1

mrHARDIfLOW : A pipeline tailored for the preprocessing and analysis of Multi-Resolution High Angular diffusion MRI and its application to a variability study of the PRIME-DE database

Alex Valcourt Caron^{1,*}, Amir Shmuel², Ziqi Hao², Maxime Descoteaux¹

¹SCIL, Université de Sherbrooke, Computer Science Department, Sherbrooke, QC, Canada

²Brain Imaging Signals Lab, McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, QC, Canada

Correspondence*:

Alex Valcourt Caron

alex.valcourt.caron@usherbrooke.ca

2 ABSTRACT

3

4 The validation of advanced methods in diffusion MRI requires finer acquisition resolutions, which
5 is hard to acquire with decent Signal-to-Noise Ratio (SNR) in humans. The use of Non-Human
6 Primates (NHP) and anaesthesia is key to unlock valid microstructural maps, but tools must
7 be adapted and configured finely for them to work well. Here, we propose a novel processing
8 pipeline implemented in Nextflow, designed for robustness and scalability, in a modular fashion to
9 allow for maintainability and a high level of customization and parametrization, tailored for the
10 analysis of diffusion data acquired on multiple spatial resolutions. Modules of processes and
11 workflows were implemented upon cutting edge and state-of-the-art MRI processing technologies
12 and diffusion modelling algorithms, namely Diffusion Tensor Imaging (DTI), Constrained Spherical
13 Deconvolution (CSD) and Distribution of Anisotropic Microstructural Environments in Diffusion-
14 compartment imaging (DIAMOND), a multi-tensor distribution estimator. Using our pipeline, we
15 provide an in-depth study of the variability of diffusion models and measurements computed
16 on 32 subjects from 3 sites of the PRIME-DE, a database containing anatomical (T1, T2),
17 functional (fMRI) and diffusion (DWI) imaging of Non-Human Primate (NHP). Together, they offer
18 images acquired over a range of different spatial resolutions, using single-shell and multi-shell
19 b-value gradient samplings, on multiple scanner vendors, that present artifacts at different level of
20 importance. We also perform a reproducibility study of DTI, CSD and DIAMOND measurements
21 outputted by the pipeline, using the Aix-Marseilles site, to ensure our implementation has minimal
22 impact on their variability. We observe very high reproducibility from a majority of diffusion
23 measurements, only gamma distribution parameters computed on the DIAMOND model display
24 a less reproducible behaviour. This should be taken into consideration when future applications
25 are performed. We also show that even if promising, the PRIME-DE diffusion data exhibits a
26 great level of variability and its usage should be done with care to prevent instilling uncertainty in
27 statistical analyses.

28 **Keywords:** Non-Human Primates, Magnetic Resonance, Diffusion, Medical Imaging, Pipeline, Automatization, High Parallel Computing,
29 Variability

1 INTRODUCTION

30 Magnetic resonance imaging is a modality of medical imaging that has seen an exponential increase of
31 interest in the past few decades since its the only non-invasive technology able to acquire both functional and
32 structural information from soft tissues. That makes it the preferential choice for the study of brain functions
33 and connectivity with the aim, notably, of mapping the human connectome. However, the study of the
34 generated images is riddled with challenges that the scientific community has yet to overcome [O'Donnell
35 and Pasternak (2015)][Jones et al. (2013)][Jones (2010)], especially when it comes to tractography
36 [Schilling et al. (2021b)][Rheault et al. (2020)][Calamante (2019)][Maier-Hein et al. (2017)][Schilling
37 et al. (2017)], even more in ensemble studies using multiple sites and multiple scanner vendors [Schilling
38 et al. (2021a)][Andica et al. (2020)][De Santis et al. (2019)]. This is particularly true for dMRI, for which
39 the current equipment available still fails to provide the intense and fast changing magnetic fields required
40 for good SNR; quality of the produced images is restricted by eddy currents and phase related distortions,
41 to name a few. Still, all MRI related techniques are limited by resolution orders of magnitude above
42 the size of the tissues of interest, resulting in partial volume effects at the interface between mediums.
43 Lowering the size of acquired voxels requires a longer acquisition time, and is often not a viable possibility,
44 since it leads to increased thermal noises in the reception coils and a higher probability of motion-related
45 artifacts in the resulting images. Selecting a good set of acquisition parameters is thus a multi-objective
46 optimization problem, a trade-off between the image resolution required for subsequent modelling and the
47 acceptable level of noise, the duration of the imaging sequence and the type of subject, a research setting
48 versus a clinical application, amongst others. Up to this day, most of it is done by trial and error, based
49 on recommendation from previous experiments, due to the lack of gold standards defining the validity of
50 most empirical MRI measurements. Inasmuch, the absence of standardized processing pipelines, carefully
51 designed to operate universally or given images acquired with a good subset of parameters, further increases
52 the difficulty of generating good research material exempt of biases from the equipment, subject or even
53 processing train.

54 This validation problem, affecting the field from the MRI scanner parametrization all the way to
55 modelization and quantity measurement, has been extensively discussed and a plethora of techniques
56 have been devised to tackle it. Comparing MRI acquisitions to tracers studies are known to provide good
57 benchmarks in white matter, but can lack in specificity [Heilingoetter and Jensen (2016)] and its application
58 requires the injection of an invasive substance in the tissue of interest thus being only performed ex-vivo.
59 Quite recently, multi-resolution frameworks have gained in popularity, combining low-resolution dMRI to
60 high-resolution microscopy in order to refine the angular representation of diffusion [Howard et al. (2019)].
61 However, this technique is also limited to ex-vivo studies, since microscopy requires slicing and staining of
62 the brain tissues.

63 Another good technique that has made its marks is comparison to results in small animals, since they
64 enable the use of specialized equipment that operates at higher magnetic fields and of stronger techniques
65 to mediate with subject motion and cardiac and respiratory gating, which are either inaccessible or ethically
66 impossible in human studies. Those allow to go to finer resolutions, both spatial and angular, while
67 retaining good SNR. Nonetheless, images still require to be adequately preprocessed before fitting models
68 and producing measurements. The algorithms used to that extent must be finely tuned to their specific
69 requirements. While multiple processing pipelines have been designed by the research community [Cieslak

et al. (2021)][Autio et al. (2020)][Theaud et al. (2020)][Gorgolewski et al. (2011)][Pierpaoli et al. (2010)], integrating MRI processing technologies into push-button applications, a great deal of them are configured to handle human subjects images or images acquired using a predefined set of parameters specific to a study, and running them on non-human high-resolution diffusion imaging requires extensive re-configuration. Most of them are also offered as black-boxes, difficult to customize or extend, which makes adapting them to different studies time consuming.

The pipeline we consider in this work alleviates most, if not all of those problems. Implemented in Nextflow [Tommaso et al. (2017)], with the latest implementation of their DSL2 framework, we designed a collection of processing modules consisting of pipeline processes and workflows, each targeting a specific set of preprocessing, model reconstruction, quantity measurement or utility algorithms, using cutting edge DWI processing technologies, such as Dipy [Garyfallidis et al. (2014)], FSL [Smith et al. (2004)], Mrtrix [Tournier et al. (2019)] and ANTs [Avants et al. (2008)]. We then built a complete processing pipeline using those modules, tailored to handle multi spatial resolution and high angular resolution diffusion encoding. The use of Nextflow ensures its scalability on a wide range of computing infrastructures, from an execution on a local computer to multiple nodes in an HCP facility, and container technologies (Docker and Singularity [Kurtzer et al. (2017)]) allow for the encapsulation of all dependencies and their versioning. This combination allow for a robust and automated execution, while keeping installation and maintenance as simple as possible for the end-user.

To test the pipeline, we processed 32 datasets from 3 sites of the PRIME-DE [Neff (2019)][Milham et al. (2018)] database. Those contain diffusion weighted images with multiple spatial and angular resolutions, single-shell and multi-shell b-value samplings, multiple scanner vendors and artifacts at different ranges of intensities depending on the site where the acquisition was executed. We provide a brief reproducibility study of the pipeline results, as well as an in-depth variability study of diffusion modelling and measurements obtained from its execution.

2 METHOD

2.1 Processing library

3 core concepts were taken into account for the development of the processing library for this project :

1. Efficiency
High resolution images are heavy (hundreds of megabytes to a few gigabytes), in particular for diffusion MRI, where a complete MRI volume is acquired for each gradient direction. Combining high spatial and orientational resolution leads to image files weighing multiple gigabytes and a number of voxels of the order of several hundreds of millions. Processing must thus be efficient and parallelized as much as possible. To achieve such a degree of efficiency, we coded the pipeline structure using Nextflow [Tommaso et al. (2017)], a meta-scheduling language enabling automatic optimization and execution of a processing tree over multiple datasets and multiple processing nodes.
2. Modularity and Morphability
State-of-the-art methods in diffusion MRI are still changing rapidly. For a good pipeline to stay up to date, it requires being able to morph to new standards and integrate novel techniques being discovered. Using the DSL2 framework of Nextflow [Tommaso et al. (2017)], we produced a highly modular library of processes and workflows, used as building blocks in our pipeline. This allowed for a high level architecture in the final pipeline coded for the current study, with a finer description of the dataflows encapsulated in the modules referring to specific processing steps. Using this paradigm

makes it easier for new pipelines to be developed that only change a subset of processing steps or add new workflows on top of older ones.

