparation pulses with first TE/TR = 3.6/112 ms, 2 × 2 mm resolution, flip angle of 15°, FOV = 224 × 224 mm, 65 slices of 2 mm of thickness and three echoes with echo spacing 6.0 ms for an acquisition time of 6 min 04 s (see Data S1 and Figure S1). Inhomogeneous MT uses a magnetization preparation (10 Hann pulses of 0.9 ms duration with 1.5 ms interval at a frequency offset of +/- 7000 Hz) (Varma et al., 2015). Two additional reference images were acquired for each echo without MT preparation, with the same parameters as the MT sequence and a second with a higher flip angle (30°) and a shorter TR (20 ms) for quantification purposes.

2.3 | MRI processing

2.3.1 | Tractoflow: DWI and T1 processing

After visual quality assessment, for each participant, *Tractoflow* (Theaud et al., 2020) a pipeline developed by SCIL (<https://github.com/scilus/tractoflow>), was used to process DWI and T1w images. This pipeline generates both diffusion measures and tractography of WM, from raw DWI, T1w, bvec/bval files and the reversed phase-encoded $b = 0$ and has been proved to be highly reproducible in time and immediate test-retest (Theaud et al., 2020). Briefly, after denoising and correcting the raw DWI images for motion, eddy-currents, geometric distortions and field inhomogeneity, the fiber orientation distribution function (fODF) was generated using constrained spherical deconvolution (Descoteaux et al., 2007; Tournier et al., 2007) with

a fixed fiber response of $(15, 4, 4) \times 10^{-4} \text{ s/mm}^2$ for all subjects (Pierpaoli & Basser, 1996), as recommended in (Dell'Acqua et al., 2013), a maximal spherical harmonics order of 8 and all b-value DWI data. Four DTI measures including fractional anisotropy (FA), mean diffusivity (MD), radial diffusivity (RD) and axial diffusivity (AD) were computed and HARDI-derived measures including total apparent fiber density (AFD total) and the number of fiber orientations (NuFO), were extracted from the fODF (see Table 1 for a list of available measures). In parallel, the T1w anatomical image was also denoised, corrected, and registered to the $b = 0$ and the FA images before tissue segmentation to generate the tracking maps including inclusion, exclusion maps and a WM seeding mask (G. Girard et al., 2014). The whole-brain ensemble tractogram was generated from a fODF map and tracking masks using both the anatomically constrained particle filter tracking algorithm (PFT, Girard et al., 2014) and local tracking with 5 and 2 seeds per voxel respectively. Except for the number of seeds per voxel, all parameters used the default *Tractoflow* settings (see Theaud et al., 2020 for a complete pipeline description).

2.3.2 | Neurite orientation dispersion density imaging

NODDI is a multi-shell compartment modeling technique that identifies three types of microstructural environments: intracellular, extracellular, and CSF compartments (Zhang et al., 2012). NODDI measures were extracted using *NODDI flow* from SCIL (https://github.com/scilus/noddi_flow), which used Accelerated Microstructure Imaging via Convex Optimization (AMICO, Daducci et al., 2015) from multi-shell DWI images. Complementing, the extracellular volume fraction (ECV), a measure of the volume fraction within a voxel that is not neuronal and assumed to be due to glial cells infiltration, was computed as follows $ECV = 1 - ICV$. Finally, four microstructural maps

were generated: ECv, intracellular volume fraction (ICv), isotropic volume fraction (ISOv), and orientation dispersion (OD).

2.3.3 | Magnetization transfer imaging (MTI)

Inhomogeneous MT images were processed using a custom in-house pipeline including tools from the FSL, Advanced Normalization Tools software (ANTs, Avants et al., 2011) and SCIL pipeline scripts (https://github.com/scilus/ihmt_flow). For each echo, raw ihMT images were firstly co-registered using ANTs linear registration (Avants et al., 2008). Next, the reference image was used to perform tissue segmentation with the *AtroposN4* command from ANTs. Tissue maps from the above T1w processing were concatenated and used as brain mask during ihMT processing. Two ihMT images were generated from all frequencies and reference images as previously described in (Varma et al., 2015): ihMT ratio (ihMTR) and ihMT ΔR_1 saturation (ihMTdR1sat)—developed to enhance ihMTR contrast by decoupling the ihMTR signal from the T1 longitudinal relaxation rate. Like ihMT saturation (ihMTsat), the ihMTdR1sat measure is computed to mitigate the confounding factors from T1 relaxation and B1+ inhomogeneity effects (Alsop et al., 2022). While not mathematically equivalent to ihMTsat, ihMTdR1sat measurement shows a very high correlation with ihMTsat (see Data S1 and Figure S2). In addition, ihMTdR1sat presents a higher SNR compared to ihMTR, given the use of a high flip angle, short-TR reference image. Additionally, from the positive frequency data and reference images, we also computed two “standard” MT images: MT ratio (MTR) and MT saturation (MTsat)—generated as described in Helms et al., 2008. However, it should be noted that these MT measures are computed from ihMT acquisitions whose saturation frequency of 7000 Hz is outside the range generally used for MTR and MTsat (1000–2500 Hz). Finally, the four resulting myelin-sensitive maps were registered to the $b = 0$ and the FA images using nonlinear SyN ANTs.

TABLE 1 List of pipelines, models and measures evaluated with the corresponding abbreviations.

Public pipeline	Models	Measures	Abbreviation
Tractoflow	Diffusion tensor imaging (DTI)	Fractional anisotropy	FA
		Axial diffusivity	AD
		Radial diffusivity	RD
		Mean diffusivity	MD
	High angular resolution diffusion imaging (HARDI)	Apparent fiber density total	AFD total
		Number of fiber orientation	NuFO
NODDI Flow	Neurite orientation dispersion and density imaging (NODDI)	Intracellular volume fraction	ICv
		Extracellular volume fraction	ECv
		Isotropic volume fraction	ISOv
		Orientation distribution	OD
ihMT flow	Magnetization transfer imaging (MTI)	inhomogeneous magnetization transfer (MT) ratio	ihMTR
		inhomogeneous MT delta R1 saturation	ihMTdR1sat
		Magnetization transfer ratio	MTR
		Magnetization transfer saturation	MTsat

Table 1 provides the complete list of measures included in the analyses.

2.4 | White matter virtual dissection

The major fascicles were automatically extracted using RecoBundlesX (Rheault, 2020) (<https://zenodo.org/record/4104300#.YNp1XVKiiM>) a multi-atlas and multi-parameter version of RecoBundles (Garyfallidis et al., 2018). For the sake of clarity and to avoid overloading, we focus the paper on a subset of bundles to show consistency in one association, commissural and projection pathway: Arcuate Fasciculus (AF), section 3 of the Corpus Callosum (CC), and Cortico-Spinal Tract (CST) are selected as bundles of interest (Catani & de Schotten, 2012). In addition, the inferior fronto-occipital fasciculus (IFOF) is also included to show consistency in a “hard-to-track” long bundle (Figure 1). Bundle colors will be matched throughout the results. All analyses are nonetheless conducted on all bundles and measures, and the respective results are available at <https://high-frequency-mri-database-supplementary.rtfid.io/>.

2.5 | Common space and average measures

To perform consistency voxel-based analysis, all images were registered in a common space. Symmetric diffeomorphic normalization

(SyN) of ANTs is used to build a template in diffusion space based on our population. For the registration, we used iterative rigid, affine, and SyN (neighborhood cross-correlation) transformations with optimal similarity measures for the linear (mutual information) (Avants et al., 2011) (https://high-frequency-mri-database-supplementary.readthedocs.io/en/latest/pipeline/common_space.html). The processed b0 images resampled at 1 mm isotropic from *Tractoflow* were used as an input with four iterations with decreasing degrees of downsampling and smoothing. All subject-specific measures maps and bundles were aligned in the common diffusion space using the resulting nonlinear registration. Finally, averaged maps were computed for each measure using all subjects and sessions and shown in Figure 2 and available at https://high-frequency-mri-database-supplementary.readthedocs.io/en/latest/results/average_maps.html.

2.6 | Consistency for whole-bundle average and profiles along bundles

For each bundle, the consistency of the different measurements was evaluated from (1) the bundle-averaged that is, one measure for the whole bundle and (2) along the bundle as a profile, also called tract-profile or connectometry (Cousineau et al., 2017; Yeatman et al., 2012; Yeh et al., 2016). For the bundle-averaged, the density map was used to generate a binary mask of each whole bundle in the common space. Then, to minimize the effect of partial volume, each

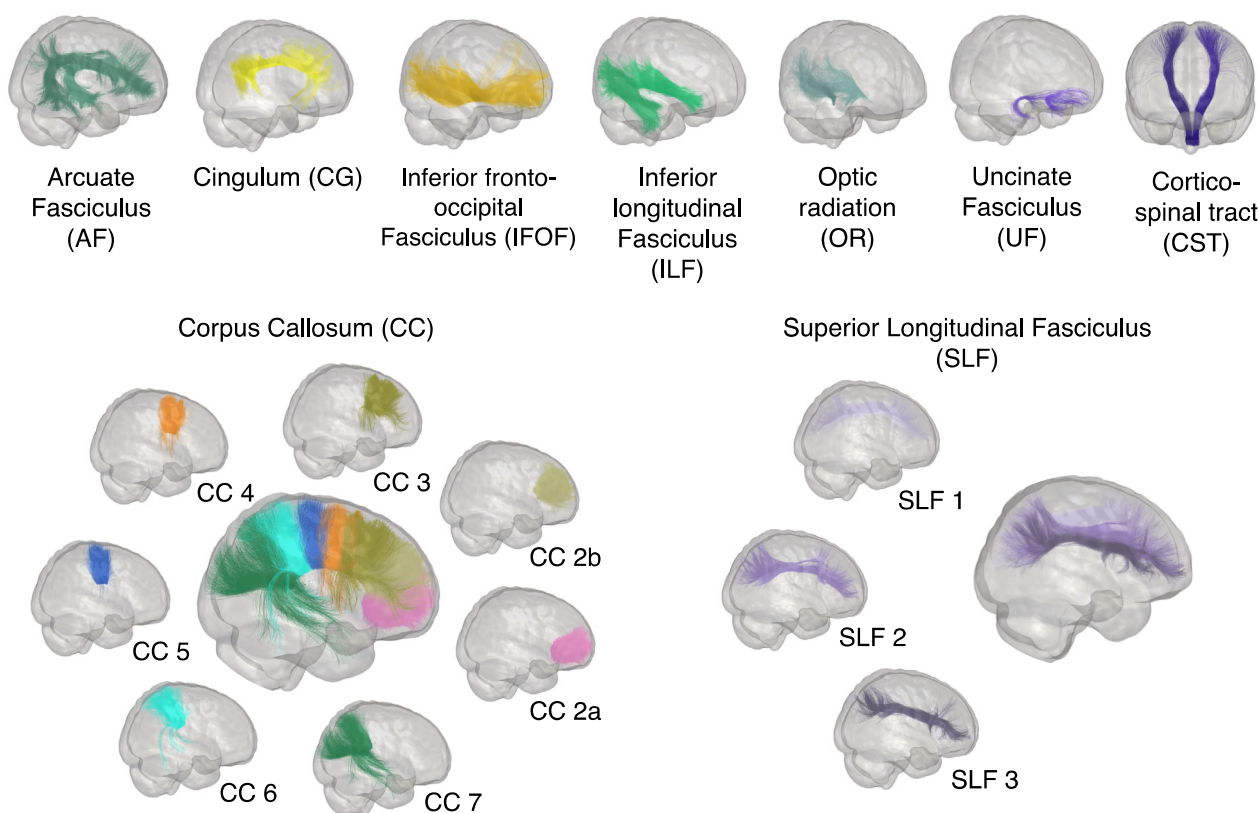


FIGURE 1 Representation of the major bundle models used by RecoBundlesX as shape priors to extract the bundles from the whole tractogram. Bundles of both hemispheres are shown.

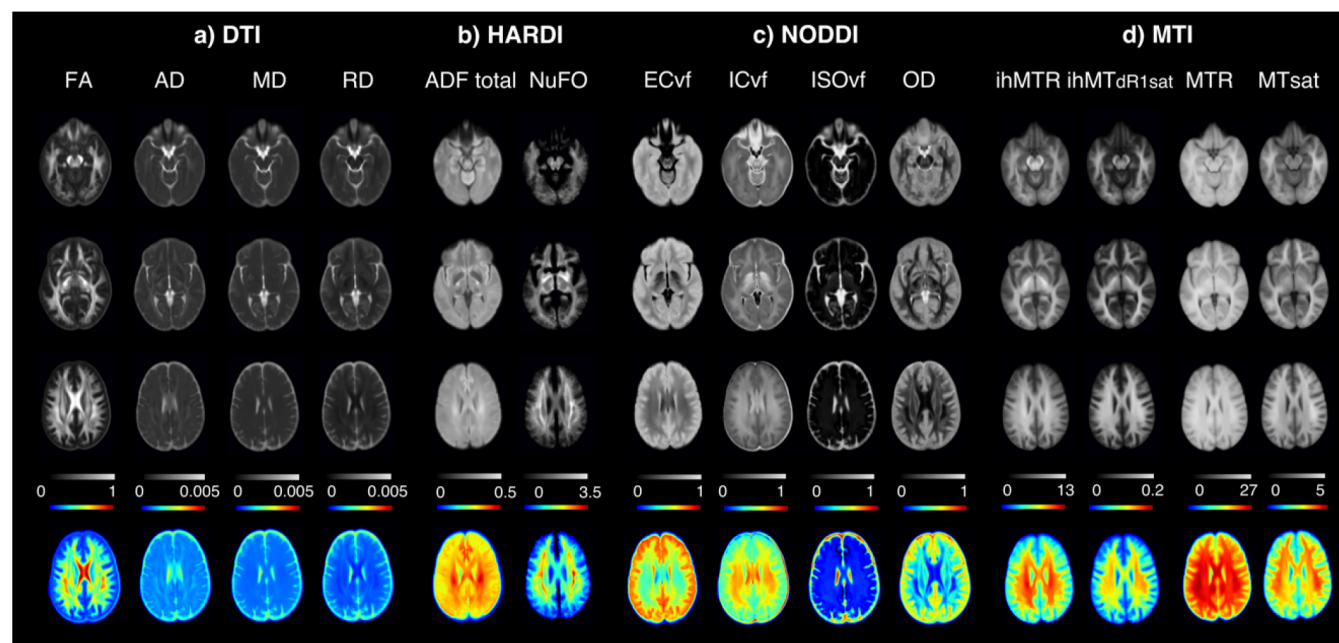


FIGURE 2 Microstructural maps, in (a) DTI measures including FA (fractional anisotropy), AD (axial diffusivity, units = mm^2/s), MD (mean diffusivity, units = mm^2/s) and RD (radial diffusivity, units = mm^2/s); (b) HARDI measures with ADF total (total apparent fiber density) and NuFO (number of fiber orientations); (c) NODDI measures including ECvf (extracellular volume fraction) and ICvf (intracellular volume fraction) compartment, ISOvf (isotropic volume fraction) and OD (orientation distribution) and (d) MTI measures including ihMTR (inhomogeneous MT ratio), ihMTdR1sat (inhomogeneous MT saturation), MTR (MT ratio), MTsat (MT saturation); averaged across participant. All contrasts are registered to diffusion space.

whole bundle mask was eroded by one voxel to generate a conservative bundle mask that we called the “safe mask.” For the consistency profiles, *Tractometry_flow*, a public pipeline developed by SCIL (https://github.com/scilus/tractometry_flow, Cousineau et al., 2017; Yeatman et al., 2012) was applied to each subject-specific bundle to obtain binary mask corresponding to each bundle section (each section corresponding to a specific label). Each bundle was resampled into 10 equidistant sections and intersected with the safe mask.

Next, left and right masks were merged for each average and section bundle mask (Figure S3). Finally, DTI, HARDI, NODDI and MTI measures were extracted for each bundle mask over session.