3. Usability and Reproducibility

The libraries used for diffusion MRI are ever-changing. It is thus impossible to guarantee a robust execution of our pipeline on versions of its dependencies released in the future. Docker and Singularity [Kurtzer et al. (2017)] - two technologies enabling the encapsulation of libraries in portable packages - were used to wrap up the numerous applications called in the different processing steps of the pipeline. They enable locking the versions of the dependencies into a pre-packaged image and remove the need for the end-user to install and manage them. In addition, they allow for multi-stage building of images, which makes it easy to add new dependencies and modify a set of versions in the case of an update of the underlying pipeline or for the needs of a specific study.

To address these concepts, the library was fragmented into 4 nested and interlocked scopes : pipeline scope, input scope, process scope and workflow scope, as depicted in figure 1. Of them, only the process scope contains the actual calls to algorithms and their input requirements. It encapsulates the actual workhorse of the pipeline and forms the processes module.

All other scopes are using Nextflow workflow objects and only specify the flow of data between processes and other workflows they include. The input scope gathers pieces of code necessary to translate the internal representation in disk memory of the input data into a data structure that is digestible by a pipeline. When included, it enforces the input convention of its pipeline. It can be changed to fit a specific study or project. The pipeline scope contains all complete pipelines. A valid pipeline uses one or more input convention from the input scope and describes as minimally as possible the flow of data between processes and sub-workflows it includes. All data flow descriptions that can potentially be reused are transferred to the workflow scope. This scope contains sub-parts of the processing that form a coherent ensemble, yet cannot be specified by a single process. The workflows in this scope uses processes and other workflows as building blocks and define the passage of data between them to form a processing sequence. An example of such workflow would be the execution of Topup (a light purple workflow in figure 1), including data preparation, which executes the following processes sequentially : selection of b0 volumes from the different phase encoded DWI, concatenation of b0 volumes, computation of Topup distortion field, application of the field to the DWI.

2.2 Processing pipeline

2.2.1 Data input

In addition to the DWI volume and its associated bval/bvec files describing the gradient sampling, the pipeline requires as input a T1 anatomical image. The user can also supply masks in either or both diffusion and anatomical space, as well as partial volume maps of white matter (WM), grey matter (GM) and cerebrospinal fluid (CSF), in anatomical space. Reverse phase acquired diffusion images can be provided, as either a single b0 volume, multiple b0 organized as a 4D volume or a full DWI 4D volume with bval/bvec files. For some of the processing steps, the pipeline also requires the input of some metadata associated to the acquisition of the DWI volumes. Metadata parameters consistent across all images can be supplied using the Nextflow configuration file. In case of varying parameters between the images, a json file must be provided alongside both forward and reverse acquired DWI volumes.

2.2.2 Processing steps

The following section presents all currently implemented steps in our pipeline, which are also available in our processes and workflows library. Even though we highly recommend using them all in the order

Figure 1. 4 scopes composing the Nextflow library and how they interface with each others. A pipeline definition (Pipeline X for example) is placed at the root of the library. To load the data, it uses one or more convention from the input scope (convention 1 and the bids convention). Then, to process it, it calls utilities in either or both the process scope (ANTs registration, Concatenate images and others) and the workflow scope (execute Topup, then register T1 to DWI). A call to an utility in the workflow scope will trigger calls to processes in the process scope they include to process the data (execute Topup firsts calls a concatenation of two opposite phase encoding DWI and extracts their b0 volumes, to pass the result to Topup for phase distortions correction). The input, workflow and process scopes form the core of the Nextflow library. Their content can be used by any pipeline defined in the pipeline scope, allowing high levels of code reusability.

154 prescribed to process raw anatomical and diffusion weighted images, their execution is by default optional
 155 and can be individually turned on and off using the Nextflow configuration file, for instance to allow the
 156 computation of the models and measures on already preprocessed data or to shorten the execution time by
 157 turning off steps deemed unnecessary after input data quality control.

158 2.2.2.1 Preprocessing

159 2.2.2.1.1 Diffusion processes

160 **Background noise** - dMRI images are affected with background noise following either a rician
 161 [Gudbjartsson and Patz (1995)] or a chi-squared distribution [Luisier et al. (2012)] and more often
 162 than not present quite poor signal-to-noise ratio. Some patch based algorithms [Manjón et al. (2008)]
 163 enable correction of such types of noise, but require long execution time and vast amounts of memory
 164 space, which can be cumbersome given the large size of diffusion volumes. A good trade-off is to use a
 165 Principal Component Analysis to extract new dimensions isolating the different sources of noise from the
 166 data. This technique has proven its abilities at enhancing the quality of diffusion data, even if it does not
 167 take into account the true nature of the noise and only corrects for normally distributed distortions. We thus
 168 run dwidenoise [Cordero-Grande et al. (2019)][Veraart et al. (2016b)][Veraart et al. (2016a)], a MP-PCA

Figure 2. DWI preprocessing steps. (Background denoising) Input 4D DWI from both phase encoding directions are denoised to remove background noise contamination. (b0 to b0 normalization) The signal is normalized so all b0 volumes in both phase encoding directions exhibit the same mean value. Each 3D diffusion volume is normalized using as reference either the b0 volume coming before it, after it or a linear interpolation of both. (Topup) b0 volumes from both phase encoding directions are concatenated and a deformation field is estimated. (Apply Topup) This field is applied on the forward phase encoded DWI to produce an undistorted 4D DWI. (Extract b0 + BET) From the undistorted 4D DWI is extracted a mean b0 volume, upon which is computed a brain mask. All voxels outside the brain mask are set to 0 in the mean b0 volume and the undistorted 4D DWI. (DTI FA) The masked undistorted 4D DWI is used to compute a DTI model and obtain a FA volume. (Register T1 to DWI) A transformation is computed by registering the T1 volume to both the mean b0 and the FA volumes. (Apply Transform to Mask) The T1 mask is transformed to the DWI space using the computed transformation. (Eddy) Motion and eddy currents correction is computed using the deformation field from Topup, the T1 mask in DWI space and the 4D DWI from both phase encoding directions. (N4) The 4D DWI is corrected for intensity distortions.

169 denoising algorithm available in Mrtrix3 [Tournier et al. (2019)], at the beginning of the preprocessing tree
 170 for diffusion, on forward and reverse phase acquired diffusion volumes, and Non-Local means [Manjón
 171 et al. (2008)] denoising on reverse phase acquired b0 volumes. Novel techniques using self-supervised
 172 learning such as Patch2Self [Fadnavis et al. (2020)] and deep convolutional neural networks [Kawamura
 173 et al. (2021)] could also easily be integrated in our modular framework, but have been discarded for the
 174 current implementation, since they still require to be thoroughly validated.

175 **b0 to b0 normalization** - High spatial and angular resolution diffusion images require long acquisition
 176 times. This causes heating in the acquisition coils, translated into a drift of signal intensities over time
 177 [Vos et al. (2017)]. To compensate for this effect, each pair of forward and reverse acquired images are
 178 normalized to the mean value of the first group of b0 volumes found in the forward image. To better correct
 179 the diffusion weighted volumes, a linear combination of the means of the groups of b0 volumes found

before and after it is used. This behavior can be modified to accommodate datasets with different deviations between b0 volumes and diffusion images, such as taking only the b0 before or after a group of diffusion directions to compensate for the signal drift.

Gibbs ringing - Since the spatial MRI images are reconstructed by inverse Fourier Transform from a subset only of points sampled through k-space, images can suffer from wave-like artifacts at the interface between different tissues [Czervionke et al. (1988)]. When uncorrected, this leads to biases in models and metrics. We thus included Gibbs ringing correction as a step just after background denoising, using mrdegibbs [Kellner et al. (2016)] from Mrtrix3 [Tournier et al. (2019)]. However, since most images are acquired using partial fourier and that this algorithm does not support it, this step is skipped by default.

Susceptibility - To correct for those artifacts, we use Topup [Andersson et al. (2003)] from the FSL [Smith et al. (2004)] library. Topup is run on the extracted b0 of forward and reverse phase acquisitions to correct for susceptibility induced distortions. The extraction of b0 can be configured to :

- Take the first b0 appearing in a dwi image
- Take all the b0 in the dwi image
- Take the averages of b0 in the dwi image
- Take averages of continuous series of b0 in the dwi image

Since Topup is not parallelized, the pipeline comes initially configured to extract the average b0 for this preprocessing step. For this study, this configuration has proven to provide a good approximation of the distortion field while lowering substantially the execution time of the algorithm. However, in the case of variations in the susceptibility field across the diffusion volumes, sampling more than one b0 per image is recommended.

Brain masking - Since most of the subsequent steps in the pipeline require a mask for better and/or faster computation, we compute one using FSL bet [Smith (2002)] on the mean b0 volume, using the one corrected by Topup if it was run. However, the execution of bet on monkey images has proven to be unstable and often leads to bleeding of the computed mask into skull and background regions. To circumvent this problem, we highly recommend to pass as input data either a mask aligned to the diffusion or computed on the T1 anatomical image, that has been quality checked and manually fixed. In the latter case, the b0 bet mask will be used only to help register the T1 to the mean b0 image + FA (via ANTs [Avants et al. (2008)] using a combination of rigid, affine and non-linear transformations) in order to transform the T1 mask to diffusion space.