2.7 | Impact of fiber populations on consistency

Many studies have reported that voxels containing multiple fiber populations (Jeurissen et al., 2013; Volz et al., 2018) affect the microstructural measures (Volz et al., 2018). Here, we evaluate the effect of fiber populations projecting in multiple orientations on the consistency of the WM measures. To this end, we used the NuFO maps which are estimated from the number of local maxima of the fODF profile in each voxel (Dell'Acqua et al., 2013). The intensity of each voxel corresponds to the number of fiber populations, ranging from 1 for the single fiber population to 2 and more for the multiple fiber populations (Jeurissen et al., 2013). We apply two thresholds of 1 and ≥ 2 on the NuFO map to compartmentalize the “average” bundle (i.e., whole bundle) into “single” and “multi” fiber populations

compartments, respectively. For this, each voxel of the whole and section masks for each bundle was sorted according to these two thresholds (https://high-frequency-mri-database-supplementary.readthedocs.io/en/latest/pipeline/fiber_population.html).

An overview of our analysis pipeline is illustrated in Figure S4 and <https://high-frequency-mri-database-supplementary.readthedocs.io/en/latest/data/overview.html#pipeline-summary>.

2.8 | Quality control

A visual quality assessment procedure was carried out for major steps including raw input data, preprocessing, registration steps, bundles segmentation and tract profiles using *dMRIqc_flow* (https://github.com/scilus/dmriqc_flow, see Table S2 for the criteria). Because resampling the bundles involves smaller mask volumes and therefore, introduces a potential confounding factor, it is important to ensure that each section of bundles contains enough voxels to assess consistency measurements. For this purpose, we extracted the volume of each section corresponding to the bundle profile analyses and the fiber population compartments that generate a novel subdivision of the masks of each section in single- and multi-compartment. Two minimum thresholds of 1000 and 400 voxels respectively were used to perform analyses, therefore, sections of bundles that had fewer voxels than these thresholds were excluded.

Since ISOvf accounts for the isotropic volume fraction, this parameter has generally very low values in the WM (Tariq, 2018), and

many studies emphasized the poor reliability of this NODDI parameter (Andica et al., 2020; Chung et al., 2016; Lehmann et al., 2021; Lucignani et al., 2021). To improve the consistency of this parameter, we proposed an evaluation of different thresholds to remove values close to zero. A range of thresholds between 0 and 0.1 with a step size of 0.01 was used.

2.9 | Evaluation of bundles

The evaluation of the reproducibility of bundles is carried out first to minimize the impact of the variability of reconstruction by tractography on the measurements. The reproducibility of the bundles was achieved using the same method as Rheault et al., 2020, 2022. We computed the Dice similarity score, correlation between the density maps and adjacency streamlines from all pairwise combinations to provide the agreement between segmentations of the same bundle across sessions.

2.10 | Correlation analysis

Pearson's correlation coefficient (r) was used to evaluate the covariance of the averaged diffusion measures for all bundles. For this, we used each bundle's averaged measure, extracted from each voxel and averaged along bundles. Pearson correlations were computed for each session and then averaged across sessions to generate an average correlation map for all sessions. The average correlation interactive map and correlation interactive maps corresponding to the sessions are available at <https://high-frequency-mri-database-supplementary.readthedocs.io/en/latest/results/correlation.html>.

2.11 | Consistency evaluation

2.11.1 | Consistency measures

The reliability of computed measures was investigated using the Image Intra-Class Correlation coefficient (I2C2, Shou et al., 2013), a generalization of the Intra-Class Correlation coefficient (ICC, Koo & Li, 2016) to n -dimensional images (one-way random effect, absolute agreement). The ICC estimates the correlation between measures corresponding to different sessions in terms of their consistency across subjects. The variability induced by within-subject and between-subject effects on the measures was quantified using two coefficients of variation per measure. The coefficient of variation within-subject (CVw) was used to evaluate the dispersion of observations when repeatedly measuring a single individual (i.e., reproducibility), thus representing the amount of random error or noise contributing to the measure. For the CVw, the CV is first estimated per subject over their respective imaging sessions and then averaged. The coefficient of variation between-subject (CVb) was used to evaluate the sample heterogeneity. The CVb is obtained by first averaging each subject session-wise, to then estimate the CV over those averages.

Confidence intervals and p -values were obtained for the I2C2 using non-parametric bootstrap (Briggs et al., 1997; Efron & Tibshirani, 1994) and the accelerated bias-corrected percentile method (Diciccio & Romano, 1988; Efron & Tibshirani, 1994) using SciPy tools (<https://scipy.org/>). To avoid overloading, confidence intervals and p -values are only reported for selected bundles and the whole bundle analysis (Table S3).

2.11.2 | Voxel-based consistency analysis

Computation of consistency measures was restricted to the safe white matter masks. The consistency analyses of each measure were carried out at the voxel-level within the bundles' mask. The individual masks corresponding to each subject and session in the common space were provided as input. The overlap between masks across sessions and subjects was then performed as described in the consistency measures. The incomplete overlap of mask between subjects and sessions was compensated by densifying each measure in the affected regions voxel-wise, using the average value estimated from the available subjects or sessions. The averaged masks used for the computation of statistical measurements are then obtained subject-wise or session-wise by mathematical union.

3 | RESULTS

3.1 | Quality control

None of the data were excluded based on input quality controls. However, two subjects were excluded from IFOF and UF analyses due to failed reconstruction of these bundles—caused by poor WM-GM-CSF segmentation in the internal capsule. In this case, all sessions were excluded for these two bundles but included for the correctly reconstructed bundles. CC Part 1 reconstruction failed for most subjects and was therefore excluded from all analyses. The set of bundles finally included in this paper is shown in Figure 1. Regarding, the number of voxels in each section of bundles, section 10 of the cingulum bundle had an average volume under the threshold. The consistency profile of this bundle was therefore generated for sections 1 to 9 (Figure S5). No other bundle sections were excluded based on volume. Finally, based on the graph and consistency results for the different thresholds of the ISOvf map, before analysis, an additional thresholding step was applied to exclude values under 0.045, for each subject (Figure S6).

3.2 | Diffusion- and myelin-based maps in diffusion template space

Three representative axial slices of the resulting averaged maps in common diffusion space are shown in Figure 2, the bottom row represents the third line with a colormap. As expected, the maps highlight the regional variation of measurements corresponding to each map, both in grey (top rows) and with a range of colors (bottom row). The maps are

smooth but show sharp contrasts between different tissue types such as CSF, GM and WM. Moreover, higher values are observed in regions of highly structured white matter as can be seen with FA, AFD total, NuFO, ICvf or OD in corona radiata regions, internal capsule, forceps, or thalamic radiation, and inversely for ISOvf or MD. As for the other maps, higher ihMTR or ihMTdR1sat values are observed in these same regions. One can appreciate qualitatively the similarities between all the different diffusion and myelin-based measures. The resulting averaged maps in common space are also available at https://high-frequency-mri-database-supplementary.readthedocs.io/en/latest/results/average_maps.html.

3.3 | Consistency of bundle tractography

Dice scores presented in Figure 3 reveal that most of the bundles are highly reproducible (mean Dice >0.7). Dice scores of voxels are high and close for all bundles, ranging from 0.56 (for SLF 1) to 0.84 (for CC 2a), which indicates that the overall spatial agreement is good. Across all bundles, SLF 1 exhibits the lowest Dice scores on average (density correlation and adjacency results are available at https://high-frequency-mri-database-supplementary.readthedocs.io/en/latest/results/bundles_reproducibility.html). SLF 1 was not excluded from the analysis, but its results should be taken with caution (Figure 3).

3.4 | Correlation analysis measures

Pearson's correlations of all measurements across WM bundles are shown in Figure 4. This figure highlights three main aspects of the

data: (1) within-model measures form a highly correlated pocket, (2) most diffusion measures are correlated, and (3) some between-model measures are correlated, while most correlations of the between-model measures are weak. More precisely, within the DTI model, the measures of MD and RD show the strongest association with each other (mean across bundles, $r = .97$) and a lower association with FA ($r = -.72$ and $r = .86$, respectively) and AD ($r = .82$ and $r = .64$, respectively). MD and RD also show strong associations with ECvf and ICvf ($r > .8$). On the other hand, AD (except for RD and MD), OD and the two measures of HARDI seem uncorrelated, either between or within a model ($r < .6$, $r < .5$, $r < .6$, respectively). MTI measurements show a strong association between ihMTdR1sat and MTsat ($r = .88$), ihMTR ($r = .81$); and much weaker correlation between ihMTdR1sat and MTR ($r = .45$). MTsat appears to be the only MTI measure that correlates, albeit moderately, with diffusion measures (r ranging from .3 to .6). This weak correlation between MTI measures and diffusion measures shows that these are sensitive to different microstructural features.

Based on these results, subsequent analyses will present the following measurements: RD, AFD total, NuFO, ISOvf, MTR and ihMTdR1sat were extracted from the 4 selected bundles: Arcuate Fasciculus (AF), Section 3 of the Corpus Callosum (CC), and Cortico-Spinal Tract (CST). These measures were selected either because they show strong correlations with other measures and may share overlapping information which can cause redundancies, or because, in the opposite case, they show a pattern of decorrelation with other measures and may therefore provide different information (Cercignani & Bouyagoub, 2018; Chamberland et al., 2019). The distribution of selected measures for each subject and session is shown in Figure 5 and other measures are available at https://high-frequency-mri-database-supplementary.readthedocs.io/en/latest/results/average_maps.html.

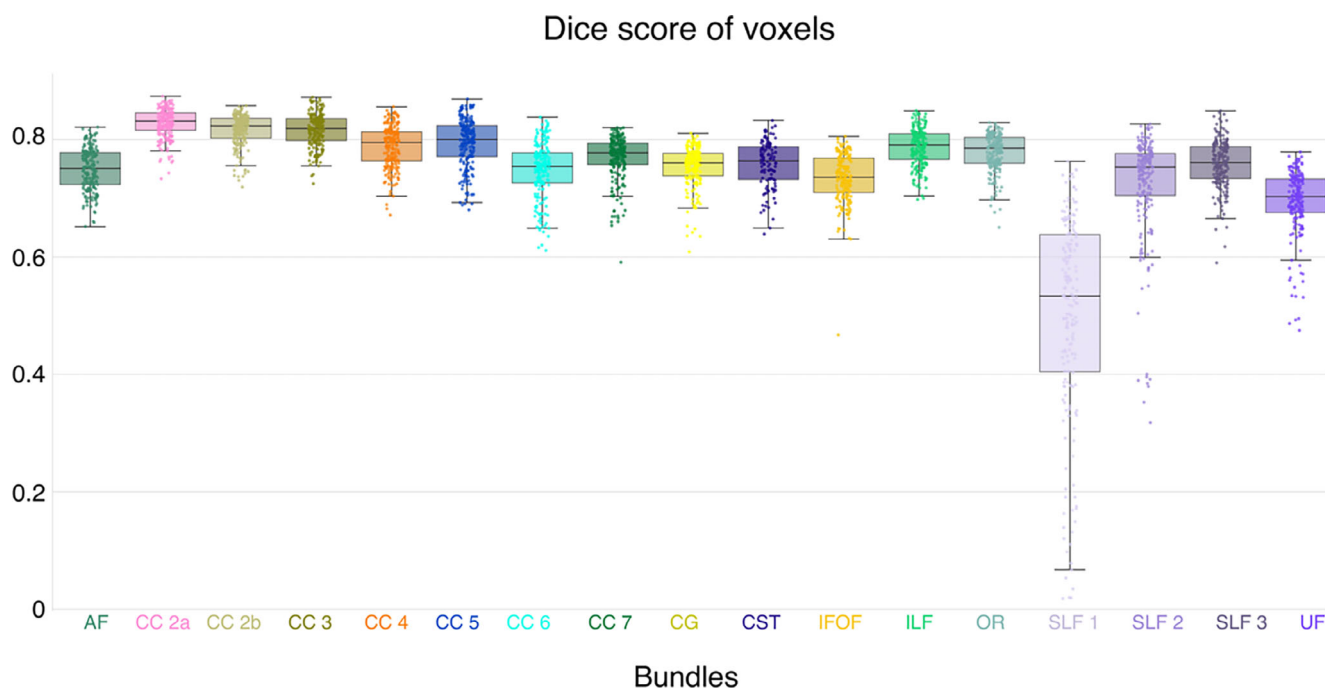


FIGURE 3 Bundles dice similarity coefficient scores for all subjects and sessions. Each dot represents one subject and session and colors correspond to each bundle.

Mean correlation of MRI measures across bundles and sessions

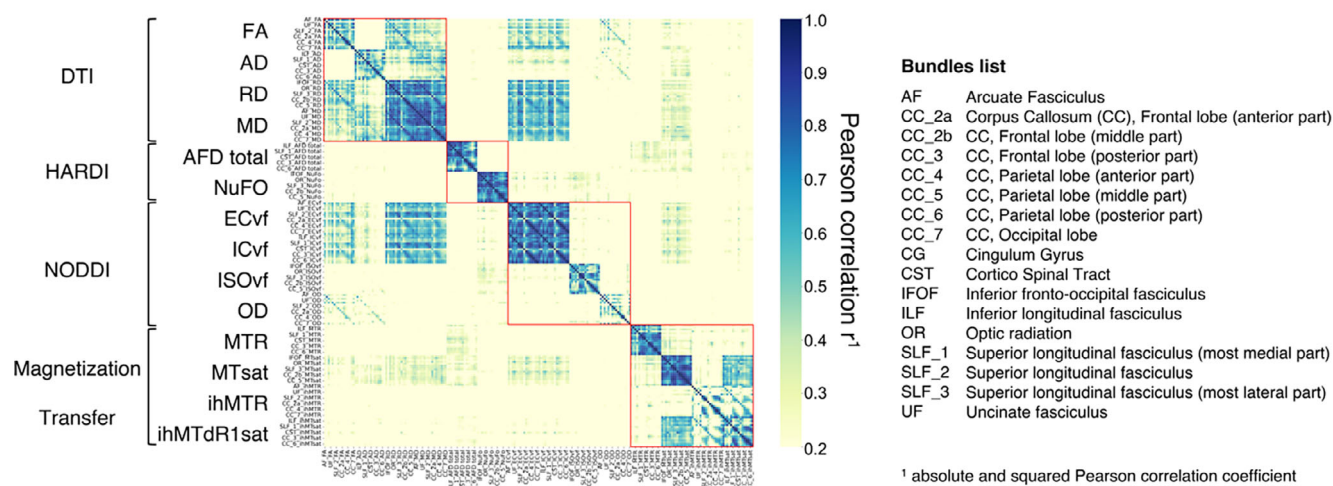


FIGURE 4 Pearson's correlation coefficients among all MRI measures and bundles. The red squares separate each DWI model and MTI. Each measurement is ordered according to the bundle list provided in the right panel of the figure. See <https://high-frequency-mri-database-supplementary.readthedocs.io/en/latest/results/correlation.html> for all correlation interactive maps.