Motion and Eddy currents - Eddy currents are naturally occurring in tissues imaged at the MRI due to the rapid variations of the imaging gradients [Ahn and Cho (1991)]. They cause local distortions in the magnetic field perceived by the protons spins, leading to spatial distortions in the acquired images. To correct for those, we use FSL Eddy [Andersson et al. (2018)][Andersson et al. (2017)][Andersson and Sotiropoulos (2016)][Andersson et al. (2016)], which also corrects for patient motion. This step is parallelized both on CPU (using openMP) and GPU (using CUDA on a Nvidia graphics card). When using the latter, execution is more efficient and one can also perform outlier detection and replacement, as well as slice-wise corrections.

Intensity distortions - MRI equipment - gradients and reception coils - are not perfectly accurate and images can suffer from variation in intensities that are not related to the diffusion process of the acquired tissues. This leads to distortions in the diffusion measures recovered from the models fitted to the signal. This effect can also be caused by the presence of high susceptibility materials near or inside the imaged

222 brain. To correct for those, we run N4 [Tustison et al. (2010)] intensity normalization from ANTs on the
223 mean b0 of each diffusion. We then apply the computed intensity transformation to all diffusion directions
224 acquired.

225 2.2.2.1.2 T1 processes

Figure 3. T1 preprocessing steps. (Background denoising) The T1 images is denoised to remove background noise contamination. (N4 normalization) The denoised T1 image is then corrected for intensity distortions.

226 **Background noise** - Since the T1 image is composed of a single 3D volume, Non-Local Means Denoising
227 [Manjón et al. (2008)] becomes tractable in a reasonable amount of time. It is run using a rician noise
228 model and parallelized over multiple CPU cores, via the algorithm available in Dipy [Garyfallidis et al.
229 (2014)]. This steps has proven to improve segmentation and registration [Theaud et al. (2020)][Constanzo
230 et al. (2018)] and has thus been integrated as an optional step in the pipeline.

231 **Intensity distortions** - As for diffusion, the T1 is corrected for local intensity distortions caused by the
232 MRI equipment using ANTs N4 [Tustison et al. (2010)] intensity normalization algorithm.

233 2.2.2.2 Upsampling

234 Once denoised, both T1 and the diffusion volumes are upsampled to a finer resolution defined in the
235 Nextflow configuration file using Dipy [Garyfallidis et al. (2014)]. This step enhances anatomical features
236 present in the diffusion volumes, brings both T1 and diffusion on a common spatial grid, allowing the use
237 of masks and segmentations subsequently computed on the T1 in reconstruction and measures algorithms
238 acting on the diffusion, and is also known to improve the results obtained from tractography [Dyrby et al.
239 (2014)].

240 2.2.2.3 Segmentation

241 **Tissue segmentation** - Segmentation is performed on the T1 using ANTs Atropos [Avants et al. (2011)].
242 The default configuration of the pipeline uses AFNI NMT v2.0 segmentation [Jung et al. (2021)] as priors,
243 after registering the NMT template to the T1 image, using ANTs [Avants et al. (2008)] (rigid + affine +
244 non-linear, MultiLabel interpolation). Tissue masks are also extracted from the partial volume fractions
245 outputted by Atropos by thresholding. A safe white matter mask, exempt of partial volume effects with both
246 CSF and gray matter, is computed for future usage such as limiting the number of possible voxels used to
247 compute the single fiber response necessary for constrained spherical deconvolution.

248 **White matter parcellation** - The default white matter atlas available in the pipeline is the UWDTI atlas
249 [Adluru et al. (2012)]. It provides templates in the form of DTI measures, of which we chose the FA to
250 perform registration. The FA is calculated from a DTI fit on the upsampled DWI data. Then a transformation

Figure 4. Tissues and white matter segmentation steps. (Register template to T1) A transformation from the template T1 to the preprocessed T1 is computed. (Apply transform) The transformation is used to move the segmentation prior into the preprocessed T1 space. (Atropos) Ants Atropos is run, with the transformed segmentation priors, to compute from the preprocessed T1 segmentation masks and partial volume fraction maps of white matter (WM), grey matter (GM) and cerebro-spinal fluid (CSF).

from template space to subject space is computed using ANTs [Avants et al. (2008)] (rigid + affine + non-linear) and applied to the atlas using MultiLabel interpolation.

2.2.2.4 Registration of T1 to DWI space

Once segmentation is computed, the upsampled T1 image is registered to the upsampled DWI space using a combination of the mean b0 volume and the FA obtained from a DTI fit. The registration is computed using ANTs and a sequence of Rigid, Affine and Non-Linear transformations. This transformation is then applied to all segmentation masks using Nearest Neighbour interpolation, to label maps resulting from tissue and white matter segmentation using MultiLabel interpolation and to partial volume maps from tissue segmentation with Linear interpolation.

2.2.2.5 Reconstruction

The pipeline offers 3 different models for the reconstruction of the diffusion process. To enable its usage on low angular resolution data or data sampled across few shells at low b-values, the single diffusion tensor model [LeBihan et al. (2001)] (DTI) can be fitted. It requires a minimum of 6 directions, though a uniform sampling of gradients over the sphere is preferable. To limit the effects of exchange and restriction occurring at high b-value, which this model cannot represent, only shells up to $b = 1300 \text{ mm/s}^2$ are used [Jones and Bassar (2004)]. This behavior can be modified by the user if need be. For data acquired on a greater number of shells, with higher b-values, such as HARDI [Hosey et al. (2005)] [Tuch et al. (2002)] and CUSP [Scherrer and Warfield (2012)], two higher-order models are available : the fiber orientation distribution function (fODF) through Constrained Spherical Deconvolution [Tournier et al. (2007)] (CSD) and the DIstribution of Anisotropic MicrOstructural eNvironments in Diffusion-compartment [Scherrer et al. (2016)] (DIAMOND), a multivariate gamma distribution over tensors.

2.2.2.6 Measurement

DTI fit - Diffusion tensors are computed using Dipy and a Weighted Least Square Fit. In addition to the tensors, eigenvalues and eigenvectors, it outputs axial (λ_1) and radial diffusivity ($\mu(\lambda_2, \lambda_3)$), fractional and geodesic anisotropy, the mode and norm of diffusion tensors. Figure 5 presents some of those metrics. It

Figure 5. DTI measurements on a selected subject from the Aix-Marseille site. (A) RGB map, (B) fractional anisotropy (FA), (C) mean diffusivity (MD), (D) radial diffusivity (RD), (E) axial diffusivity (AD) and (F) the principal direction of the estimated tensor in the centrum semi-ovale (region circled in blue in (A)).

also allows the output of validation maps, such as the residuals of the fit, the standard deviation across diffusion volumes displaying pulsation and misalignment artifacts, and a map of physically implausible voxels, representing the areas where the diffusion weighted signal presents higher values than the b_0 mean.

CSD fit - The response function and spherical deconvolution fit are also computed using Dipy. If a tissue segmentation is available at this step in the pipeline and the input DWI contains multiple b-value shells, a Multi-Shell Multi-Tissue approach is used [Jeurissen et al. (2014)]. Otherwise, the pipeline reverts to using a Single-Shell Single-Tissue algorithm [Tournier et al. (2007)]. The response is obtained using a subset of single-fiber white matter voxels (200 per default) which are presenting a high enough FA value (between 0.55 and 0.75 per default) and reside inside the safe white matter mask priorly computed. The peaks are then extracted from the fitted fODF, selecting only the ones presenting an amplitude greater than 1.5 times the maximal amplitude in isotropic voxels [Dell'Acqua et al. (2012)] (figure 6 : F). The latter are obtained using a threshold on mean diffusivity (a minimum of $2.6e^{-3}mm^2$) and fractional anisotropy (0.15). Those threshold are also used to compute the Number of Fibers Orientations and the maps of voxel-based total (first SH coefficient), sum (integral on the sphere) and maximal Apparent Fiber Density [Dell'Acqua et al. (2012)][Raffelt et al. (2012)], as well as the RGB mapping of orientations on the sphere (figure 6 : A-E).

Figure 6. fODF measurements on a selected subject from the Aix-Marseille site. (A) RGB map, (B) total apparent fiber density (AFDt), (D) number of fiber orientations (NuFO) and (C) fODF peaks in the centrum semi-ovale (region circled in blue in (A)).

291 **Multivariate Gamma Distribution Tensor Fit** - The Multi-Tensor fit is acquired using Diamond [Scherrer
292 et al. (2016)], configured to allow for a maximum of 2 tensors (fascicles) per voxel. As default, the fit also
293 outputs the fraction of free-water. Classical tensor measures (FA, AD, RD, MD, RGB) are estimated on
294 each fascicle separately and are used to compute their mean and max counterparts. In addition is generated
295 a map of the fraction of each fascicle per voxel and of their main peaks, and other statistical quantities such
296 as the mean and standard deviation of isotropic and anisotropic diffusivities [Reymbaut and Descoteaux
297 (2019)]. A subset of metrics is displayed in figure 7.

298 **2.3 Pipeline reproducibility**

299 To evaluate correctly the potential variations in metrics between the several subjects of the PRIME-DE
300 databases, it is a requirement that the processing carried on the data be as reproducible as possible. To do
301 so, all algorithms used by the pipeline that require the execution of a random process have seen their seed
302 enforced to the same number across all subjects and sites. To evaluate the reproducibility, we did a test-test
303 analysis, using the 4 subjects from the aix-marseille database, on which the pipeline was run 3 separate
304 times, using the same configuration and singularity image. The image intraclass correlation coefficient
305 (I2C2) [Shou et al. (2013)] was computed using equation 1 on relevant DTI, fODF and multi-tensor metrics.

Figure 7. Multi-tensor measurements on a selected subject from the Aix-Marseille site. (A) RGB map, (B) maximum fascicle FA (max fFA), (C) mean diffusivity (MD), (F) axial diffusivity (AD) on the left, radial diffusivity (RD) on the right, (E) free water fraction (fFW) and (D) principal direction of the estimated mean tensor of each fascicle in the centrum semi-ovale (region circled in blue in (A)).

$$I2C2 = 1 - \frac{\sum_{s \in S} N_s - 1}{\sum_{s \in S} (N_s - 1)} \sum_{s \in S} \sum_{ss \in s} \sum_{v \in I_{ss}} \frac{(I_{ss}(v) - \hat{I}_s(v))^2}{(I_{ss}(v) - \hat{I}(v))^2} \quad (1)$$

Here, I_{ss} refers to a modality, mask or measurement to evaluate for a specific session ss belonging to a subject s , $\hat{I}_s$ is the mean image for a specific subject s taken over all its sessions and $\hat{I}$ is the mean image taken over all sessions and all subjects.

2.4 Study design and data selection

This study focuses at evaluating variability in diffusion weighted (DW) measurements across the different sites available in the Primate Data Exchange (PRIME-DE) project, as well as between the MRI scanner vendors and models used by them to acquire the images. From the 25 sites available in the database, 8 of them were identified as containing diffusion data, and only 4 with a fair number of subjects ($N \geq 4$): Aix-Marseille University (4 subjects, Siemens Prisma scanner), University of California, Davis (19 subjects, Siemens Skyra scanner), Mount Sinai School of Medicine - Philips (9 subjects, Philips Achieva scanner) and Mount Sinai School of Medicine - Siemens (6 subjects, Siemens Skyra scanner). All subjects are Macaca Mulatta primates, anesthetized before acquisition. The scan sequences parameters can be found

SITE	SCANNER	# subjects	T1	DWI			
			voxel size	voxel size	b-values	# directions	# b0
Aix-Marseille	Siemens Prisma 3T	1	0.8mm 0.8mm 0.8mm	1.25mm 1.25mm 1.25mm	250 1000	6 64	5
		1	0.8mm 0.8mm 0.8mm	1mm 1mm 1mm	250 1000	6 64	5
		1	0.8mm 0.8mm 0.8mm	1.25mm 1.25mm 1.25mm	300 1000 2000	6 32 64	7
		1	0.8mm 0.8mm 0.8mm	1mm 1mm 1mm	500 1000 2000	6 30 30	5
UC-Davis	Siemens Skyra 3T	19	0.3mm 0.3mm 0.3mm	0.7mm 0.7mm 1.4mm	1600	60	6
Sinai-Philips	Philips Achieva 3T	9	0.3mm 0.3mm 0.3mm	1mm 1mm 1mm	1000	120	2
Sinai-Siemens	Siemens Skyra 3T	6	0.5mm 0.5mm 0.5mm	1mm 1mm 1mm	1000	80	10

Table 1. Acquisition parameters for the sites of the PRIME-DE database containing diffusion data.

in table 1. All datasets were acquired on 3T systems, with DWI at spatial resolutions from 0.7 to 1.4 mm^3 , with either isotropic or rectangular voxels, and anatomical images at voxels sizes ranging from 0.3 mm^3 to 0.8 mm^3 isotropic.

2.5 Quality control

All input data was quality controlled on both DW and anatomical images. Images were controlled for a variety of possible artifacts, including motion, ghosting, aliasing, gibbs ringing, B0 and B1 field inhomogeneities and magnetic susceptibility. After evaluation, images from Mount Sinai School of Medicine - Siemens were excluded due to non-uniformity of contrasts across both DW and anatomical images from hyper-intensities caused by restraints positioned around the subject's head. The 3 other sites presented high enough image quality and were judged good enough for the study.

2.6 Brain masking

Brain masks were computed prior to processing on the anatomical T1 images using DeepBet v1.0 [Wang et al. (2021)], a U-Net trained on macaque data from the PRIME-DE. All masks were quality controlled and manually fixed to prevent exclusion of brain voxels and inclusion of skull and eyes. The inclusion of this technology in the pipeline was considered, but was ultimately rejected due to the weight in memory of its dependencies, practically doubling the size of the container required to run the pipeline. Moreover, the lack of validation on databases other than the PRIME-DE, which was used to generate the segmentation

models, and its insertion near the beginning of the processing chain was deemed too risky, since its failure at providing a good brain mask would impede on most of all subsequent steps executed.

2.7 Processing hardware

The processing was done launching the pipeline separately for all 3 selected databases to evaluate the effectiveness of execution. We used 3 nodes of the Compute Canada Beluga cluster, each equipped with 2 Intel Gold 6148 Skylake processors at 2.4 GHz clock speed, 186 Gb of RAM and 4 NVidia V100SXM2 graphics cards, 16 Gb VRAM.

2.8 Data quality evaluation

To quantify the increase in quality of the processed data, two measures, SNR and CNR, were computed inside each region of the registered tissue maps on the mean b_0 volume of the diffusion weighted images. Topup correction and upsampling was performed on the raw data, to allow for the use of the tissue maps when computing their SNR and CNR. Our method thus does not quantify the effect of those steps on the amelioration of the data quality.

To calculate SNR, two different approaches were used to quantify the noise standard deviation on the b_0 . For DWI containing multiple b_0 volumes, intermediate noise maps were extracted by subtracting each possible permutations of two b_0 . Those maps were then averaged to obtain a mean noise map, which was used to compute the noise standard deviation in each region. SNR was then calculated using equation 2 [Narasimhan and Jacobs (2002)]. For DWI with a single b_0 volume, a background mask was extracted using the technique presented in [Balan et al. (2012)] and the noise standard deviation was calculated from the voxels of the b_0 inside it. SNR was calculated using equation 3. In both equation 2 and 3, b_0 refers to the b_0 image, N to the extracted mean noise map, T_{mask} to the mask of the tissue of interest and B_{mask} to the image background mask. $\mu_M(I)$ and $\sigma_M^2(I)$ respectively refer to the mean and variance of all voxels of an image I included in a mask M .

$$\text{SNR}(b_0, N, T_{\text{mask}}) = \frac{\mu_{T_{\text{mask}}}(b_0)}{\sigma_{T_{\text{mask}}}^2(N)} \quad (2)$$

$$\text{SNR}(b_0, T_{\text{mask}}, B_{\text{mask}}) = \frac{\mu_{T_{\text{mask}}}(b_0)}{\sigma_{B_{\text{mask}}}^2(b_0)} \quad (3)$$

CNR was computed on the b_0 by comparing white matter with gray matter regions using equation 4.

$$\text{CNR}(I, \text{WM}_{\text{mask}}, \text{GM}_{\text{mask}}) = \frac{|\mu_{\text{WM}_{\text{mask}}}(I) - \mu_{\text{GM}_{\text{mask}}}(I)|}{\sigma_{\text{WM}_{\text{mask}}}^2(I) + \sigma_{\text{GM}_{\text{mask}}}^2(I)} \quad (4)$$

We also performed a variability study in tissue maps (WM and GM) of several metrics extracted from the diffusion models available in the pipeline. Using the average over sessions of subjects, we analyzed inter-subject variability separately for each database (equation 5), intra-vendor variability for Siemens scanners (UC-Davis and Aix-Marseille) (equation 8) and inter-vendor variability using all sites (equation 7). Since UC-Davis and Sinai-Philips presented multiple time points per subject, intra-subject variability was also computed (equation 6).

$$\text{CV}_{\text{inter-subject}}(\{\hat{S}_1, \dots, \hat{S}_n\} \in D, M) = \frac{\sigma_M(D)}{\mu_M(D)} \quad (5)$$

$$CV_{\text{intra-subject}}(\{S_{1,i}, \dots, S_{1,k}\} \in S, M) = \frac{\sigma_M(S)}{\mu_M(S)} \quad (6)$$

$$CV_{\text{inter-vendor}}(\{\hat{S}_{1,D_1}, \dots, \hat{S}_{n,D_1}, \dots, \hat{S}_{1,D_v}, \dots, \hat{S}_{i,D_v}\} \in V, M) = \frac{\sigma_M(V)}{\mu_M(V)} \quad (7)$$

$$CV_{\text{intra-vendor}}(\{\hat{S}_{1,D_1}, \dots, \hat{S}_{n,D_1}, \dots, \hat{S}_{1,D_v}, \dots, \hat{S}_{i,D_v}\} \in V_i, M) = \frac{\sigma_M(V_i)}{\mu_M(V_i)} \quad (8)$$

Here, S denotes a set of images (T1, DWI, masks and measures) of a single subject, across all of its repetitions, D a single acquisition site (Aix-Marseille, Sinai-Philips or UC-Davis), V a vendor (Philips, Siemens) and M a mask defining a region of interest in which is evaluated the variation measure.