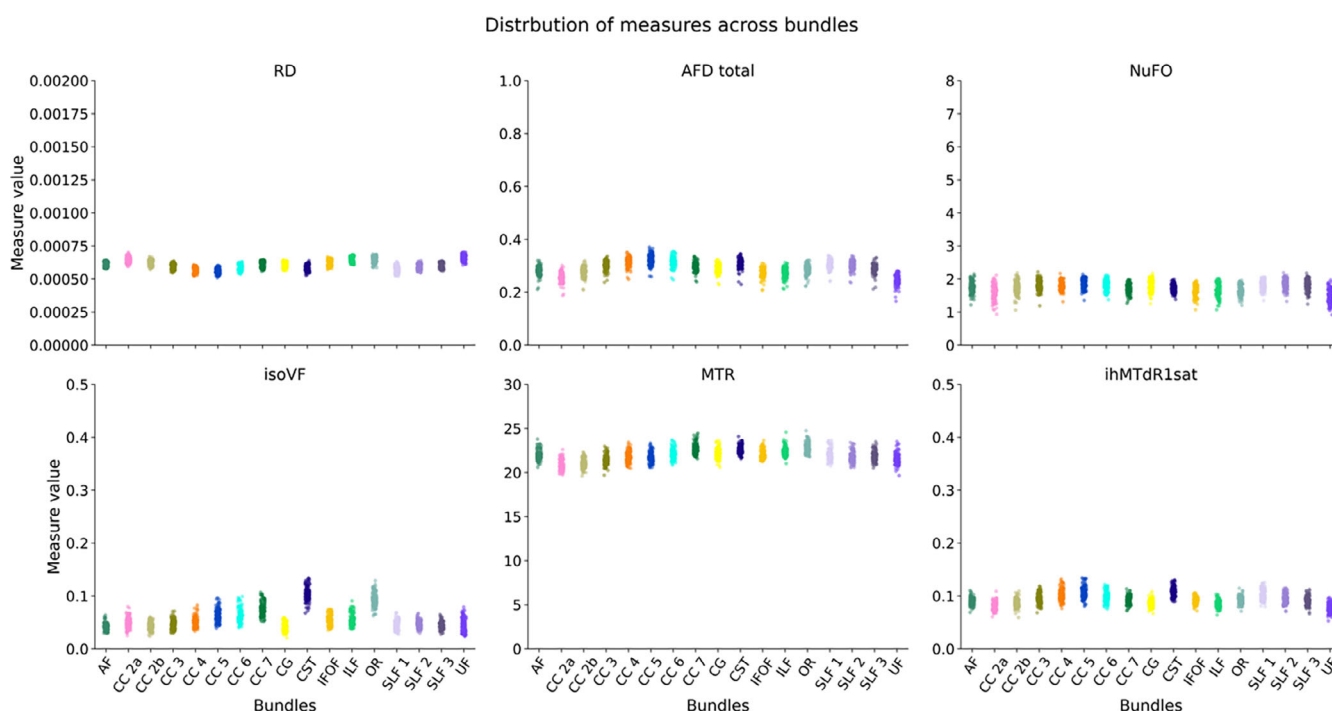


FIGURE 5 Individual measures for each bundle. Each dot represents one subject and session, and each plot represents one bundle. The colors correspond to the bundles.

[database-supplementary.readthedocs.io/en/latest/results/measure.html#whole-bundle-measures](https://high-frequency-mri-database-supplementary.readthedocs.io/en/latest/results/measure.html#whole-bundle-measures).

3.5 | Consistency of bundle-averaged measures

All consistency and MRI measurements are available at <https://high-frequency-mri-database-supplementary.readthedocs.io/en/latest/res>

[ults/consistency.html](https://high-frequency-mri-database-supplementary.readthedocs.io/en/latest/results/measure.html#whole-bundle-measures) and, confidence intervals and p-values for selected bundles are shown in Supplementary Table 3. Across DWI measures, most bundles exhibit a high degree of reliability with an ICC ranging from 0.55 for NuFo to 0.93 for FA and overall low variability ranging from 1.2% for FA (CVw) to 16% for ISOvf (CVb, Figure 6 and Table 2 for current study results and for a comparison with existing literature). As expected, DTI measures showed consistently the highest reliability (95% of ICC were higher than 0.80) and lowest

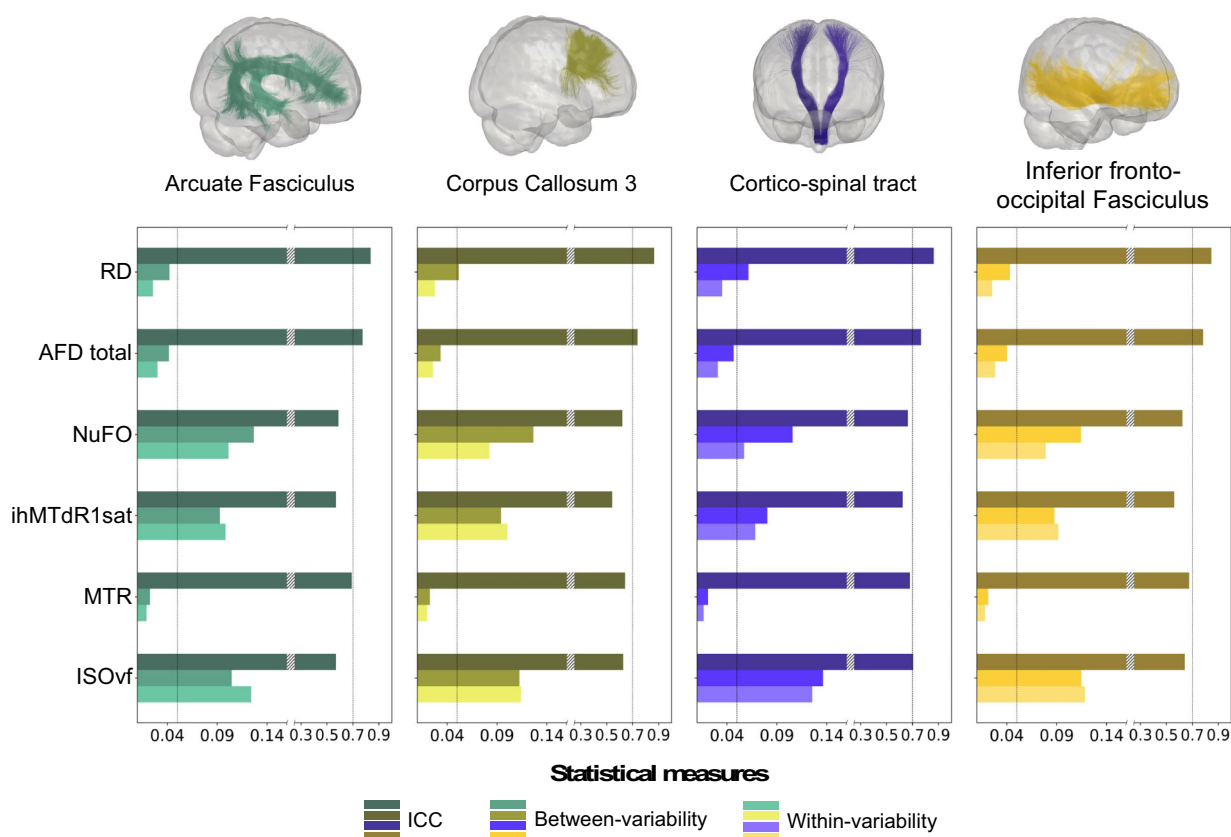


FIGURE 6 Consistency of bundle-averaged measurements of the selected bundles and measures. Each bar represents a measure of consistency with a decreasing saturation: the value of ICC in dark colors (the higher the better), Between-variability in medium colors (the lower the better) and within-variability in light colors (the lower the better). A break in the x-axis has been introduced to facilitate reading and the ICC and variability scales have been adapted.

variability (90% of CVw and CVb were lower than 5%) with higher CVb compared to CVw and small variation across bundles (Figure 6, Table 2). For HARDI measures, AFD total showed high reliability (ICC > 0.7 [0.73–0.8]) and low variability (CVw < 4.1% [1.8%–4%] and CVb < 6% [2.3%–5.7%]) for all bundles. NuFO showed lower reliability (ICC ~ 0.62 [0.55–0.7]) and higher variability (CVw ~ 8.1% [5.7%–11%] and CVb ~ 11.5% [9%–14%]) compared to AFD total (Figure 6, Table 2). NODDI measures showed lower reliability and higher variability compared to DTI measures. However, in general, OD showed consistently high reliability (ICC > 0.7) and low variability (CVw < 9% and CVb < 15%) followed by ICvf and ECvf with good reliability (ICC > 0.7) and variability (CVw and CVb < 10%) across all bundles. ISOvf measure represents the lowest reproducible measure of all NODDI maps with moderate reliability (ICC ~ 0.64) and greater variability (CVw and CVb < 16%) (Figure 6, Table 2).

Compared to DWI measures, a more variable pattern of results emerged for MTI measures with ICC ranging from 0.36 to 0.84 and CVs ranging from 1.4% to 13.8% across bundles. MTsat measures showed consistently the highest reliability (ICC > 0.7 except for UF and CC2a) followed by MTR (ICC ~ 0.67 across bundles). Both MT measures showed the lowest variability with CVs < 5% for all bundles. Noted that MTsat showed higher variability (CVw ~ 2.8% and

CVb ~ 4%) compared to MTR (CVw ~ 1.9% and CVb ~ 2.1%) (Figure 6, Table 2). In contrast, ihMT measures showed lower reliability with a mean ICC of 0.59 (0.44–0.67) and higher variability with a mean CVb of 7.9% (4.6%–11.8%) and CVw of 8.9% (6%–13.8%) compared with MT measures. As for the MT measures, ihMTdR1sat exhibited higher ICC but higher variability compared to ihMTR (Figure 6, Table 2).

3.6 | Consistency profiles along bundles

The distribution of measures along the bundle is provided here <https://high-frequency-mri-database-supplementary.readthedocs.io/en/latest/results/measure.html#profile-bundle-measures>. Globally, whatever the model and the measurement, the consistency measures quantified along the bundles display good stability or a low variability of their profile depending on the bundles (Figure 7). The different parts of the CC consistently show more variable profiles depending on the bundle sections, while the SLF bundles show the most stable profiles. On the other hand, unlike the whole bundle consistency measures, the values of ICCs profiles are lower, and the values of CVs profiles are higher (Figure 7, see also <https://high-frequency-mri-database-supplementary>.

TABLE 2 Measurements of reliability and variability across bundles (the current study is in bold) and a non-exhaustive list of values from the literature.

MRI	Study	MRI measure	ICC	CV %	Within-CV %	Between-CV %
DWI	Current study (n = 20) ^a	FA	0.87 (0.79–0.93)	-	5.1 (3.1–6.9)	9.9 (7.0–10.6)
		MD	0.80 (0.72–0.86)	-	1.8 (1.2–2.3)	2.7 (1.8–3.5)
		AD	0.85 (0.75–0.90)	-	2.4 (1.6–2.9)	4.1 (3.2–4.8)
		RD	0.84 (0.77–0.91)	-	2.9 (1.9–3.5)	4.8 (3.7–6.2)
		AFD total	0.77 (0.73–0.80)	-	2.8 (1.8–4.0)	3.8 (2.3–5.7)
		NuFO	0.62 (0.55–0.70)	-	8.1 (5.7–10.9)	11.5 (9.0–14)
		ICvf	0.75 (0.67–0.81)	-	2.7 (1.9–3.3)	3.7 (2.6–4.6)
		ECvf	0.75 (0.66–0.80)	-	4.3 (2.9–8.9)	5.1 (3.5–9.7)
		ISOvf	0.64 (0.57–0.70)	-	11.8 (9.0–13.8)	1.2 (8.7–15.9)
		OD	0.87 (0.81–0.92)	-	7.3 (5.7–8.3)	13.2 (10.6–14.7)
	Acheson et al., 2017 (n = 12)	FA	(0.77–0.99)	(0.91–3.86)	-	-
	Grech-Sollars et al., 2015 (n = 9)	FA	(0.3–0.8)	-	(2.5–9.2)	(3.6–9.7)
		MD	0.41	-	(1.3–5.5)	(1.4–2.6)
	Veenith et al., 2013 (n = 26)	FA	0.78 (0.56–0.98)	-	-	7.9 (3.3–31.7)
		ADC	0.79 (0.17–0.99)	-	-	7.3 (2.4–33.7)
		AD	0.74 (0.47–0.98)	-	-	4.5 (1.5–15)
		RD	0.82 (0.46–0.99)	-	-	12.4 (3.6–63.2)
	Hakulinen et al., 2021 (n = 40)	FA	(0.74–0.94)	-	(4.8–11)	(4.1–14)
		ADC	(0.8–0.97)	-	(3–8.3)	(3.8–7.2)
		AD	(0.76–0.94)	-	(4–7.8)	(2.5–6)
		RD	(0.75–0.96)	-	(4.5–28.7)	(5.3–34.4)
	Shahim et al., 2017 (n = 4)	DTI	-	(3.5–7.4)	-	-
	Boukadi et al., 2019 (n = 18)	FA	(0.62–0.89)	-	-	-
		MD	(0.8–0.91)	-	-	-
		AD	(0.71–0.95)	-	-	-
		RD	(0.76–0.92)	-	-	-
		NuFO	(0.5–0.69)	-	-	-
	Andica et al., 2020 (n = 10)	FA	(0.74–0.99)	(0.2–1.9)	-	-
		MD	(0.77–0.96)	(0.2–3.3)	-	-
		AD	(0.83–0.99)	(0.1–2.5)	-	-
		RD	(0.82–0.99)	(0.2–4.1)	-	-
		OD	(0.78–0.99)	(0.2–3.8)	-	-
		Icwf	(0.81–0.98)	(0.4–1.5)	-	-
	Chung et al., 2016 (n = 8)	ISOvf	(0.15–0.98)	(1.6–18.5)	-	-
		DTI	-	-	(0.5–3.7)	(2.7–11.3)
		NODDI (excluded ISOvf)	-	-	(0.6–7.3)	(3.8–40)
	Tariq, 2013 (n = 4)	Icwf	-	5.13	-	-
		OD	-	5.71	-	-
		FA	-	4.43	-	-
		MD	-	4.09	-	-
MTI	Current study (n = 20) ^a	MTR	0.67 (0.60–0.75)	-	1.9 (1.4–2.5)	2.1 (1.5–2.7)
		MTsat	0.77 (0.69–0.84)	-	2.8 (1.9–3.8)	4.0 (2.7–5.0)
		ihMTR	0.48 (0.36–0.56)	-	8.3 (6.0–13.0)	6.7 (4.6–9.7)
		ihMTdR1sat	0.59 (0.44–0.67)	-	9.0 (6.5–13.8)	8.7 (5.9–11.8)
	Schwartz et al., 2019 (n = 1)	MTR	0.98	-	0.8 ^b	-
	Weiskopf et al., 2013 (n = 5)	MTsat	-	-	-	(0.6–0.8)

TABLE 2 (Continued)

MRI	Study	MRI measure	ICC	CV %	Within-CV %	Between-CV %
	Mchinda et al., 2018 (n = 22)	ihMTR	-	-	<2	<10
	Zhang et al., 2019 (n = 5)	ihMTR	(0.74–0.95)	-	-	-
	Wood et al., 2020 (n = 12)	MTR	(0.5–0.9)	-	-	-
		ihMTR	(0.3–0.75)	-	-	-

Note: Studies measures are reported as mean statistics for WM or (range) across WM regions. *n* = number of subjects; CV, Mean coefficient of variation; ADC, apparent diffusion coefficient.

^aReported as mean for all bundles and range across bundles.

^bAuthors used repeatability (RPT, agreement within site).