3 RESULTS

Figure 8. Execution time of processes run on datasets from the (A) Aix-Marseille, (B) Mount Sinai - Philips and (C) UC-Davis sites. (D) presents the global time consumption per process over all sites.

3.1 Processing time

Per process execution time is displayed in figure 8 separately for each database. Most of the processing time was occupied by 4 algorithms, namely Topup (~ 66 min avg), Eddy (~ 36 min avg), DIAMOND (~ 31 min avg) and fODF metrics computation (~ 15 min avg); those times are heavily bound to the diffusion image resolution, here ranging from 1mm to 1.25mm isotropic. Every other major step of the pipeline executed in about 5 minutes or less. It took about 6 hours processing the 4 dataset of the Aix-Marseille site, 25 hours for the 27 (9 subjects, 3 sessions) of Sinai-Philips and 34 hours for the 38 (19 subject, 2 sessions) of UC-Davis. Note that for UC-Davis, the number of iterations for Topup had to be increased, due to the intense susceptibility artifacts present in the images, thus drastically increasing the execution time.

3.2 Image quality

Signal to noise ratio (SNR) is presented in figure 9, on the average b0 volume extracted from both raw and preprocessed DW images. With the exception of UC-Davis, the measure suggests an increase in image quality resulting from the different denoising algorithms used in the preprocessing part of the pipeline. For Sinai-Philips and UC-Davis, SNR distribution displays less variance after preprocessing, with a substantial increase of mean SNR for Sinai-Philips.

Figure 9. SNR of b0 images inside (WM) white matter, (GM) gray matter and (CSF) cerebro spinal-fluid tissues

Contrast to noise ratio (CNR) in white matter voxels is displayed in a similar fashion in figure 10. As for SNR, distributions of CNR display less variance after preprocessing for Sinai-Philips and UC-Davis. Mean CNR is also increased for Aix-Marseille and Sinai-Philips.

Figure 10. CNR of b0 images

3.3 Reproducibility

Table 2 shows reproducibility scores on relevant output images of the pipeline. For classical diffusion models (DTI and fODF), as well as for anatomical mask computing and registration processes, reproducibility is

high, with I2C2 scores of 98% and more. Scores for the DIAMOND model averages around 97% for metrics computed over the fascicles (MD, AD, RD, FA, max fFA, free-water fraction) and 91% model selection, indicating that the outcomes of this algorithm are less reproducible. Gamma distribution parameters (kappa, kappaAD, hei and heiAD) do not perform as well, with kappa and hei scoring at 83% and heiAD and kappaAD at 52% and 63%.

Modality		Measure	I2C2
Anatomy	T1	WM PVF	0.997
	T1	GM PVF	0.995
	T1	CSF PVF	0.990
	T1	registration	1.000
Diffusion	DTI	MD	0.991
	DTI	AD	0.990
	DTI	RD	0.991
	DTI	FA	0.999
	fODF	AFD	0.981
	fODF	AFD total	0.997
	fODF	NuFO	0.986
	DIAMOND	MD	0.962
	DIAMOND	AD	0.975
	DIAMOND	RD	0.956
	DIAMOND	FA	0.990
	DIAMOND	max fFA	0.979
	DIAMOND	hei	0.831
	DIAMOND	heiAD	0.523
	DIAMOND	kappa	0.831
	DIAMOND	kappaAD	0.630
	DIAMOND	fascicle fractions	0.968
	DIAMOND	free-water fraction	0.981
	DIAMOND	model selection	0.903

Table 2. Reproducibility measures on images generated by the pipeline.

3.4 Variation of measurements

Figures 11 to 13 display the coefficient of variation of several metrics outputted by the pipeline, and figure 14 on the gamma distribution parameters estimated by the DIAMOND model, calculated in both white matter (WM) and gray matter (GM) masks.

The Sinai-Philips site presents the least mean intra-subject variability overall at $\leq 2\%$, except for NuFO, AFD and free-water fraction (fFW), but the worst inter-subject variability in almost all metrics and parameters evaluated.

For DTI metrics, Aix-Marseille displays better inter-subject variability in both white matter ($\leq 4\%$ in MD and RD and $\leq 7\%$ in RD and FA) and gray matter ($\sim 3\%$) than Sinai-Philips. The portrait is reversed for fODF measurements, with the exception of AFDt, where the variability observed is greater (more than twice for NuFO in WM). For metrics computed on the DIAMOND model, the variability is better, with the exception of MD in WM and of all measures excluding free-water fraction (fFW) in GM. For gamma distribution parameters, variability is also higher, except for kappaAD and heiAD in WM.

Figure 11. Coefficient of variation (CV) for DTI measurements : (MD) mean diffusivity, (AD) axial diffusivity, (RD) radial diffusivity, (FA) fractional anisotropy

Figure 12. Coefficient of variation (CV) for fODF measurements : (AFD) max apparent fiber density, (AFDt) total apparent fiber density, (NuFO) number of fiber orientations

407 UC-Davis exhibits good inter-subject variability in DTI measurements at less than 4% for MD, AD
 408 and RD and around 6% in WM and 8% in GM for FA. For fODF metrics in WM, AFDt scores similarly
 409 at 4%, AFD at about 9% and NuFO presents as the least reliable at 14%. Oddly, mean intra-subject
 410 variability is greater than inter-subject variability for this site for almost all quantities observed, except for
 411 fractional anisotropy (with the exception of the mean FA on DIAMOND fascicles), AFDt (in WM only)

Figure 13. Coefficient of variation (CV) for diamond measurements : (MD) mean mean diffusivity among fascicles, (FA) mean fractional anisotropy among fascicles, (max fFA) maximum fascicle FA, (fFW) fraction of free-water

412 and free-water measurements, as well as for gamma distribution parameters excluding heiAD in WM and
 413 kappa in GM. With further inspections, it was noticed that for all measurements exhibiting this trend, two
 414 clusters of similar measurements could be formed, corresponding each to one of the sessions at which the
 415 subjects' images were acquired, as presented in figure 15.

Figure 14. Coefficient of variation (CV) for diamond gamma distributions parameters : (hei) heterogeneity index, (heiAD) axial heterogeneity index, (kappa) kappa parameter of tensor distribution, (kappaAD) axial kappa parameter of tensor distribution

Intra-vendor variability of Siemens scanners is computed on subjects from both Aix-Marseille and UC-Davis. For DTI metrics, variability in WM ranges from 12% to 14% for diffusivity measures and at around 8% for FA and is slightly lower in GM. For fODF, NuFO displays the highest variability (21% in WM, 32% in GM), followed by AFD (10% in WM, 22% in GM). Variability for AFDt is substantially lower at 4% in WM and 7% in GM. For measurements on the DIAMOND model in WM, MD is the least variable at 6% and fractional anisotropy scores at 10%, 2% below the fascicle based measure (max fFA). Results shows that performance from those measure is approximately two times more variable in GM. Free-water fraction is the most variable measure estimated at 28% in WM and 21% in GM. For gamma distribution parameters, values range widely depending on the parameter and on the tissue. For WM, kappa and kappaAD are the less variable at 7% and 10%, and hei and heiAD score at 33% and 18% respectively. For GM, the more variable parameter is kappaAD at 32%, followed by kappa and hei at around 17% and heiAD at 13%.

Figure 15. Clustering behaviour of the acquisition sessions of the subjects from the UC-Davis site.

Inter-vendor variability is consistently higher than intra-vendor on all DTI measurements, with an increase in variability of 2% in FA and from 4% to 10% in diffusivity measures in both WM and GM. The same trend is observed for fODF metrics, AFDt displaying the less increase (2%) and NuFO the most (5%) in WM. For measures on the DIAMOND model, MD shows the greater increase (6% in WM and 3% in GM). FA and maxfFA variability remains stable, with a slight increase of 2% for the former and less than 1% for the latter. Free-water fraction variability increases substantially of about 5%. For gamma distribution parameters, a slight decrease of variability is observed for all parameters, except for hei which remains stable in WM and increases of about 3% in GM.

4 DISCUSSION

4.1 Reproducibility

Reproducibility is of prior importance when it comes to a good processing pipeline. It is a measure of how much a pipeline can, now and later in time, given the same data, configuration and underlying algorithm's versions, output data that behaves the same and give similar results when analyzed. As for tractoflow [Theaud et al. (2020)], a novel pipeline for dMRI analysis of human subjects, this pipeline was designed with reproducibility in mind. Results displayed in table 2 confirm good reproducibility from DTI and fODF models computed, as well as from T1 registration processes (using ANTs registration) and tissues partial volume fractions computation (using ANTs Atropos). Scores are slightly lower for metrics estimated on the DIAMOND model, and significantly lower for gamma distribution parameters (kappa, kappaAD, hei, heiAD). This behaviour is expected since the DIAMOND implementation included in the pipeline does not allow to set the random number seed, which injects a quantity of variability in the outcomes of the algorithm. We would nonetheless advise against their usage in a statistical study as is, since they are bound to introduce a level of variability in the analysis. Note that this does not affect the measurements calculated (MD, FA, max fFA and fFW), since in the current implementation of the pipeline, their computation only rely on the mean tensor estimated on each fascicle, thus are independent of the gamma distribution parameters.

4.2 Pipeline implementation

In addition to providing a highly reproducible processing chain, our pipeline implementation also offers a high level of adaptability to the parametrization of its input's acquisition sequences. Whilst designing the collection of processing modules, we took great care to make their configuration resilient to different spatial (voxel size, number of slices, orientation of slice and/or phase encoding) and orientational (number of b-values and/or gradient directions) configurations. This proved critical to adequately scale the parameter space of algorithms used by the pipeline, such as Topup, which rely on a definition of its deformation field given in millimetres in voxel space, and to allow disabling their execution when input requirements were unmet.

Nevertheless, we are aware the default values we offer could vary as algorithms are updated and that some use-case could benefit changing them. To that extent, is provided with most of them an additional configuration layer, presenting all possible parameters, to allow the end-user to finely adapt the pipeline to its specific study. We also allowed the user to disable all steps included in the pipeline, in the case one desires only to compute models and measures and has already preprocessed its data, or to speed up computation time by disabling steps deemed unnecessary after thorough quality control. Still, we do not recommend skipping any preprocessing steps we included in the pipeline when providing raw data from the scanner.

Furthermore, our implementation is ready for new use-cases for which our pipeline does not provide the adequate preprocessing and modeling steps. Using our modular design, a pipeline can easily be derived from ours, using new algorithms, workflows and/or input conventions, by adding the necessary objects in their respective module and binding their dataflow correctly.

4.3 Evaluation of variability

Lowering variability of measurements as much as possible is of prior importance in a study as it can destroy any predictive power of an analysis. In this study, we evaluated from the PRIME-DE database, using 3 sites that provided enough good quality DWI images, as well as T1w images, the variability of measurements provided by our pipeline on 4 levels : intra-subject for Sinai-Philips and UC-Davis, inter-subject for all 3

478 sites, intra-vendor for Siemens scanners (Aix-Marseille and UC-Davis) and inter-vendor, that is across all
479 3 sites.

480 Our results reflect common problems that still need to be addressed by the community to make MRI a fully
481 reliable imaging modality. The increase of variability reported when including data from multiple scanner
482 vendors and from multiple sites is representative of the need for data harmonization protocols capable of
483 reducing the effects induced by the acquisition software and technology [Vollmar et al. (2010)][Tax et al.
484 (2019)][Grech-Sollars et al. (2015)] as well as signal deviation occurring from environmental differences
485 (temperature, humidity, time of day, etc.) [Book et al. (2021)][Meyer et al. (2016)]. Moreover, it cries at
486 the need for consensus on acquisition protocols and to the creation and application of guidelines by the
487 community, without which, group efforts like the PRIME-DE project cannot become fully relevant. As
488 seen in table 1, acquisition parameters vary widely from one site to another, especially when it comes to the
489 sampling of diffusion gradients' directions. This has a direct impact on the validity of the reconstruction
490 of diffusion models, since different b-values encodes different processes occurring in the imaged tissues
491 [Veraart et al. (2019)][McKinnon et al. (2017)].

492 In our study, gradient sampling seems to have a limited impact on DTI measurements for Aix-Marseille,
493 since shells with b-values higher than $1300 \text{ mm}^2/s$ are omitted. This however has a greater effect when
494 comparing other sites to UC-Davis, since DTI was reconstructed for it from the only available shell at
495 $1600 \text{ mm}^2/s$, explaining the 3 to 4 fold increase in Siemens intra-vendor and global inter-vendor variability
496 relative to the average inter-subject variability of associated sites. The impact becomes greater when
497 looking at fODF and multi-tensor reconstruction (DIAMOND), since those model make use of all b-values.
498 This makes Aix-Marseille inter-subject variability increase since the analysis becomes a comparison
499 between 2 subjects acquired with single-shell sampling, and 2 others with multi-shell, the latter presenting
500 a richer portrait of the diffusion process by the usage of higher gradient weighting. Both fODF and
501 multi-tensor reconstruction are known to benefit from the higher b-values [Yang et al. (2015)][Taquet et al.
502 (2015)][Scherrer and Warfield (2012)][Scherrer and Warfield (2010)]; the quantities estimated from those
503 models are destined to vary from the case of a low b-value single-shell acquisition.

504 Another trend observed in our study hinting to the need of guidelines and protocols is the high mean
505 intra-subject variability observed for UC-Davis. While analyzing this site, we noticed a strong discrepancy
506 of diffusion measurements between the two sessions images were acquired at. This behaviour could be the
507 result of equipment upgrades, software or hardware, as well as poor sequence configuration, resulting in an
508 ill-conditioning of the tissues required for diffusion acquisition. The only quantities were this behaviour
509 was not detected are fractional anisotropies (FA and max fFA) and total apparent fiber density (AFDt).
510 Nevertheless, this makes the usage of separate sessions from UC-Davis impossible in a study of diffusion
511 measurements as the variability would erase any statistical power of an analysis. We hope in the future,
512 good quality control could be integrated in the scanners to prevent or correct the acquisition of images
513 presenting this kind of behaviour.

5 LIMITATIONS AND FUTURE WORK

514 In the current version of our pipeline, we implemented a robust and efficient preprocessing chain of
515 state-of-the-art algorithms to correct for well know artifacts affecting the MRI signal, such as background
516 noise, gibbs ringing, susceptibility distortions, motion, signal dropout and intensity disparities within
517 tissues. We are well aware that our technique is not universal and that some use-cases may require the
518 execution of different algorithms than the one we considered. MRI preprocessing still need to be thoroughly
519 validated; the effect of applying specific techniques and algorithms, as well as their ordering relative to one

another, requires to be quantified in order to attain an adequate level of consensus in the community. We designed our pipeline and its modular framework taking those facts into account, to ease the process of swapping and upgrading its different steps, in order to fit specific study cases. We also plan to release in the near future a version of the pipeline integrating alternative denoising techniques that are popular in the diffusion MRI research community based on recommendations from the ISMRM diffusion study consensus effort, once their work becomes available. Inasmuch, we intent to include brain extraction workflows, for both anatomical and diffusion weighted images, adapted to the morphology of input subjects. Good solutions have been developed that display good resiliency on human subjects, but their extension to other primates and to small animals is still to be validated and can lead to segmentation lacking specificity. As the field continues to evolve and those issues become addressed, we will consider the addition of those solutions in our pipeline.

Moreover, for our implementation, we chose a limited subset of 3 local models to represent the diffusion process in dMRI voxels. More algorithms, especially in the category of multi-tensor models, could be considered in the future, such as the ball and stick [Behrens et al. (2003)][Behrens et al. (2007)][Jbabdi et al. (2012)], ball and zeppelin [Sotiropoulos et al. (2016)] models and NODDI [Zhang et al. (2012)]. The inclusion of Diffusion Kurtosis Imaging [Jensen et al. (2005)] could also be considered. More general models of the Ensemble Averaged Propagator (EAP) could also be added, using for example the Simple Harmonic Oscillator Based Reconstruction and Estimation (3D-SHORE) or the Mean Apparent Propagator (MAP) [Özarslan et al. (2013)], along with relevant measurements computed on them. Furthermore, we plan to release supplemental modules to compute tractograms upon the computed models the pipeline offers.

It is part of future work to adress better computing resources management, there is room for improvement. The Nextflow scripting language offers extensive capabilities to define per process resources prerequisites. Using them, we were able to define clearly the number of processing cores (CPU) and need for hardware accelerators (GPU) for all processes in our library and ensure a high level of parallelization. However, we do not yet control for requirements in memory (RAM) of each individual process executed. The lack of specification for memory usage of multiple algorithms we used, combined with the difficulty of obtaining the actual memory space of compressed NIFTI images using the Nextflow language makes it hard to define a clear requirement to respect at execution time. Specifying an overestimated precondition on memory needs could lead to a fewer number of processes executed at the same time, which would impede the parallelization capabilities of the pipeline and result in a waste of computing resources. We are aware that this behaviour could lead to execution problems, more so when processing high resolution images that spans multiple gigabytes on disk. To mitigate the effects of that limitation, we opted for a retry strategy for processes failing for that reason, combined with a restriction of the number of parallel executions of memory hungry processes.

6 CONCLUSION

In this work, we built a reliable and reproducible pipeline, able to tackle the task of preprocessing and computing diffusion models on multi-resolution diffusion MRI data. We also presented our modular library of Nextflow processes and workflows using state-of-the-art and cutting edge DWI processing technologies, designed to be easily upgradable and adaptable. We used our pipeline to analyze the variability of 3 sites (32 subjects) included in the PRIME-DE database presenting good quality DWI data. We showed that even if promising, that data exhibited a great level of variability and its usage should be done with care to prevent instilling uncertainty in statistical analyses. This is a good example of the effects of the lack of consensus in the diffusion MRI field when it comes to specifying guidelines for acquisition procedures.

The hurdle of defining a gold standard in dMRI makes this problem even more difficult to overpass, but a level of standardization and harmonization must be attained in order to reduce variability in imaging between subjects, scanner and sites as much as possible. Without it, conjugate efforts like the PRIME-DE become less relevant for the study of the brain and its properties. It is through conjoint endeavours such as the consensus effort from the ISMRM diffusion study group and others, and by the development of reproducible, robust and maintainable software, that we as a community will make of diffusion MRI a reliable modality.