Intra Class Coefficient (ICC) profile

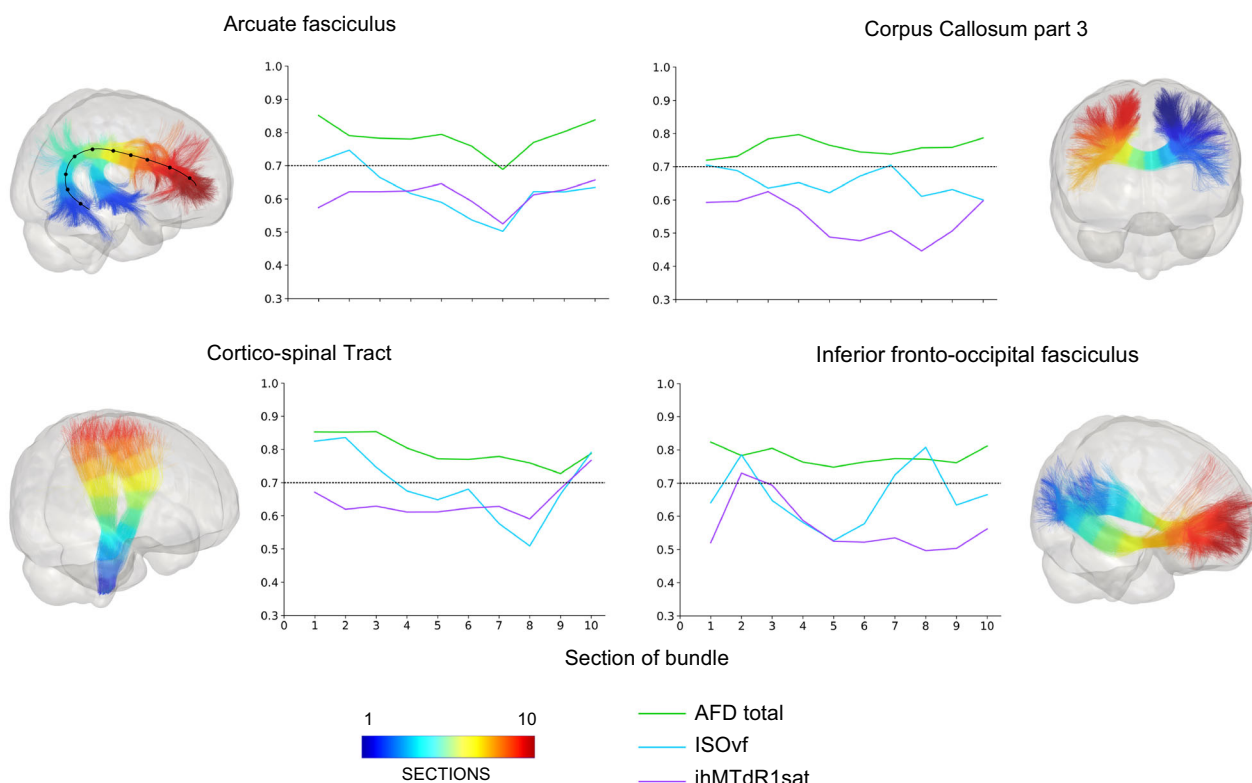


FIGURE 7 Consistency profile for selected bundles and three measures. The colors displayed on the bundles represent the section numbers from 1 (blue) to 10 (red) corresponding to the graphs. Each line of the graphs represents a measure with AFD total in green, ISOvf in light blue and ihMTdR1sat in purple. Only these three measurements for ICC results are displayed to facilitate the results' reading and represent the different profiles observed. On the AF bundle, we show a black line with dots representing each of the 10 track-section.

readthedocs.io/en/latest/results/consistency.html#profile-bundle-consistency).

More precisely, similarly to whole bundle results, consistency of DTI measures quantified along the bundles most often displays very stable profiles with higher ICCs (mean across section = 0.8 [0.64–0.95]) and lower CVs values (CVb: 4.5% [0.8%–13%], CVw: 2.8% [0.7%–9.5%]) with respect to other measures (Figure 7). The consistency profile of HARDI measures show analogous patterns to the DTI-derived ones, with similar ICCs and CVs values for AFD total

(Figure 7, green line) and lower ICCs and higher CVs values for NuFO (ICC: 0.72 [0.5–0.88], CVb: 6.2% [1.1%–14.1%], CVw: 4.4% [0.9%–14.5%]). For NODDI measurements, ICCs and CVs profiles show more variability along the bundles (Figure 7), but overall tend to be rather high to moderate for ICCs (0.76 [0.5–0.95]) and low for CVs, with mean value across the section of 8% [0.9%–28%] and 6.4% [0.8%–21%] for CVb and CVw respectively (Figure 7). Again, OD shows stable and higher consistency measurements in contrast to ISOvf whose consistency profile is more variable (Figure 7, light blue line).

Regarding the MTI measures, the consistency exhibits globally stable profiles for most bundles, with more stable CV profiles than ICC profiles. As for the whole bundle results, the highest ICCs and lowest CVs values are found for MTsat and MTR measures, while ihMT measures showed lower reliability ($ICC \sim 0.51$ [0.3–0.82]) and higher variability ($CVw \sim 7.6\%$ [3.1%–15.6%] and $CVb \sim 6.5\%$ [1.9%–15%]) (purple line in Figure 7).

3.7 | Impact of fiber population on consistency

For all measures, the bundle compartmentalization into “single” and “multi” fiber population regions affects ICCs and CVs values regardless of the measure (Figure 8, Table S4, https://high-frequency-mri-database-supplementary.readthedocs.io/en/latest/results/fiber_population_consistency.html#whole-bundle-consistency). Compared with ICC computed from average bundle, the compartmentalization revealed higher ICCs (mean across bundles, ICC: 0.72, 0.82 and 0.81 for average, multi- and single-compartment, respectively) and, equal to lower CVs across bundles (mean across bundles, CVb/CVw : 6.7%/5.2%, 3.7%/2.7% and 4.6%/3.4% for average, multi and single-compartment respectively). More precisely, ICCs measures from multi-compartment were higher, while the ICCs value observed for single compartment is lower and inversely for CVs (see <https://high-frequency-mri-database-supplementary.readthedocs.io/en/latest/res>

[ults/fiber_population_consistency.html#whole-bundle-consistency](https://high-frequency-mri-database-supplementary.readthedocs.io/en/latest/results/fiber_population_consistency.html#whole-bundle-consistency)). Differences between single and multi-compartment can be moderate, especially for DTI measurements, or more important such as MTI measurements. Some measures exhibit a different pattern than others. This is the case for the OD measurement of the NODDI model, whose reliability pattern is inverted. As expected, since OD is a measure of the dispersion of fiber orientation in the voxel, a higher or equal ICCs to the average for single compartment compared with multiple compartments for most bundles is not surprising. In addition, the impact of compartmentalization has a more moderate effect on the consistency of SLF bundles measures. Indeed, the values of ICCs and CVs are generally equal (or present a slight difference) between the two compartments. This moderate effect is also found for DTI measures, for most, but not all, bundles (Figure 8, see RD as an example, red line).

The compartmentalization also affects the mean values of the measurements. The mean of the measures in the multi-compartment was lower than those found in the single-compartment (e.g., mean CST FA: 0.47, 0.53 and 0.35 for average, single- and multi-compartment respectively). Globally, there is a decrease (or increase depending on the measure) in the mean values of the measures with increasing fiber populations. The distribution profile of measures is available https://high-frequency-mri-database-supplementary.readthedocs.io/en/latest/results/fiber_population_measures.html#profile-bundle-measures.

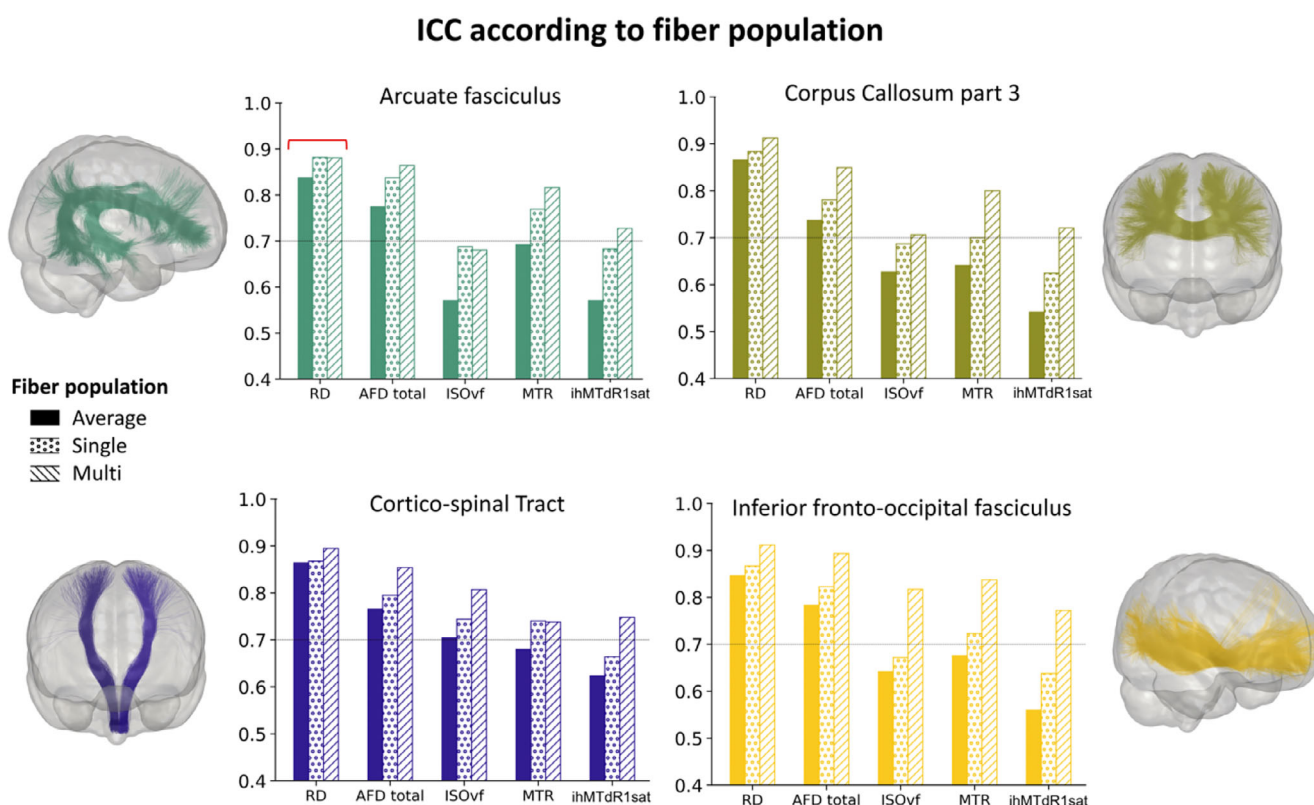


FIGURE 8 Impact of fiber populations on bundle-averaged consistency values for selected bundles. The colors correspond to the bundles. The bars represent the ICCs values, full for “average,” dotted hatch for “single” and diagonal hatch for “multi” compartment.

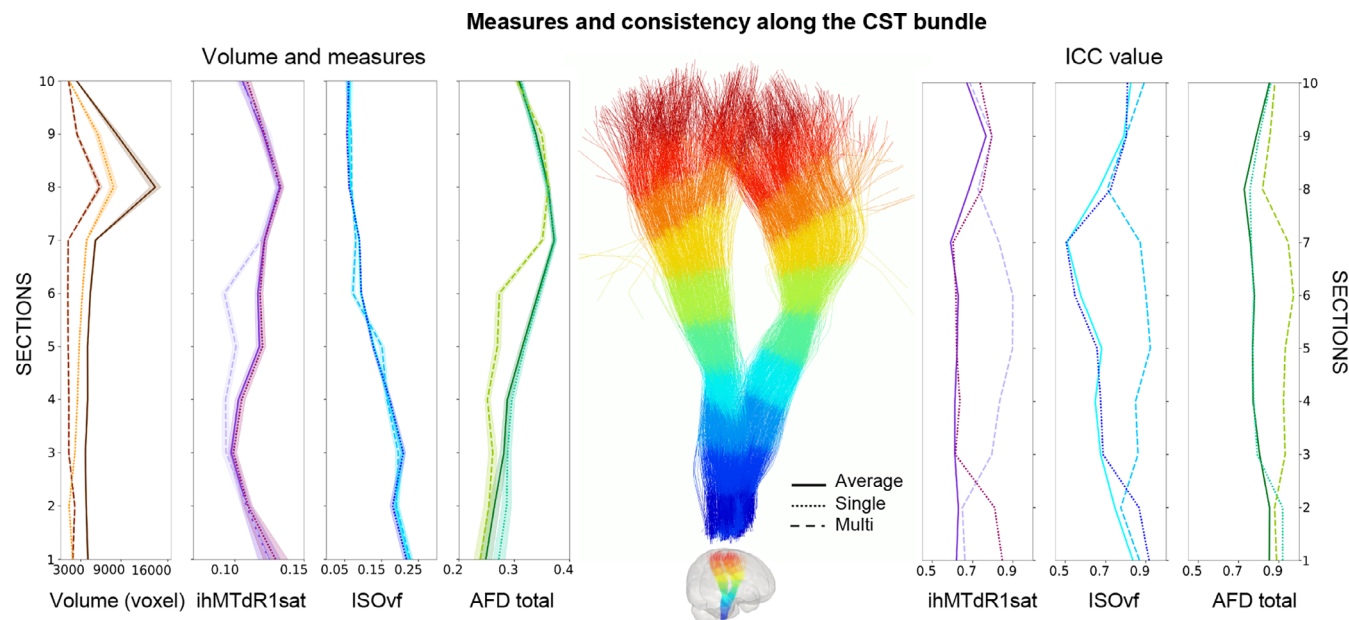


FIGURE 9 Volume, measures and consistency values per section and compartment of the CST bundle. The left panel shows volume and measures, the right panel shows ICC value. Continuous line represents values obtained from bundle average, dashed line represents value from multi-compartment and dot line represents value from single compartment. Colors correspond to measures, volume in brown, ihMTdR1sat in purple, isoVF in light blue and AFD total in green.

Figure 9 presents profile examples for some measures and the associated consistency values for CST bundle. The compartmentalization of the bundle profiles into “single” and “multi” fiber population regions showed a similar effect to the results on the whole bundle profiles, with higher multi-compartment ICCs compared to the single-compartment and, inversely for CVs (Figure 9, https://high-frequency-mri-database-supplementary.readthedocs.io/en/latest/results/fiber_population_consistency.html#profile-bundle-consistency). The single-compartment profile is generally close to or equal to the whole bundle profile, while the multi-compartment is more distant from the average profile. Some bundles present a different pattern for the inferior (sections 1–3) and superior (sections 8–10; or anterior/posterior) sections compared to middle sections. For example, CST exhibits a higher ICC in the single compartment for inferior sections (1–2), whereas the superior sections (8–10) show weak differences between the compartments compared to the middle part of the bundle (sections 3–7, Figure 9). Again, SLF bundles show globally the most stable profiles with a low impact of compartmentalization on consistency, whereas the CC bundles show the reverse pattern. Finally, this compartmentalization effect is also found in the measurement profiles (Figure 9). This effect is greater for some measures such as FA or OD while others are less affected such as ISOvf, ICvf, MD or MTR. Like consistency results, the compartmentalization does not uniformly affect the measurement profiles of bundles according to the sections, with a different effect between the sections in the middle and those at the ends of the bundle (Figure 9).

4 | DISCUSSION

Using a “high-frequency” repeated-measure study collected from 20 healthy subjects, we assessed the consistency of multiple WM

microstructural measures across the bundles, with special emphasis on the most frequently used image analysis approaches. The Dice scores reveal a good spatial agreement of the segmentation of the bundles across the sessions, except for SLF 1 whose results should be taken with caution. The results show that the reliability and variability of DWI measures are good (ICC >0.7; CVw and CVb < 15%) across the bundles and especially excellent for most DTI measures as well as AFD total and OD index from the HARDI and NODDI models, respectively (CVw and CVb <5%, ICC >0.75). In addition, the profile consistencies of these measures are comparable to the whole bundles, with voxel values well above ICC >0.7 and CVw <4% along the bundle. In contrast, MTI showed good reliability and variability for MT measurements (CV <8%, ICC ≥0.7) and moderate for ihMT measurements (CV < 15%, ICC ≥0.5). We also showed that the number of fiber populations affects the consistency of the measurements, with a moderate effect on the DTI and HARDI measurements, and a more important effect on the NODDI and MTI measurements. Finally, SLF 2–3 and CC 4 to CC 6 bundles showed the most consistent MRI measurements followed by AF, CST, CC 7, OR, CC3, IFOF and CG, while CC 2a, CC 2b, UF and ILF have more moderate consistency measurements (Figure 10).