DATA AVAILABILITY STATEMENT

The pipeline developed for this study and the nextflow modules library are available online at github.com/boiteaclou/mrHARDIfLOW. The processed diffusion data, including computed models and diffusion metrics, are available on Zenodo at doi.org/10.5281/zenodo.5719453.

REFERENCES

- Adluru, N., Zhang, H., Fox, A. S., Shelton, S. E., Ennis, C. M., Bartosic, A. M., et al. (2012). A diffusion tensor brain template for rhesus macaques. *NeuroImage* 59, 306–318. doi:10.1016/j.neuroimage.2011.07.029
- Ahn, C. and Cho, Z. (1991). Analysis of the eddy-current induced artifacts and the temporal compensation in nuclear magnetic resonance imaging. *Medical Imaging, IEEE Transactions on* 10, 47–52. doi:10.1109/42.75610
- Andersson, J. L., Graham, M. S., Drobnjak, I., Zhang, H., and Campbell, J. (2018). Susceptibility-induced distortion that varies due to motion: Correction in diffusion mr without acquiring additional data. *NeuroImage* 171, 277–295. doi:10.1016/j.neuroimage.2017.12.040
- Andersson, J. L., Graham, M. S., Drobnjak, I., Zhang, H., Filippini, N., and Bastiani, M. (2017). Towards a comprehensive framework for movement and distortion correction of diffusion mr images: Within volume movement. *NeuroImage* 152, 450–466. doi:10.1016/j.neuroimage.2017.02.085
- Andersson, J. L., Skare, S., and Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage* 20, 870–888. doi:10.1016/S1053-8119(03)00336-7
- Andersson, J. L. and Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion mr imaging. *NeuroImage* 125, 1063–1078. doi:10.1016/j.neuroimage.2015.10.019
- Andersson, J. L. R., Graham, M. S., Zsoldos, E., and Sotiropoulos, S. N. (2016). Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion mr images. *NeuroImage* 141, 556–572. doi:10.1016/j.neuroimage.2016.06.058
- Andica, C., Kamagata, K., Hayashi, T., Hagiwara, A., Uchida, W., Saito, Y., et al. (2020). Scan-rescan and inter-vendor reproducibility of neurite orientation dispersion and density imaging metrics. *Neuroradiology* 62, 483–494. doi:10.1007/s00234-019-02350-6
- Autio, J. A., Glasser, M. F., Ose, T., Donahue, C. J., Bastiani, M., Ohno, M., et al. (2020). Towards hcp-style macaque connectomes: 24-channel 3t multi-array coil, mri sequences and preprocessing. *NeuroImage* 215, 116800. doi:10.1016/j.neuroimage.2020.116800
- Avants, B., Tustison, N., and Song, G. (2008). Advanced normalization tools (ants). *Insight J* 1-35
- Avants, B. B., Tustison, N. J., Wu, J., Cook, P. A., and Gee, J. C. (2011). An open source multivariate framework for n-tissue segmentation with evaluation on public data. *Neuroinformatics* 9, 381–400. doi:10.1007/s12021-011-9109-y

- 604 Balan, A. G., Traina, A. J., Ribeiro, M. X., Marques, P. M., and Traina Jr., C. (2012). Smart histogram
605 analysis applied to the skull-stripping problem in t1-weighted mri. Computers in Biology and Medicine
606 42, 509–522. doi:10.1016/j.combiomed.2012.01.004
- 607 Behrens, T., Berg, H. J., Jbabdi, S., Rushworth, M., and Woolrich, M. (2007). Probabilistic diffusion
608 tractography with multiple fibre orientations: What can we gain? NeuroImage 34, 144–155. doi:10.
609 1016/j.neuroimage.2006.09.018
- 610 Behrens, T., Woolrich, M., Jenkinson, M., Johansen-Berg, H., Nunes, R., Clare, S., et al. (2003).
611 Characterization and propagation of uncertainty in diffusion-weighted mr imaging. Magnetic Resonance
612 in Medicine 50, 1077–1088. doi:10.1002/mrm.10609
- 613 Book, G. A., Meda, S. A., Janssen, R., Dager, A. D., Poppe, A., Stevens, M. C., et al. (2021). Effects of
614 weather and season on human brain volume. PLOS ONE 16, 1–16. doi:10.1371/journal.pone.0236303
- 615 Calamante, F. (2019). The seven deadly sins of measuring brain structural connectivity using diffusion mri
616 streamlines fibre-tracking. Diagnostics 9, 115. doi:10.3390/diagnostics9030115
- 617 Cieslak, M., Cook, P. A., He, X., Yeh, F.-C., Dhollander, T., Adebimpe, A., et al. (2021). Qsiprep:
618 an integrative platform for preprocessing and reconstructing diffusion mri data. Nature Methods 18,
619 775–778. doi:10.1038/s41592-021-01185-5
- 620 Constanzo, J., Dumont, M., Lebel, R., Tremblay, L., Whittingstall, K., Masson-Côté, L., et al. (2018).
621 Diffusion mri monitoring of specific structures in the irradiated rat brain. Magnetic Resonance in
622 Medicine 80, 1614–1625. doi:10.1002/mrm.27112
- 623 Cordero-Grande, L., Christiaens, D., Hutter, J., Price, A. N., and Hajnal, J. V. (2019). Complex diffusion-
624 weighted image estimation via matrix recovery under general noise models. NeuroImage 200, 391–404.
625 doi:10.1016/j.neuroimage.2019.06.039
- 626 Czervionke, L. F., Czervionke, J. M., Daniels, D. L., and Haughton, V. M. (1988). Characteristic features
627 of mr truncation artifacts. American Journal of Roentgenology 151, 1219–1228. doi:10.2214/ajr.151.6.
628 1219
- 629 De Santis, S., Bastiani, M., Droby, A., Kolber, P., Zipp, F., Pracht, E., et al. (2019). Characterizing
630 microstructural tissue properties in multiple sclerosis with diffusion mri at 7t and 3t: The impact of the
631 experimental design. Neuroscience 403, 17–26. doi:10.1016/j.neuroscience.2018.03.048
- 632 Dell’Acqua, F., Simmons, A., Williams, S. C. R., and Catani, M. (2012). Can spherical deconvolution
633 provide more information than fiber orientations? hindrance modulated orientational anisotropy, a
634 true-tract specific index to characterize white matter diffusion. Human brain mapping 34, 2464–2483.
635 doi:10.1002/hbm.22080
- 636 Dyrby, T. B., Lundell, H., Burke, M. W., Reisle, N. L., Paulson, O. B., Ptito, M., et al. (2014). Interpolation
637 of diffusion weighted imaging datasets. NeuroImage 103, 202–213. doi:10.1016/j.neuroimage.2014.09.
638 005
- 639 Fadnavis, S., Batson, J., and Garyfallidis, E. (2020). Patch2self: Denoising diffusion mri with self-
640 supervised learning. In Advances in Neural Information Processing Systems, eds. H. Larochelle,
641 M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin (Curran Associates, Inc.), vol. 33, 16293–16303
- 642 Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., Van Der Walt, S., Descoteaux, M., et al. (2014).
643 Dipy, a library for the analysis of diffusion mri data. Frontiers in Neuroinformatics 8, 8. doi:10.3389/
644 fninf.2014.00008
- 645 Gorgolewski, K., Burns, C., Madison, C., Clark, D., Halchenko, Y., Waskom, M., et al. (2011). Nipype: A
646 flexible, lightweight and extensible neuroimaging data processing framework in python. Frontiers in
647 Neuroinformatics 5, 13. doi:10.3389/fninf.2011.00013