In this study, some choices have been made for acquisition schemes, MRI techniques or processing tools to do all the consistency analyses. We know that many other options are available for similar tasks and ways to assess reliability and variability. Our protocol is a compromise between the literature and the constraints imposed by the scanner. Indeed, the choice of b-values at 300, 1000 and 2000 with a specific non-equal distribution of the 100 directions is optimized for the balance between signal-to-noise ratio (SNR) and angular contrast ratio, while maintaining a reasonably short acquisition time. Low b-values of 300 and 1000 with 8 and 32 directions (respectively)

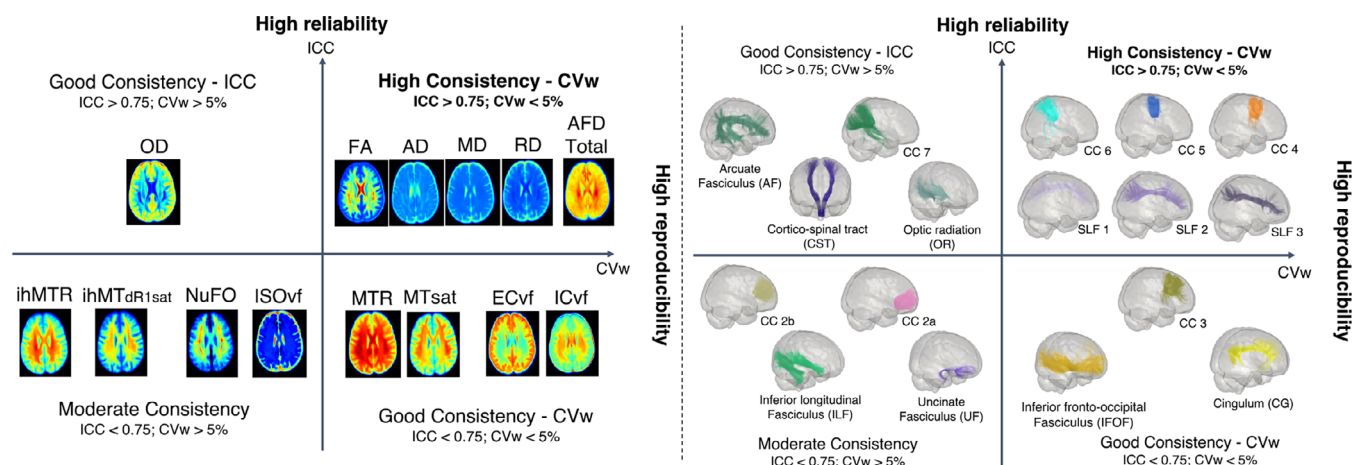


FIGURE 10 Classification of bundles and measures into four groups according to ICC values and within-variability (CVw). The x-axis represents CVw values (i.e., reproducibility), and the y-axis represents ICC values (i.e., reliability). High consistency group represents low CVw value <0.05 (5%) and high ICC value >0.75 (high reliability); Good consistency–ICC: high ICC value >0.75 but high CVw value >0.05 (>5%); Good consistency–CVw: low CVw value <0.05 but low ICC value <0.75 and finally, Moderate consistency: high CVw value >0.05 (>5%) and low ICC value <0.75.

are used to obtain a robust and unbiased estimate of the DTI and the mono-exponential part of the signal (Jones et al., 1999). Indeed, dMRI data with a high b value often have a non-Gaussian distribution, therefore applying a DTI model to such dMRI data may result in measurement biases compared to those using a b value ≤ 1200 (Jones et al., 1999). Then, a higher b -value of 2000 with 60 directions is used to obtain the angular contrast of the signal. Previous papers have shown a good compromise with acquisition schemes around 60 directions with b -values between 1500 and 3000 (Dell'Acqua et al., 2007; Dell'Acqua & Tournier, 2019; Tournier et al., 2004). Moreover, a b -value of 2000 helps the NODDI model and more advanced microstructure models to estimate non-Gaussian phenomena in the signal (Parvathaneni et al., 2018). It is also important to note that the dMRI protocol would have been more optimal had we acquired every DWI volume in both forward and reversed phase-encoding directions (Irfanoglu et al., 2021). However, this would have doubled the acquisition time and the current protocol was designed to run on a clinical population after our healthy control assessment. An acquisition protocol with an equal distribution of forward and reversed phase-encoding could be further investigated in the future. In the same way, there are different methods for performing tractography (Maier-Hein et al., 2017; Yendiki et al., 2011) or identifying bundles (Schilling et al., 2021). Additionally, several other microstructural measures can be characterized. All these parameters could influence both variability and reliability. The intent here was not to provide an analysis between different processing tools or parameters, and therefore we do not recommend using the consistency values from this study as an absolute value. Instead, we aimed to contribute to a global understanding of the variability and reliability of popular WM imaging techniques such as DWI as well as newer techniques such as ihMT in healthy subjects. However, considering that each pathology will affect variability and reliability in a different way, it is also important to investigate it in the disease states.

4.1 | Diffusion consistency measures of whole and along-bundle profiling

Here, we demonstrated good to excellent consistency, for both the whole bundle and along the bundle, of most diffusion measures used in MRI studies. More precisely, the reliability and variability of DTI measurements are comparable or superior to the previous DTI studies, with ICC above 0.7 and within- and between-CV around 5% and 10% (see Table 2 for a comparison with existing literature) (Acheson et al., 2017; Grech-Sollars et al., 2015; Hakulinen et al., 2021; Palacios et al., 2017; Shahim et al., 2017; Veenith et al., 2013; Zhou et al., 2018). Regarding NuFO, our results are concordant with the only study that has assessed the reliability of this measure in the test-retest study and reports moderate reliability with an ICC of ~ 0.6 across four bundles (Boukadi et al., 2019, Table 2). This highlights the need for further validation of this measure before adopting it in longitudinal or clinical studies. On the other hand, even though NODDI measurements are inherently noisier than DTI measurements for white matter modeling—likely due to a more complex model and requiring high b -value data—within- and between-variability for ICvf and ECvf were lower or comparable to the FA variability in most bundles (Andica et al., 2020; Chung et al., 2016; Lucignani et al., 2021). However, this sensitivity to noise inherent to NODDI could explain the higher within-variability of ISOvf and OD. Indeed, despite the additional thresholding applied on ISOvf, this measure systematically presents the largest CV in all the bundles, which is consistent with studies that suggest that ISOvf is a poorly reliable parameter (Tariq, 2013, 2018). Finally, our findings of overall greater CVb for most measures compared with their corresponding CVw are in accordance with other DTI- and NODDI-based reliability studies (Andica et al., 2020; Chung et al., 2016; Lucignani et al., 2021; Tariq, 2013; Veenith et al., 2013). This result is expected, as there will be greater microstructural heterogeneity in a

localized region across a population compared with multiple observations within the same subject.

4.2 | MTI consistency measures of whole and along-bundle profiling

Although our MT measurements were derived from ihMT images, our results are consistent with previous studies see Table 2 for a comparison with existing literature (Schwartz et al., 2019; van der Weijden et al., 2021; Weiskopf et al., 2013; Wood et al., 2020). Our ihMT measurements show equivalent (ICC, Wood et al., 2020) or lower values than the previous studies (CVw: Mchinda et al., 2018; ICC: Zhang et al., 2019). Discrepancies with literature could be explained by, first the different ways to assess variability and reliability measures, then the difference for ihMT sequences with a long cosine-modulated RF pulse, without any T1D filtering in Zhang's report instead of a train of short RF pulses with moderate T1D filtering in this study (Zhang et al., 2019); and Mchinda's report an ihMTR between-subject variability based on 1.5 T MRI instead of 3 T MRI in this study (Mchinda et al., 2018). In addition, Varma et al. reported that ihMTR could be varied depending on different sets of saturation parameters, especially when the saturation time is in the range of 0–200 ms (Ercan et al., 2018; Varma et al., 2015). This discrepancy between studies suggests a strong effect of experimental design and local factors on the measurements and, therefore, makes it difficult to compare our results with each other. Nevertheless, we can agree that among the MTI-derived measures, the ihMT measures appear to be more variable and less reliable than the other measures. Noted that considering these measurements are derived from the same acquisition, the ihMT effect (i.e., SNR) being much weaker than a standard MT effect, it is expected that the variability of ihMT will always be greater than the MT variability. However, ihMT is more specific to myelin (Duhamel et al., 2019; Prevost et al., 2018), which is a significant advantage in clinical studies. Moreover, B1+ inhomogeneities can bias measurements when appropriate protocols or specific processing are not used, affecting the reliability of ihMT measurements. Thus, as with all new contrast which are not standardized yet, further optimization including the acquisition parameters might improve the accuracy of ihMT measures and, by extension, their consistency under different experimental conditions.

4.3 | Consistency and fiber population

Results emerging from the single versus multi-fiber population impact show that the fiber populations affect the consistency of measures. The higher the fiber populations of a bundle (i.e., multi fibers), the higher the ICC and lower the within- and between-subject CV (and inversely). Despite it seems like counter-intuitive results, a smaller range of values for the compartment with multiple fiber populations for most measures compared to the single compartment explains the greater consistency (CST FA range in multi: 0.34–0.42 vs. single 0.48–

0.61). Thus, the intrinsic variability of the measurements in the multi-compartment is already more restricted than in the single compartment. Although the method and the variability values are not directly comparable, Volz et al. also report greater variability of the compartment with a single fiber population compared to compartments with multiple fiber populations (Volz et al., 2018). In agreement with this study, diffusion and myelin measurements per se are also impacted by compartmentalization, with higher value in the single compartment compared to the multiple compartments (and vice versa depending on the measurements). This reinforces the idea that the organization of the underlying WM affects both measures and consistency. Moreover, some bundles are more affected than others, this is notably the case of the CC which, unlike the SLF bundles, displays a more pronounced compartmentalization effect. Beyond the bundle, we also showed that some measurements are more affected than others, suggesting that the affected measurements are sensitive to the orientation of the fibers of the white matter. This has been described for the ihMT which shows a dependence on the orientation of the fibers with respect to B0. Thus, in the presence of several fiber populations, this effect could be averaged, thereby improving the consistency of the measurements. Together, these results suggest that the dissociation of voxels according to the number of fibers populations may be relevant and must be tested in future studies.

4.4 | What is the contribution of this study to clinical and research studies?

Assessing the consistency of measurements extracted from MRI images represents an important step toward validating this approach in longitudinal studies and clinical trials. Clinical applications favor measures with high reliability, which optimizes a trade-off between the two variability components with low within-variability (i.e., more stable across different measuring times) and high between-variability (i.e., more differentiable across participants). This variability pattern is a necessary criterion for the high validity of a biomarker, which can be used to diagnose, monitor and predict neurological consequences by clarifying the effects of a disease or treatment. However, another important point in clinical practice is the distinction between healthy and pathological states, for which high between-variability is not required.

As expected, the measures corresponding to this criterion are the “simplest” that is, the DTI-derived measures, particularly FA, RD, AD and MD. However, more advanced measures could also be promising candidates as biomarkers. Indeed, axonal density and neuroinflammation indices can be interesting in pathologies where the axons are altered and disrupted, such as Alzheimer's disease (AD, AFD total: Roy et al., 2021), or predictive of response to treatment (ICvf and ECvf: Dowell et al., 2019; Kraguljac et al., 2019; Sarrazin et al., 2019), reinforcing their potential as clinically relevant biomarkers. On the other hand, recent studies show that it is important from the biological mechanisms point of view to consider fiber orientation dispersion, especially in the crossing fibers regions such as the semioval centrum

(Andersen et al., 2020; Chad et al., 2021; Mito et al., 2018; Schilling et al., 2020). A DTI study showed in MS patients that considering the orientation of the fibers for the FA reveals changes in WM related to disability, while the standard FA failed to do so (Andersen et al., 2020). Moreover, a fixel-based analysis (Raffelt et al., 2017) has revealed a specific degeneration of the SLF alongside the preservation of the CST and the CC in AD, whereas this results in an increase in FA with conventional DTI measures, that is, the decrease in FA of the SLF bundle leads to a “virtual” increase in FA in the CST (Mito et al., 2018). Therefore, distinguishing different fiber populations reveals a change in bundle-specific MRI measurements, without causing significant abnormality in other bundles that intersect in the same region (Doan et al., 2017; Lee et al., 2015; Mito et al., 2018). Here, the orientation dispersion (OD) is a measure of fiber orientation in the voxel whose reliability is good and which covaries with the NuFO (Chamberland et al., 2019). The latter displays more moderate reliability, but unlike OD, NuFO can be extracted from clinical acquisitions without requiring multi-shell DWI acquisition and advanced modeling, which is an important advantage for clinical studies. This emphasizes the critical role of different fiber orientations, and their clinical potential, in understanding brain disease processes where conventional measurements are “blind.”

Although the MTR is the most used measure, this study shows that the MTsat presents a lower within-variability and a greater between-variability compared to the MTR (<https://high-frequency-mri-database-supplementary.readthedocs.io/en/latest/results/consistency.html#within-variability>). Again, this indicates that MTsat would be a more favorable biomarker than MTR. Moreover, recent studies have shown that MTsat is more sensitive than MTR in MS (Granziera et al., 2021; Lema et al., 2017; Saccenti et al., 2020). Despite lower consistency, recent studies support the use of ihMT in clinical studies due to its specificity to myelin, especially in patients with MS (Obberghen et al., 2018; Prevost et al., 2018; Varma et al., 2015). Indeed, albeit preliminary due to the small number of subjects, a recent study showed that in MS patients ihMTR was correlated with clinical disability, whereas MTR failed to do so (Obberghen et al., 2018). Moreover, many efforts have been made recently to overcome significant technical limitations—especially on its feasibility on different scanners—and make ihMT an applicable tool in daily clinical practice that may outperform MT measures soon (Girard et al., 2015; Soustelle et al., 2022; Varma et al., 2018; Wood et al., 2020).

All bundles that show high consistency, whether extracted on whole or along the bundles, are involved in different pathologies, aging or development. This is particularly the case for AF, ILF, IFOF, CST, UF and CG which have very good consistency per se (Dice score) (Atkinson-Clement et al., 2017; Beaudoin et al., 2021; Bergamino et al., 2020; Coelho et al., 2021). A more moderate observation can be made regarding some parts of the corpus callosum, which showed more variable consistency of data. Beyond this, the profiles reveal good to excellent levels of coherence, like the measurements from whole bundle. Interestingly, the consistency levels and measurements vary according to the sections along the bundle. This suggests that along-bundle profiling could help to reliably highlight more subtle

changes such as the presence of a white matter lesion, which may impact one or more parts of the bundle without affecting the whole bundle. This is supported by a recent study which shows that AFD profiles are affected by the presence of a lesion and are significantly different at the location of WMH in fronto-pontine tracts bundle in mild-cognitive impairment subjects with high WM lesion load (Kim et al., 2022).

5 | CONCLUSION

Using a “high-frequency” repeated-measure study collected from 20 healthy subjects, we showed that the reliability and variability of DWI measures are excellent to good across the bundles as well as along the bundle. In contrast, MTI showed good reliability and variability for MT measurements and moderate for ihMT measurements. We also showed that the number of fiber populations affects the consistency of the measurements, with a moderate effect on the DTI and HARDI measurements, and a more important effect on the NODDI and MTI measurements. Finally, the most consistent MRI measurements are found for SLF 2–3 and CC 4 to CC 6 bundles, then for AF, CST, CC 7, OR, CC3, IFOF and CG, while CC 2a, CC 2b, UF and ILF have more moderate consistency measurements.

AUTHOR CONTRIBUTIONS

Manon Edde drafted the manuscript, contributed to the design of the study, reviewed the literature, and contributed to data collection, analyzed, and interpreted the data. Maxime Descoteaux designed the study, supervised data analysis and interpretation, and contributed to drafting the manuscript. Maxime Descoteaux and Guillaume Gilbert contributed to MRI data acquisition. Antoine Théberge contributed to the design of the study and data collection. Manon Edde, Guillaume Theaud, and Guillaume Gilbert contributed to developing the inhomogeneous magnetization transfer pipeline. Alex Valcourt-Caron, Guillaume Theaud and Matthieu Dumont helped with the data analysis and Alex Valcourt-Caron developed the voxel-level consistency analysis script. Jean-Christophe Houde, Guillaume Theaud, Loika Maltais and Matthieu Dumont contributed to image processing. François Rheault performed the bundle's consistency analysis. All authors revised the final version of the manuscript. Stefano Magon designed the study, data interpretation, and review the manuscript. Muhamed Barakovic and Federico Spagnolo helped with the data interpretation and review the manuscript.

ACKNOWLEDGMENTS

The authors would like to thank the participants who accepted to take part in this study. The authors also thank the radiology team at Centre Hospitalier Universitaire of Sherbrooke (CHUS) for their support with the scanning of participants.

FUNDING INFORMATION

Part of this research was supported by the NSERC Discovery grant (www.nserc-crsng.gc.ca), the Université de Sherbrooke Institutional Chair in Neuroinformatics from Pr Descoteaux (www.usherbrooke.ca)

and Mitacs Accelerate program (www.mitacs.ca). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

CONFLICT OF INTEREST STATEMENT

Maxime Descoteaux is co-owner and chief scientific officer at Imeka Solutions Inc. (available online: www.imeka.ca [accessed on August 5, 2021]). Federico Spagnolo is an employee of F. Hoffmann-La Roche Ltd. Muhamed Barakovic is an employee of Hays plc and a consultant for F. Hoffmann-La Roche Ltd. Stefano Magon is an employee and shareholder of F. Hoffmann-La Roche Ltd. Guillaume Gilbert is an employee of Philips Healthcare. Guillaume Theaud, Matthieu Dumont, Jean-Christophe Houde and Loika Maltais are employees at Imeka Solutions Inc.

DATA AVAILABILITY STATEMENT

The raw and generated data sets during the current study are not publicly available because we are not the exclusive owners of the data. However, they are available from the corresponding author under reasonable request and data sharing agreements with the owners.

CODE AVAILABILITY

All Nextflow and code used to process the T1 and DWI images, ihMT images, RecoBundlesX and Tractometry are available at <https://github.com/scilus>. Operations like generating and eroding masks were done using Scilpy scripts available at <https://github.com/scilus/scilpy>. Consistency analysis code is available at https://github.com/AlexVCaron/longitudinal_image_statistics.

ORCID

Manon Edde  <https://orcid.org/0000-0002-4181-8031>

François Rheault  <https://orcid.org/0000-0002-0097-8004>

REFERENCES

- Acheson, A., Wijtenburg, S. A., Rowland, L. M., Winkler, A., Mathias, C. W., Hong, L. E., Jahanshad, N., Patel, B., Thompson, P. M., McGuire, S. A., Sherman, P. M., Kochunov, P., & Dougherty, D. M. (2017). Reproducibility of tract-based white matter microstructural measures using the ENIGMA-DTI protocol. *Brain and Behavior*, 7(2), e00615. <https://doi.org/10.1002/brb3.615>
- Alexander, G. E. (2017). An emerging role for imaging White matter in the preclinical risk for Alzheimer disease: Linking β -amyloid to myelin. *JAMA Neurology*, 74(1), 17–19. <https://doi.org/10.1001/jamaneurol.2016.4123>
- Alsop, D. C., Ercan, E., Girard, O. M., Mackay, A. L., Michal, C. A., Varma, G., Vinogradov, E., & Duhamel, G. (2022). Inhomogeneous magnetization transfer imaging: Concepts and directions for further development. *NMR in Biomedicine*, e4808. <https://doi.org/10.1002/nbm.4808>
- Andersen, K. W., Lasič, S., Lundell, H., Nilsson, M., Topgaard, D., Sellebjerg, F., Szczepankiewicz, F., Siebner, H. R., Blinkenberg, M., & Dyrby, T. B. (2020). Disentangling white-matter damage from physiological fibre orientation dispersion in multiple sclerosis. *Brain Communications*, 2(2), fcaa077. <https://doi.org/10.1093/braincomms/fcaa077>
- 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(2), 870–888. [https://doi.org/10.1016/S1053-8119\(03\)00336-7](https://doi.org/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(4), 483–494. <https://doi.org/10.1007/s00234-019-02350-6>
- Arnold, D. L., Piani-Meier, D., Bar-Or, A., Benedict, R. H., Cree, B. A., Giovannoni, G., Gold, R., Vermersch, P., Arnould, S., Dahlke, F., Hach, T., Ritter, S., Karlsson, G., Kappos, L., & Fox, R. J. (2022). Effect of siponimod on magnetic resonance imaging measures of neurodegeneration and myelination in secondary progressive multiple sclerosis: Gray matter atrophy and magnetization transfer ratio analyses from the EXPAND phase 3 trial. *Multiple Sclerosis Journal*, 28, 1526–1540. <https://doi.org/10.1177/13524585221076717>
- Atkinson-Clement, C., Pinto, S., Eusebio, A., & Coulon, O. (2017). Diffusion tensor imaging in Parkinson's disease: Review and meta-analysis. *NeuroImage: Clinical*, 16, 98–110. <https://doi.org/10.1016/j.nicl.2017.07.011>
- Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, 12(1), 26–41. <https://doi.org/10.1016/j.media.2007.06.004>
- Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., & Gee, J. C. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage*, 54(3), 2033–2044. <https://doi.org/10.1016/j.neuroimage.2010.09.025>
- Basser, P. J., Mattiello, J., & LeBihan, D. (1994). MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, 66(1), 259–267. [https://doi.org/10.1016/S0006-3495\(94\)80775-1](https://doi.org/10.1016/S0006-3495(94)80775-1)
- Beaudoin, A.-M., Rheault, F., Theaud, G., Laberge, F., Whittingstall, K., Lamontagne, A., & Descoteaux, M. (2021). Modern technology in multi-shell diffusion MRI reveals diffuse white matter changes in young adults with relapsing-remitting multiple sclerosis. *Frontiers in Neuroscience*, 15, 5017. <https://doi.org/10.3389/fnins.2021.665017>
- Beck, D., de Lange, A.-M. G., Maximov, I. I., Richard, G., Andreassen, O. A., Nordvik, J. E., & Westlye, L. T. (2021). White matter microstructure across the adult lifespan: A mixed longitudinal and cross-sectional study using advanced diffusion models and brain-age prediction. *NeuroImage*, 224, 117441. <https://doi.org/10.1016/j.neuroimage.2020.117441>
- Bergamini, M., Keeling, E. G., Mishra, V. R., Stokes, A. M., & Walsh, R. R. (2020). Assessing white matter pathology in early-stage Parkinson disease using diffusion MRI: A systematic review. *Frontiers in Neurology*, 11, 314. <https://doi.org/10.3389/fneur.2020.00314>
- 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(1), 24–33. <https://doi.org/10.1111/psyp.12769>
- Boukadi, M., Marcotte, K., Bedetti, C., Houde, J.-C., Desautels, A., Deslauriers-Gauthier, S., Chapleau, M., Boré, A., Descoteaux, M., & Brambati, S. M. (2019). Test-retest reliability of diffusion measures extracted along White matter language fiber bundles using HARDI-based Tractography. *Frontiers in Neuroscience*, 12, 1055. <https://doi.org/10.3389/fnins.2018.01055>
- Briggs, A. H., Wonderling, D. E., & Mooney, C. Z. (1997). Pulling cost-effectiveness analysis up by its bootstraps: A non-parametric approach to confidence interval estimation. *Health Economics*, 6(4), 327–340. [https://doi.org/10.1002/\(SICI\)1099-1050\(199707\)6:4<327::AID-HEC282>3.0.CO;2-W](https://doi.org/10.1002/(SICI)1099-1050(199707)6:4<327::AID-HEC282>3.0.CO;2-W)
- Brown, J. W. L., Pardini, M., Brownlee, W. J., Fernando, K., Samson, R. S., Prados Carrasco, F., Ourselin, S., Gandini Wheeler-Kingshott, C. A. M., Miller, D. H., & Chard, D. T. (2017). An abnormal periventricular

- magnetization transfer ratio gradient occurs early in multiple sclerosis. *Brain*, 140(2), 387–398. <https://doi.org/10.1093/brain/aww296>
- Bruton, A., Conway, J. H., & Holgate, S. T. (2000). Reliability: What is it, and how is it measured? *Physiotherapy*, 86(2), 94–99. [https://doi.org/10.1016/S0031-9406\(05\)61211-4](https://doi.org/10.1016/S0031-9406(05)61211-4)
- Cai, L. Y., Kerley, C. I., Yu, C., Aboud, K. S., Beason-Held, L., Shafer, A. T., Resnick, S. M., Jordan, L. C., Anderson, A. W., Schilling, K. G., Lyu, I., & Landman, B. A. (2021). Joint cortical surface and structural connectivity analysis of Alzheimer's disease. *Medical Imaging 2021: Image Processing*, 11596, 1159630. <https://doi.org/10.1117/12.2580956>
- Caruyer, E., Lenglet, C., Sapiro, G., & Deriche, R. (2013). Design of multi-shell sampling schemes with uniform coverage in diffusion MRI. *Magnetic Resonance in Medicine*, 69(6), 1534–1540. <https://doi.org/10.1002/mrm.24736>
- Catani, M., & de Schotten, M. T. (2012). *Atlas of human brain connections*. Oxford University Press.
- Cercignani, M., & Bouyagoub, S. (2018). Brain microstructure by multi-modal MRI: Is the whole greater than the sum of its parts? *NeuroImage*, 182, 117–127. <https://doi.org/10.1016/j.neuroimage.2017.10.052>
- Chad, J. A., Pasternak, O., & Chen, J. J. (2021). Orthogonal moment diffusion tensor decomposition reveals age-related degeneration patterns in complex fiber architecture. *Neurobiology of Aging*, 101, 150–159. <https://doi.org/10.1016/j.neurobiolaging.2020.12.020>
- Chamberland, M., Raven, E. P., Genc, S., Duffy, K., Descoteaux, M., Parker, G. D., Tax, C. M. W., & Jones, D. K. (2019). Dimensionality reduction of diffusion MRI measures for improved tractometry of the human brain. *NeuroImage*, 200, 89–100. <https://doi.org/10.1016/j.neuroimage.2019.06.020>
- 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. *Human Brain Mapping*, 37(12), 4550–4565. <https://doi.org/10.1002/hbm.23328>
- Coelho, A., Fernandes, H. M., Magalhães, R., Moreira, P. S., Marques, P., Soares, J. M., Amorim, L., Portugal-Nunes, C., Castanho, T., Santos, N. C., & Sousa, N. (2021). Signatures of white-matter microstructure degradation during aging and its association with cognitive status. *Scientific Reports*, 11(1), 1. <https://doi.org/10.1038/s41598-021-83983-7>
- Cousineau, M., Jodoin, P.-M., Morency, F. C., Rozanski, V., Grand'Maison, M., Bedell, B. J., & Descoteaux, M. (2017). A test-retest study on Parkinson's PPMI dataset yields statistically significant white matter fascicles. *NeuroImage: Clinical*, 16, 222–233. <https://doi.org/10.1016/j.nicl.2017.07.020>
- 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. <https://doi.org/10.1016/j.neuroimage.2014.10.026>
- Dell'Acqua, F., Rizzo, G., Scifo, P., Clarke, R. A., Scotti, G., & Fazio, F. (2007). A model-based deconvolution approach to solve fiber crossing in diffusion-weighted MR imaging. *IEEE Transactions on Biomedical Engineering*, 54(3), 462–472. <https://doi.org/10.1109/TBME.2006.888830>
- Dell'Acqua, F., Simmons, A., Williams, S. C. R., & Catani, M. (2013). Can spherical deconvolution provide more information than fiber orientations? Hindrance modulated orientational anisotropy, a true-tract specific index to characterize white matter diffusion. *Human Brain Mapping*, 34(10), 2464–2483. <https://doi.org/10.1002/hbm.22080>
- Dell'Acqua, F., & Tournier, J.-D. (2019). Modelling white matter with spherical deconvolution: How and why? *NMR in Biomedicine*, 32(4), e3945. <https://doi.org/10.1002/nbm.3945>
- Descoteaux, M., Angelino, E., Fitzgibbons, S., & Deriche, R. (2007). Regularized, fast, and robust analytical Q-ball imaging. *Magnetic Resonance in Medicine*, 58(3), 497–510. <https://doi.org/10.1002/mrm.21277>
- Diciccio, T. J., & Romano, J. P. (1988). A review of bootstrap confidence intervals. *Journal of the Royal Statistical Society: Series B: Methodological*, 50(3), 338–354. <https://doi.org/10.1111/j.2517-6161.1988.tb01732.x>
- Doan, N. T., Engvig, A., Persson, K., Alnæs, D., Kaufmann, T., Rokicki, J., Córdova-Palomera, A., Moberget, T., Brækhus, A., Barca, M. L., Engedal, K., Andreassen, O. A., Selbæk, G., & Westlye, L. T. (2017). Dissociable diffusion MRI patterns of white matter microstructure and connectivity in Alzheimer's disease spectrum. *Scientific Reports*, 7(1), 1. <https://doi.org/10.1038/srep45131>
- Dowell, N. G., Bouyagoub, S., Tibble, J., Voon, V., Cercignani, M., & Harrison, N. A. (2019). Interferon-alpha-induced changes in NODDI predispose to the development of fatigue. *Neuroscience*, 403, 111–117. <https://doi.org/10.1016/j.neuroscience.2017.12.040>
- Duhamel, G., Prevost, V. H., Cayre, M., Herten, A., Mchinda, S., Carvalho, V. N., Varma, G., Durbec, P., Alsop, D. C., & Girard, O. M. (2019). Validating the sensitivity of inhomogeneous magnetization transfer (ihMT) MRI to myelin with fluorescence microscopy. *NeuroImage*, 199, 289–303. <https://doi.org/10.1016/j.neuroimage.2019.05.061>
- Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. Chapman and Hall/CRC. <https://doi.org/10.1201/9780429246593>
- Ercan, E., Varma, G., Mädler, B., Dimitrov, I. E., Pinho, M. C., Xi, Y., Wagner, B. C., Davenport, E. M., Maldjian, J. A., Alsop, D. C., Lenkinski, R. E., & Vinogradov, E. (2018). Microstructural correlates of 3D steady-state inhomogeneous magnetization transfer (ihMT) in the human brain white matter assessed by myelin water imaging and diffusion tensor imaging. *Magnetic Resonance in Medicine*, 80(6), 2402–2414. <https://doi.org/10.1002/mrm.27211>
- Filippi, M., Inglese, M., Rovaris, M., Sormani, M. P., Horsfield, M. A., Iannucci, G., Colombo, B., & Comi, G. (2000). Magnetization transfer imaging to monitor the evolution of MS: A 1-year follow-up study. *Neurology*, 55(7), 940–946. <https://doi.org/10.1212/WNL.55.7.940>
- Garyfallidis, E., Côté, M.-A., Rheault, F., Sidhu, J., Hau, J., Petit, I., Fortin, D., Cunanne, S., & Descoteaux, M. (2018). Recognition of white matter bundles using local and global streamline-based registration and clustering. *NeuroImage*, 170, 283–295. <https://doi.org/10.1016/j.neuroimage.2017.07.015>
- Geeraert, B. L., Lebel, R. M., & Lebel, C. (2019). A multiparametric analysis of white matter maturation during late childhood and adolescence. *Human Brain Mapping*, 40(15), 4345–4356. <https://doi.org/10.1002/hbm.24706>
- Girard, G., Whittingstall, K., Deriche, R., & Descoteaux, M. (2014). Towards quantitative connectivity analysis: Reducing tractography biases. *NeuroImage*, 98, 266–278. <https://doi.org/10.1016/j.neuroimage.2014.04.074>
- Girard, O. M., Prevost, V. H., Varma, G., Cozzone, P. J., Alsop, D. C., & Duhamel, G. (2015). Magnetization transfer from inhomogeneously broadened lines (ihMT): Experimental optimization of saturation parameters for human brain imaging at 1.5 Tesla. *Magnetic Resonance in Medicine*, 73(6), 2111–2121. <https://doi.org/10.1002/mrm.25330>
- Granberg, T., Fan, Q., Treaba, C. A., Ouellette, R., Herranz, E., Mangeat, G., Loupre, 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: A Journal of Neurology*, 140(11), 2912–2926. <https://doi.org/10.1093/brain/awx247>
- Granziera, C., Wuerfel, J., Barkhof, F., Calabrese, M., De Stefano, N., Enzinger, C., Evangelou, N., Filippi, M., Geurts, J. J. G., Reich, D. S., Rocca, M. A., Ropele, S., Rovira, A., Sati, P., Toosy, A. T., Vrenken, H., Gandini Wheeler-Kingshott, C. A. M., Kappos, L., & the MAGNIMS Study Group. (2021). Quantitative magnetic resonance imaging towards clinical application in multiple sclerosis. *Brain*, 144(5), 1296–1311. <https://doi.org/10.1093/brain/awab029>

- Grech-Sollars, M., Hales, P. W., Miyazaki, K., Raschke, F., Rodriguez, D., Wilson, M., Gill, S. K., Banks, T., Saunders, D. E., Clayden, J. D., Gwilliam, M. N., Barrick, T. R., Morgan, P. S., Davies, N. P., Rossiter, J., Auer, D. P., Grundy, R., Leach, M. O., Howe, F. A., ... Clark, C. A. (2015). Multi-centre reproducibility of diffusion MRI parameters for clinical sequences in the brain. *NMR in Biomedicine*, 28(4), 468–485. <https://doi.org/10.1002/nbm.3269>
- Gurevich, M., Waknin, R., Stone, E., & Achiron, A. (2018). Fingolimod-improved axonal and myelin integrity of white matter tracts associated with multiple sclerosis-related functional impairments. *CNS Neuroscience & Therapeutics*, 24(5), 412–419. <https://doi.org/10.1111/cns.12796>
- Hakulinen, U., Brander, A., Ilvesmäki, T., Helminen, M., Öhman, J., Luoto, T. M., & Eskola, H. (2021). Reliability of the freehand region-of-interest method in quantitative cerebral diffusion tensor imaging. *BMC Medical Imaging*, 21(1), 144. <https://doi.org/10.1186/s12880-021-00663-8>
- Helms, G., Dathe, H., Kallenberg, K., & Dechent, P. (2008). High-resolution maps of magnetization transfer with inherent correction for RF inhomogeneity and T1 relaxation obtained from 3D FLASH MRI. *Magnetic Resonance in Medicine*, 60(6), 1396–1407. <https://doi.org/10.1002/mrm.21732>
- Henkelman, R. M., Stanisz, G. J., & Graham, S. J. (2001). Magnetization transfer in MRI: A review. *NMR in Biomedicine*, 14(2), 57–64. <https://doi.org/10.1002/nbm.683>
- Hickman, S. J., Toosy, A. T., Jones, S. J., Altmann, D. R., Miszkil, K. A., MacManus, D. G., Barker, G. J., Plant, G. T., Thompson, A. J., & Miller, D. H. (2004). Serial magnetization transfer imaging in acute optic neuritis. *Brain*, 127(3), 692–700. <https://doi.org/10.1093/brain/awh076>
- Honnadevasthan, Arun, A., Connelly, A., Smith, R. E., & Calamante, F. (2021). Characterisation of white matter asymmetries in the healthy human brain using diffusion MRI fixel-based analysis. *NeuroImage*, 225, 117505. <https://doi.org/10.1016/j.neuroimage.2020.117505>
- Irfanoglu, M. O., Sadeghi, N., Sarlis, J., & Pierpaoli, C. (2021). Improved reproducibility of diffusion MRI of the human brain with a four-way blip-up and down phase-encoding acquisition approach. *Magnetic Resonance in Medicine*, 85(5), 2696–2708. <https://doi.org/10.1002/mrm.28624>
- Jeurissen, B., Leemans, A., Tournier, J.-D., Jones, D. K., & Sijbers, J. (2013). Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging. *Human Brain Mapping*, 34(11), 2747–2766. <https://doi.org/10.1002/hbm.22099>
- Jones, D. K., Horsfield, M. A., & Simmons, A. (1999). Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magnetic Resonance in Medicine*, 42(3), 515–525. [https://doi.org/10.1002/\(SICI\)1522-2594\(199909\)42:3<515::AID-MRM14>3.0.CO;2-Q](https://doi.org/10.1002/(SICI)1522-2594(199909)42:3<515::AID-MRM14>3.0.CO;2-Q)
- Jones, D. K., Knösche, T. R., & Turner, R. (2013). White matter integrity, fiber count, and other fallacies: The do's and don'ts of diffusion MRI. *NeuroImage*, 73, 239–254. <https://doi.org/10.1016/j.neuroimage.2012.06.081>
- Kim, M., & Cercignani, M. (2014). Chapter 3.4—Magnetization transfer. In J. Cohen-Adad & C. A. M. Wheeler-Kingshott (Eds.), *Quantitative MRI of the spinal cord* (pp. 164–180). Academic Press. <https://doi.org/10.1016/B978-0-12-396973-6.00012-5>
- Kim, T., Aizenstein, H. J., Snitz, B. E., Cheng, Y., Chang, Y.-F., Roush, R. E., Huppert, T. J., Cohen, A., Doman, J., & Becker, J. T. (2022). Tract specific white matter lesion load affects white matter microstructure and their relationships with functional connectivity and cognitive decline. *Frontiers in Aging Neuroscience*, 13, 760663. <https://doi.org/10.3389/fnagi.2021.760663>
- Koller, K., Rudrapatna, S. U., Chamberland, M., Raven, E. P., Parker, G. D., Tax, C. M. W., Drakesmith, M., Fasan, F., Owen, D., Hughes, G., Charron, C., Evans, J. C., & Jones, D. K. (2020). MICRA: Microstructural image compilation with repeated acquisitions. *NeuroImage*, 225, 117406. <https://doi.org/10.1016/j.neuroimage.2020.117406>
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intra-class correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Koshiyama, D., Fukunaga, M., Okada, N., Morita, K., Nemoto, K., Yamashita, F., Yamamori, H., Yasuda, Y., Matsumoto, J., Fujimoto, M., Kudo, N., Azechi, H., Watanabe, Y., Kasai, K., & Hashimoto, R. (2020). Association between the superior longitudinal fasciculus and perceptual organization and working memory: A diffusion tensor imaging study. *Neuroscience Letters*, 738, 135349. <https://doi.org/10.1016/j.neulet.2020.135349>
- Kraguljac, N. V., Anthony, T., Monroe, W. S., Skidmore, F. M., Morgan, C. J., White, D. M., Patel, N., & Lahti, A. C. (2019). A longitudinal neurite and free water imaging study in patients with a schizophrenia spectrum disorder. *Neuropsychopharmacology*, 44(11), 1932–1939. <https://doi.org/10.1038/s41386-019-0427-3>
- Laule, C., & Moore, G. R. W. (2018). Myelin water imaging to detect demyelination and remyelination and its validation in pathology. *Brain Pathology*, 28(5), 750–764. <https://doi.org/10.1111/bpa.12645>
- Lee, S.-H., Coutu, J.-P., Wilkens, P., Yendiki, A., Rosas, H. D., & Salat, D. H. (2015). Tract-based analysis of white matter degeneration in Alzheimer's disease. *Neuroscience*, 301, 79–89. <https://doi.org/10.1016/j.neuroscience.2015.05.049>
- Lehmann, N., Aye, N., Kaufmann, J., Heinze, H.-J., Düzel, E., Ziegler, G., & Taubert, M. (2021). Longitudinal reproducibility of neurite orientation dispersion and density imaging (NODDI) derived metrics in the White matter. *Neuroscience*, 457, 165–185. <https://doi.org/10.1016/j.neuroscience.2021.01.005>
- Lema, A., Bishop, C., Malik, O., Mattoscio, M., Ali, R., Nicholas, R., Muraro, P. A., Matthews, P. M., Waldman, A. D., & Newbould, R. D. (2017). A comparison of magnetization transfer methods to assess brain and cervical cord microstructure in multiple sclerosis. *Journal of Neuroimaging*, 27(2), 221–226. <https://doi.org/10.1111/jon.12377>
- Lu, P.-J., Barakovic, M., Weigel, M., Rahmanzadeh, R., Galbusera, R., Schiavi, S., Daducci, A., La Rosa, F., Bach Cuadra, M., Sandkühler, R., Kuhle, J., Kappos, L., Cattin, P., & Granziera, C. (2021). GAMER-MRI in multiple sclerosis identifies the diffusion-based microstructural measures that are most sensitive to focal damage: A deep-learning-based analysis and clinico-biological validation. *Frontiers in Neuroscience*, 15, 647535. <https://doi.org/10.3389/fnins.2021.647535>
- Lucignani, M., Breschi, L., Espagnet, M. C. R., Longo, D., Talamanca, L. F., Placidi, E., & Napolitano, A. (2021). Reliability on multiband diffusion NODDI models: A test retest study on children and adults. *NeuroImage*, 238, 118234. <https://doi.org/10.1016/j.neuroimage.2021.118234>
- Luque Laguna, P. A., Combes, A. J. E., Streffer, J., Einstein, S., Timmers, M., Williams, S. C. R., & Dell'Acqua, F. (2020). Reproducibility, reliability and variability of FA and MD in the older healthy population: A test-retest multiparametric analysis. *NeuroImage: Clinical*, 26, 102168. <https://doi.org/10.1016/j.nicl.2020.102168>
- Magnotta, V. A., Matsui, J. T., Liu, D., Johnson, H. J., Long, J. D., Bolster, B. D., Mueller, B. A., Lim, K., Mori, S., Helmer, K. G., Turner, J. A., Reading, S., Lowe, M. J., Aylward, E., Flashman, L. A., Bonett, G., & Paulsen, J. S. (2012). MultiCenter reliability of diffusion tensor imaging. *Brain Connectivity*, 2(6), 345–355. <https://doi.org/10.1089/brain.2012.0112>
- Maier-Hein, K. H., Neher, P. F., Houde, J.-C., Côté, M.-A., Garyfallidis, E., Zhong, J., Chamberland, M., Yeh, F.-C., Lin, Y.-C., Ji, Q., Reddick, W. E., Glass, J. O., Chen, D. Q., Feng, Y., Gao, C., Wu, Y., Ma, J., Renjie, H., Li, Q., ... Descoteaux, M. (2017). The challenge of mapping the human connectome based on diffusion tractography. *Nature Communications*, 8(1), 1349. <https://doi.org/10.1038/s41467-017-01285-x>

- Manning, A. P., Chang, K. L., MacKay, A. L., & Michal, C. A. (2017). The physical mechanism of "inhomogeneous" magnetization transfer MRI. *Journal of Magnetic Resonance*, 274, 125–136. <https://doi.org/10.1016/j.jmr.2016.11.013>
- Mchinda, S., Varma, G., Prevost, V. H., Le Troter, A., Rapacchi, S., Guye, M., Pelletier, J., Ranjeva, J.-P., Alsop, D. C., Duhamel, G., & Girard, O. M. (2018). Whole brain inhomogeneous magnetization transfer (ihMT) imaging: Sensitivity enhancement within a steady-state gradient echo sequence. *Magnetic Resonance in Medicine*, 79(5), 2607–2619. <https://doi.org/10.1002/mrm.26907>
- Mito, R., Raffelt, D., Dhollander, T., Vaughan, D. N., Tournier, J.-D., Salvado, O., Brodtmann, A., Rowe, C. C., Villemagne, V. L., & Connelly, A. (2018). Fibre-specific white matter reductions in Alzheimer's disease and mild cognitive impairment. *Brain*, 141(3), 888–902. <https://doi.org/10.1093/brain/awx355>
- Munsch, F., Varma, G., Taso, M., Girard, O., Guidon, A., Duhamel, G., & Alsop, D. C. (2021). Characterization of the cortical myeloarchitecture with inhomogeneous magnetization transfer imaging (ihMT). *NeuroImage*, 225, 117442. <https://doi.org/10.1016/j.neuroimage.2020.117442>
- Obberghen, E. V., Mchinda, S., le Troter, A., Prevost, V. H., Viout, P., Guye, M., Varma, G., Alsop, D. C., Ranjeva, J.-P., Pelletier, J., Girard, O., & Duhamel, G. (2018). Evaluation of the sensitivity of inhomogeneous magnetization transfer (ihMT) MRI for multiple sclerosis. *American Journal of Neuroradiology*, 39, 634–641. <https://doi.org/10.3174/ajnr.A5563>
- Palacios, E. M., Martin, A. J., Boss, M. A., Ezekiel, F., Chang, Y. S., Yuh, E. L., Vassar, M. J., Schnyer, D. M., MacDonald, C. L., Crawford, K. L., Irimia, A., Toga, A. W., Mukherjee, P., & TRACK-TBI Investigators. (2017). Toward precision and reproducibility of diffusion tensor imaging: A multicenter diffusion phantom and traveling volunteer study. *American Journal of Neuroradiology*, 38(3), 537–545. <https://doi.org/10.3174/ajnr.A5025>
- 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. *Magnetic Resonance Imaging*, 50, 96–109. <https://doi.org/10.1016/j.mri.2018.03.004>
- Pierpaoli, C., & Basser, P. J. (1996). Toward a quantitative assessment of diffusion anisotropy. *Magnetic Resonance in Medicine*, 36(6), 893–906.
- Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafò, M. R., Nichols, T. E., Poline, J.-B., Vul, E., & Yarkoni, T. (2017). Scanning the horizon: Towards transparent and reproducible neuroimaging research. *Nature Reviews Neuroscience*, 18(2), 115–126. <https://doi.org/10.1038/nrn.2016.167>
- Prevost, V. H., Cayre, M., Carvalho, V. N., Mchinda, S., Varma, G., Ranjeva, J. P., & Duhamel, G. (2018). Inhomogeneous Magnetization Transfer (ihMT) Sensitivity to Myelin Impairments in Cuprizone Mouse Model 926. <https://cds.ismrm.org/protected/18MPProceedings/PDFfiles/0926.html>
- Raffelt, D., Tournier, J.-D., Rose, S., Ridgway, G. R., Henderson, R., Crozier, S., Salvado, O., & Connelly, A. (2012). Apparent fibre density: A novel measure for the analysis of diffusion-weighted magnetic resonance images. *NeuroImage*, 59(4), 3976–3994. <https://doi.org/10.1016/j.neuroimage.2011.10.045>
- Raffelt, D. A., Tournier, J.-D., Smith, R. E., Vaughan, D. N., Jackson, G., Ridgway, G. R., & Connelly, A. (2017). Investigating white matter fibre density and morphology using fixel-based analysis. *NeuroImage*, 144, 58–73. <https://doi.org/10.1016/j.neuroimage.2016.09.029>
- Rahmanzadeh, R., Galbusera, R., Lu, P.-J., Bahn, E., Weigel, M., Barakovic, M., Franz, J., Nguyen, T. D., Spincemaille, P., Schiavi, S., Daducci, A., La Rosa, F., Absinta, M., Sati, P., Bach Cuadra, M., Radue, E.-W., Leppert, D., Kuhle, J., Kappos, L., ... Granziera, C. (2022). A new advanced MRI biomarker for Remyelinated lesions in multiple sclerosis. *Annals of Neurology*, 92(3), 486–502. <https://doi.org/10.1002/ana.26441>
- Rasoanandrianina, H., Demortière, S., Trabelsi, A., Ranjeva, J. P., Girard, O., Duhamel, G., Guye, M., Pelletier, J., Audoin, B., & Callot, V. (2020). Sensitivity of the inhomogeneous magnetization transfer imaging technique to spinal cord damage in multiple sclerosis. *American Journal of Neuroradiology*, 41(5), 929–937. <https://doi.org/10.3174/ajnr.A6554>
- Rheault, F. (2020). *Analyse et reconstruction de faisceaux de la matière blanche* [Université de Sherbrooke]. <https://savoirs.usherbrooke.ca/handle/11143/17255>
- Rheault, F., Benedictis, A. D., Daducci, A., Maffei, C., Tax, C. M. W., Romascano, D., Caverzasi, E., Morency, F. C., Corrivetti, F., Pestilli, F., Girard, G., Theaud, G., Zemmoura, I., Hau, J., Glavin, K., Jordan, K. M., Pomiecko, K., Chamberland, M., Barakovic, M., ... Descoteaux, M. (2020). Tractostorm: The what, why, and how of tractography dissection reproducibility. *Human Brain Mapping*, 41(7), 1859–1874. <https://doi.org/10.1002/hbm.24917>
- Rheault, F., Schilling, K. G., Valcourt-Caron, A., Théberge, A., Poirier, C., Grenier, G., Guberman, G. I., Begnoche, J., Legarreta, J. H., Cai, L. Y., Roy, M., Edde, M., Caceres, M. P., Ocampo-Pineda, M., Al-Sharif, N., Karan, P., Bontempi, P., Obaid, S., Bosticardo, S., ... Landman, B. A. (2022). Tractostorm 2: Optimizing tractography dissection reproducibility with segmentation protocol dissemination. *Human Brain Mapping*, 43, 2134–2147. <https://doi.org/10.1002/hbm.25777>
- Roy, M., Fortier, M., Rheault, F., Edde, M., Croteau, E., Castellano, C.-A., Langlois, F., St-Pierre, V., Cuenoud, B., Bocti, C., Fulop, T., Descoteaux, M., & Cunnane, S. C. (2021). A ketogenic supplement improves white matter energy supply and processing speed in mild cognitive impairment. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 7(1), e12217. <https://doi.org/10.1002/trc2.12217>
- Saccetti, L., Hagiwara, A., Andica, C., Yokoyama, K., Fujita, S., Kato, S., Maekawa, T., Kamagata, K., Le Berre, A., Hori, M., Wada, A., Tateishi, U., Hattori, N., & Aoki, S. (2020). Myelin measurement using quantitative magnetic resonance imaging: A correlation study comparing various imaging techniques in patients with multiple sclerosis. *Cell*, 9(2), 2. <https://doi.org/10.3390/cells9020393>
- Sarrazin, S., Poupon, C., Teillac, A., Mangin, J.-F., Polosan, M., Favre, P., Laidi, C., D'Albis, M.-A., Leboyer, M., Lledo, P.-M., Henry, C., & Houenou, J. (2019). Higher in vivo cortical intracellular volume fraction associated with lithium therapy in bipolar disorder: A multicenter NODDI study. *Psychotherapy and Psychosomatics*, 88(3), 171–176. <https://doi.org/10.1159/000498854>
- Schilling, K. G., Petit, L., Rheault, F., Remedios, S., Pierpaoli, C., Anderson, A. W., Landman, B. A., & Descoteaux, M. (2020). Brain connections derived from diffusion MRI tractography can be highly anatomically accurate-if we know where white matter pathways start, where they end, and where they do not go. *Brain Structure & Function*, 225(8), 2387–2402. <https://doi.org/10.1007/s00429-020-02129-z>
- Schilling, K. G., Rheault, F., Petit, L., Hansen, C. B., Nath, V., Yeh, F.-C., Girard, G., Barakovic, M., Rafael-Patino, J., Yu, T., Fischi-Gomez, E., Pizzolato, M., Ocampo-Pineda, M., Schiavi, S., Canales-Rodríguez, E. J., Daducci, A., Granziera, C., Innocenti, G., Thiran, J.-P., ... Descoteaux, M. (2021). Tractography dissection variability: What happens when 42 groups dissect 14 white matter bundles on the same dataset? *NeuroImage*, 243, 118502. <https://doi.org/10.1016/j.neuroimage.2021.118502>
- Schneider, T., Brownlee, W., Zhang, H., Ciccirelli, O., Miller, D. H., & Wheeler-Kingshott, C. G. (2017). Sensitivity of multi-shell NODDI to multiple sclerosis white matter changes: A pilot study. *Functional Neurology*, 32(2), 97–101. <https://doi.org/10.11138/FNeur/2017.32.2.097>

- Schwartz, D. L., Tagge, I., Powers, K., Ahn, S., Bakshi, R., Calabresi, P. A., Constable, R. T., Grinstead, J., Henry, R. G., Nair, G., Papinutto, N., Pelletier, D., Shinohara, R., Oh, J., Reich, D. S., Sicotte, N. L., & Rooney, W. D. (2019). Multisite reliability and repeatability of an advanced brain MRI protocol. *Journal of Magnetic Resonance Imaging*, 50(3), 878–888. <https://doi.org/10.1002/jmri.26652>
- Shahim, P., Holleran, L., Kim, J. H., & Brody, D. L. (2017). Test-retest reliability of high spatial resolution diffusion tensor and diffusion kurtosis imaging. *Scientific Reports*, 7(1), 11141. <https://doi.org/10.1038/s41598-017-11747-3>
- Shou, H., Eloyan, A., Lee, S., Zipunnikov, V., Crainiceanu, A. N., Nebel, M. B., Caffo, B., Lindquist, M. A., & Crainiceanu, C. M. (2013). Quantifying the reliability of image replication studies: The image intra-class correlation coefficient (I2C2). *Cognitive, Affective, & Behavioral Neuroscience*, 13(4), 714–724. <https://doi.org/10.3758/s13415-013-0196-0>
- Soustelle, L., Troalen, T., Herten, A., Mchinda, S., Ranjeva, J.-P., Guye, M., Varma, G., Alsop, D. C., Duhamel, G., & Girard, O. M. (2022). A strategy to reduce the sensitivity of inhomogeneous magnetization transfer (ihMT) imaging to radiofrequency transmit field variations at 3 T. *Magnetic Resonance in Medicine*, 87(3), 1346–1359. <https://doi.org/10.1002/mrm.29055>
- Tariq, M. (2013). *Assessing scan-rescan reproducibility of the parameter estimates from NODDI*. ISMRM. <https://cds.ismrm.org/protected/13MProceedings/PDFfiles/3187.PDF>
- Tariq, M. (2018). *In vivo quantification of complex neurite configurations using magnetic resonance imaging* [Doctoral thesis], University College London. <https://discovery.ucl.ac.uk/id/eprint/10046378/>
- Teipel, S. J., Reuter, S., Stieltjes, B., Acosta-Cabronero, J., Ernemann, U., Fellgiebel, A., Filippi, M., Frisoni, G., Hentschel, F., Jessen, F., Klöppel, S., Meindl, T., Pouwels, P. J. W., Hauenstein, K.-H., & Hampel, H. (2011). Multicenter stability of diffusion tensor imaging measures: A European clinical and physical phantom study. *Psychiatry Research: Neuroimaging*, 194(3), 363–371. <https://doi.org/10.1016/j.psychres.2011.05.012>
- Theaud, G., Houde, J.-C., Boré, A., Rheault, F., Morency, F., & Descoteaux, M. (2020). TractoFlow: A robust, efficient and reproducible diffusion MRI pipeline leveraging nextflow & singularity. *NeuroImage*, 218, 116889. <https://doi.org/10.1016/j.neuroimage.2020.116889>
- Thieleking, R., Zhang, R., Paerisch, M., Wirkner, K., Anwander, A., Beyer, F., Villringer, A., & Witte, A. V. (2021). Same brain, different look?—The impact of scanner, sequence and preprocessing on diffusion imaging outcome parameters. *Journal of Clinical Medicine*, 10(21), 4987. <https://doi.org/10.3390/jcm10214987>
- Tofts, P., Dowell, N. G., & Cercignani, M. (2018). *Quantitative MRI of the brain: Principles of physical measurement* (Second ed.). CRC Press Taylor & Francis Group.
- Tournier, J.-D., Calamante, F., & Connelly, A. (2007). Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage*, 35(4), 1459–1472. <https://doi.org/10.1016/j.neuroimage.2007.02.016>
- Tournier, J.-D., Calamante, F., Gadian, D. G., & Connelly, A. (2004). Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage*, 23(3), 1176–1185. <https://doi.org/10.1016/j.neuroimage.2004.07.037>
- Uddin, M. N., Figley, T. D., Solar, K. G., Shatil, A. S., & Figley, C. R. (2019). Comparisons between multi-component myelin water fraction, T1w/T2w ratio, and diffusion tensor imaging measures in healthy human brain structures. *Scientific Reports*, 9(1), 1. <https://doi.org/10.1038/s41598-019-39199-x>
- van der Weijden, C. W. J., García, D. V., Borra, R. J. H., Thurner, P., Meilof, J. F., van Laar, P.-J., Dierckx, R. A. J. O., Gutmann, I. W., & Vries, E. F. J. (2021). Myelin quantification with MRI: A systematic review of accuracy and reproducibility. *NeuroImage*, 226, 117561. <https://doi.org/10.1016/j.neuroimage.2020.117561>
- Varma, G., Duhamel, G., de Bazelaire, C., & Alsop, D. C. (2015). Magnetization transfer from inhomogeneously broadened lines: A potential marker for myelin. *Magnetic Resonance in Medicine*, 73(2), 614–622. <https://doi.org/10.1002/mrm.25174>
- Varma, G., Girard, O. M., Mchinda, S., Prevost, V. H., Grant, A. K., Duhamel, G., & Alsop, D. C. (2018). Low duty-cycle pulsed irradiation reduces magnetization transfer and increases the inhomogeneous magnetization transfer effect. *Journal of Magnetic Resonance*, 296, 60–71. <https://doi.org/10.1016/j.jmr.2018.08.004>
- Vavasour, I. M., Laule, C., Li, D. K. B., Traboulsee, A. L., & MacKay, A. L. (2011). Is the magnetization transfer ratio a marker for myelin in multiple sclerosis? *Journal of Magnetic Resonance Imaging*, 33(3), 710–718. <https://doi.org/10.1002/jmri.22441>
- Vavasour, I. M., Tam, R., Li, D. K., Laule, C., Taylor, C., Kolind, S. H., MacKay, A. L., Javed, A., & Traboulsee, A. (2019). A 24-month advanced magnetic resonance imaging study of multiple sclerosis patients treated with alemtuzumab. *Multiple Sclerosis Journal*, 25(6), 811–818. <https://doi.org/10.1177/1352458518770085>
- Veenith, T. V., Carter, E., Grossac, J., Newcombe, V. F. J., Outtrim, J. G., Lupson, V., Williams, G. B., Menon, D. K., & Coles, J. P. (2013). Inter subject variability and reproducibility of diffusion tensor imaging within and between different imaging sessions. *PLoS One*, 8(6), e65941. <https://doi.org/10.1371/journal.pone.0065941>
- Volz, L. J., Cieslak, M., & Grafton, S. T. (2018). A probabilistic atlas of fiber crossings for variability reduction of anisotropy measures. *Brain Structure and Function*, 223(2), 635–651. <https://doi.org/10.1007/s00429-017-1508-x>
- Wang, J. Y., Abdi, H., Bakhadirov, K., Diaz-Arrastia, R., & Devous, M. D. (2012). A comprehensive reliability assessment of quantitative diffusion tensor tractography. *NeuroImage*, 60(2), 1127–1138. <https://doi.org/10.1016/j.neuroimage.2011.12.062>
- Weiskopf, N., Suckling, J., Williams, G., Correia, M., Inkster, B., Tait, R., Ooi, C., Bullmore, E., & Lutti, A. (2013). Quantitative multi-parameter mapping of R1, PD*, MT, and R2* at 3T: A multi-center validation. *Frontiers in Neuroscience*, 7, 95. <https://doi.org/10.3389/fnins.2013.00095>
- Wolff, S. D., & Balaban, R. S. (1989). Magnetization transfer contrast (MTC) and tissue water proton relaxation in vivo. *Magnetic Resonance in Medicine*, 10(1), 135–144. <https://doi.org/10.1002/mrm.1910100113>
- Wood, T. C., Damestani, N. L., Lawrence, A. J., Ljungberg, E., Barker, G. J., Solana, A. B., Wiesinger, F., & Williams, S. C. R. (2020). Silent myelin-weighted magnetic resonance imaging. *Wellcome Open Research*, 5, 74. <https://doi.org/10.12688/wellcomeopenres.15845.2>
- Yeatman, J. D., Dougherty, R. F., Myall, N. J., Wandell, B. A., & Feldman, H. M. (2012). Tract profiles of white matter properties: Automating fiber-tract quantification. *PLoS One*, 7(11), e49790. <https://doi.org/10.1371/journal.pone.0049790>
- Yeh, F.-C., Badre, D., & Verstynen, T. (2016). Connectometry: A statistical approach harnessing the analytical potential of the local connectome. *NeuroImage*, 125, 162–171. <https://doi.org/10.1016/j.neuroimage.2015.10.053>
- Yendiki, A., Panneck, P., Srinivasan, P., Stevens, A., Zöllei, L., Augustinack, J., Wang, R., Salat, D., Ehrlich, S., Behrens, T., Jbabdi, S., Gollub, R., & Fischl, B. (2011). Automated probabilistic reconstruction of white-matter pathways in health and disease using an atlas of the underlying anatomy. *Frontiers in Neuroinformatics*, 5, 23. <https://doi.org/10.3389/fninf.2011.00023>
- 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(4), 1000–1016. <https://doi.org/10.1016/j.neuroimage.2012.03.072>

- Zhang, L., Chen, T., Tian, H., Xue, H., Ren, H., Li, L., Fan, Q., Wen, B., & Ren, Z. (2019). Reproducibility of inhomogeneous magnetization transfer (ihMT): A test-retest, multi-site study. *Magnetic Resonance Imaging*, 57, 243–249. <https://doi.org/10.1016/j.mri.2018.11.010>
- Zhang, L., Wen, B., Chen, T., Tian, H., Xue, H., Ren, H., Li, L., Fan, Q., & Ren, Z. (2020). A comparison study of inhomogeneous magnetization transfer (ihMT) and magnetization transfer (MT) in multiple sclerosis based on whole brain acquisition at 3.0 T. *Magnetic Resonance Imaging*, 70, 43–49. <https://doi.org/10.1016/j.mri.2020.03.010>
- Zhou, X., Sakaie, K. E., Debbins, J. P., Narayanan, S., Fox, R. J., & Lowe, M. J. (2018). Scan-rescan repeatability and cross-scanner comparability of DTI metrics in healthy subjects in the SPRINT-MS multicenter trial. *Magnetic Resonance Imaging*, 53, 105–111. <https://doi.org/10.1016/j.mri.2018.07.011>
- Zuo, X.-N., Xu, T., & Milham, M. P. (2019). Harnessing reliability for neuroscience research. *Nature Human Behaviour*, 3(8), 8–771. <https://doi.org/10.1038/s41562-019-0655-x>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Edde, M., Theaud, G., Dumont, M., Théberge, A., Valcourt-Caron, A., Gilbert, G., Houde, J.-C., Maltais, L., Rheault, F., Spagnolo, F., Barakovic, M., Magon, S., & Descoteaux, M. (2023). High-frequency longitudinal white matter diffusion- and myelin-based MRI database: Reliability and variability. *Human Brain Mapping*, 44(9), 3758–3780. <https://doi.org/10.1002/hbm.26310>