- 648 Grech-Sollars, M., Hales, P. W., Miyazaki, K., Raschke, F., Rodriguez, D., Wilson, M., et al. (2015).
649 Multi-centre reproducibility of diffusion mri parameters for clinical sequences in the brain. NMR in
650 Biomedicine 28, 468–485. doi:10.1002/nbm.3269
- 651 Gudbjartsson, H. and Patz, S. (1995). The rician distribution of noisy mri data. Magnetic Resonance in
652 Medicine 34, 910–914. doi:10.1002/mrm.1910340618
- 653 Heilingoetter, C. L. and Jensen, M. B. (2016). Histological methods for ex vivo axon tracing: A systematic
654 review. Neurological research 38, 561—569. doi:10.1080/01616412.2016.1153820
- 655 Hosey, T., Williams, G., and Ansorge, R. (2005). Inference of multiple fiber orientations in high angular
656 resolution diffusion imaging. Magnetic Resonance in Medicine 54, 1480–1489. doi:10.1002/mrm.20723
- 657 Howard, A. F., Mollink, J., Kleinnijenhuis, M., Palbage-Gamarallage, M., Bastiani, M., Cottaar, M.,
658 et al. (2019). Joint modelling of diffusion mri and microscopy. NeuroImage 201, 116014. doi:10.1016/j.
659 neuroimage.2019.116014
- 660 Jbabdi, S., Sotiropoulos, S. N., Savio, A. M., Graña, M., and Behrens, T. E. (2012). Model-based analysis
661 of multishell diffusion mr data for tractography: How to get over fitting problems. Magnetic Resonance
662 in Medicine 68, 1846–1855. doi:10.1002/mrm.24204
- 663 Jensen, J. H., Helpert, J. A., Ramani, A., Lu, H., and Kaczynski, K. (2005). Diffusional kurtosis imaging:
664 The quantification of non-gaussian water diffusion by means of magnetic resonance imaging. Magnetic
665 Resonance in Medicine 53, 1432–1440. doi:10.1002/mrm.20508
- 666 Jeurissen, B., Tournier, J.-D., Dhollander, T., Connelly, A., and Sijbers, J. (2014). Multi-tissue constrained
667 spherical deconvolution for improved analysis of multi-shell diffusion mri data. NeuroImage 103,
668 411–426. doi:10.1016/j.neuroimage.2014.07.061
- 669 Jones, D. (2010). Challenges and limitations of quantifying brain connectivity in vivo with diffusion mri.
670 Imaging in Medicine 2, 341–355. doi:10.2217/iim.10.21
- 671 Jones, D. K. and Basser, P. J. (2004). “squashing peanuts and smashing pumpkins”: How noise distorts
672 diffusion-weighted mr data. Magnetic Resonance in Medicine 52, 979–993. doi:10.1002/mrm.20283
- 673 Jones, D. K., Knösche, T. R., and Turner, R. (2013). White matter integrity, fiber count, and other fallacies:
674 The do’s and don’ts of diffusion mri. NeuroImage 73, 239–254. doi:10.1016/j.neuroimage.2012.06.081
- 675 Jung, B., Taylor, P. A., Seidlitz, J., Sponheim, C., Perkins, P., Ungerleider, L. G., et al. (2021). A
676 comprehensive macaque fmri pipeline and hierarchical atlas. NeuroImage 235, 117997. doi:10.1016/j.
677 neuroimage.2021.117997
- 678 Kawamura, M., Tamada, D., Funayama, S., Kromrey, M.-L., Ichikawa, S., Onishi, H., et al. (2021).
679 Accelerated acquisition of high-resolution diffusion-weighted imaging of the brain with a multi-shot
680 echo-planar sequence: Deep-learning-based denoising. Magnetic Resonance in Medical Sciences 20,
681 99–105. doi:10.2463/mrms.tn.2019-0081
- 682 Kellner, E., Dhital, B., Kiselev, V. G., and Reiser, M. (2016). Gibbs-ringing artifact removal based on
683 local subvoxel-shifts. Magnetic Resonance in Medicine 76, 1574–1581. doi:10.1002/mrm.26054
- 684 Kurtzer, G. M., Sochat, V., and Bauer, M. W. (2017). Singularity: Scientific containers for mobility of
685 compute. PLOS ONE 12, 1–20. doi:10.1371/journal.pone.0177459
- 686 LeBihan, D., Mangin, J.-F., Poupon, C., Clark, C. A., Pappata, S., Molko, N., et al. (2001). Diffusion
687 tensor imaging: Concepts and applications. Journal of Magnetic Resonance Imaging 13, 534–546.
688 doi:10.1002/jmri.1076
- 689 Luisier, F., Blu, T., and Wolfe, P. J. (2012). A cure for noisy magnetic resonance images: Chi-square
690 unbiased risk estimation. IEEE Transactions on Image Processing 21, 3454–3466. doi:10.1109/TIP.
691 2012.2191565

- 692 Maier-Hein, K., Neher, P., Houde, J.-C., Côté, M.-A., Garyfallidis, E., Zhong, J., et al. (2017). The
693 challenge of mapping the human connectome based on diffusion tractography. Nature Communications
694 8, 1349. doi:10.1038/s41467-017-01285-x
- 695 Manjón, J. V., Carbonell-Caballero, J., Lull, J. J., García-Martí, G., Martí-Bonmatí, L., and Robles, M.
696 (2008). Mri denoising using non-local means. Medical Image Analysis 12, 514–523. doi:10.1016/j.
697 media.2008.02.004
- 698 McKinnon, E. T., Jensen, J. H., Glenn, G. R., and Helpert, J. A. (2017). Dependence on b-value of
699 the direction-averaged diffusion-weighted imaging signal in brain. Magnetic Resonance Imaging 36,
700 121–127. doi:10.1016/j.mri.2016.10.026
- 701 Meyer, C., Muto, V., Jaspard, M., Kussé, C., Lambot, E., Chellappa, S. L., et al. (2016). Seasonality in
702 human cognitive brain responses. Proceedings of the National Academy of Sciences 113, 3066–3071.
703 doi:10.1073/pnas.1518129113
- 704 Milham, M. P., Ai, L., Koo, B., Xu, T., Amiez, C., Balezeau, F., et al. (2018). An open resource for
705 non-human primate imaging. Neuron 100, 61–74. doi:10.1016/j.neuron.2018.08.039
- 706 Narasimhan, P. and Jacobs, R. E. (2002). 16 - neuroanatomical micromagnetic resonance imaging. In Brain
707 Mapping: The Methods (Second Edition), eds. A. W. Toga and J. C. Mazziotta (San Diego: Academic
708 Press). Second edition edn., 399–426
- 709 Neff, E. P. (2019). Prime-de is primed with primate data. Lab Animal 48, 26. doi:10.1038/
710 s41684-018-0209-2
- 711 O'Donnell, L. J. and Pasternak, O. (2015). Does diffusion mri tell us anything about the white matter? an
712 overview of methods and pitfalls. Schizophrenia Research 161, 133–141. doi:10.1016/j.schres.2014.09.
713 007
- 714 Pierpaoli, C., Walker, L., Irfanoglu, M. O., Barnett, A., Basser, P., Chang, L.-C., et al. (2010). Tortoise:
715 An integrated software package for processing of diffusion mri data. In ISMRM 18th annual meeting
716 (Stockholm, Sweden), vol. 1597
- 717 Raffelt, D., Tournier, J.-D., Rose, S., Ridgway, G. R., Henderson, R., Crozier, S., et al. (2012). Apparent
718 fibre density: A novel measure for the analysis of diffusion-weighted magnetic resonance images.
719 NeuroImage 59, 3976–3994. doi:10.1016/j.neuroimage.2011.10.045
- 720 Reymbaut, A. and Descoteaux, M. (2019). Advanced encoding methods in diffusion MRI. doi:1908.
721 04177v3
- 722 Rheault, F., Poulin, P., Caron, A. V., St-Onge, E., and Descoteaux, M. (2020). Common misconceptions,
723 hidden biases and modern challenges of dMRI tractography. Journal of Neural Engineering 17, 011001.
724 doi:10.1088/1741-2552/ab6aad
- 725 Scherrer, B., Schwartzman, A., Taquet, M., Sahin, M., Prabhu, S. P., and Warfield, S. K. (2016).
726 Characterizing brain tissue by assessment of the distribution of anisotropic microstructural environments
727 in diffusion-compartment imaging (diamond). Magnetic Resonance in Medicine 76, 963–977.
728 doi:10.1002/mrm.25912
- 729 Scherrer, B. and Warfield, S. K. (2010). Why multiple b-values are required for multi-tensor models.
730 evaluation with a constrained log-euclidean model. In 2010 IEEE International Symposium on
731 Biomedical Imaging: From Nano to Macro. 1389–1392. doi:10.1109/ISBI.2010.5490257
- 732 Scherrer, B. and Warfield, S. K. (2012). Parametric representation of multiple white matter fascicles from
733 cube and sphere diffusion mri. PLOS ONE 7, 1–20. doi:10.1371/journal.pone.0048232
- 734 Schilling, K., Gao, Y., Janve, V., Stepniewska, I., Landman, B., and Anderson, A. (2017). Confirmation of
735 a gyral bias in diffusion mri fiber tractography. Human Brain Mapping 39, 1449–1466. doi:10.1002/
736 hbm.23936

- 737 Schilling, K., Tax, C., Rheault, F., Hansen, C., Yang, Q., Yeh, F.-C., et al. (2021a). Fiber tractography
738 bundle segmentation depends on scanner effects, vendor effects, acquisition resolution, diffusion
739 sampling scheme, diffusion sensitization, and bundle segmentation workflow. NeuroImage 242, 118451.
740 doi:10.1016/j.neuroimage.2021.118451
- 741 Schilling, K. G., Tax, C. M., Rheault, F., Landman, B., Anderson, A., Descoteaux, M., et al. (2021b).
742 Prevalence of white matter pathways coming into a single diffusion mri voxel orientation: the bottleneck
743 issue in tractography. bioRxiv doi:10.1101/2021.06.22.449454
- 744 Shou, H., Eloyan, A., Lee, S., Zipunnikov, V., Crainiceanu, A. N., Nebel, M. B., et al. (2013). Quantifying
745 the reliability of image replication studies: The image intraclass correlation coefficient (i2c2). Cognitive,
746 Affective & Behavioral Neuroscience 13, 714–724. doi:10.3758/s13415-013-0196-0
- 747 Smith, S. M. (2002). Fast robust automated brain extraction. Human Brain Mapping 17, 143–155.
748 doi:10.1002/hbm.10062
- 749 Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., et al.
750 (2004). Advances in functional and structural mr image analysis and implementation as fsl. NeuroImage
751 23, S208–S219. doi:10.1016/j.neuroimage.2004.07.051
- 752 Sotiropoulos, S. N., Hernández-Fernández, M., Vu, A. T., Andersson, J. L., Moeller, S., Yacoub, E., et al.
753 (2016). Fusion in diffusion mri for improved fibre orientation estimation: An application to the 3t and 7t
754 data of the human connectome project. NeuroImage 134, 396–409. doi:10.1016/j.neuroimage.2016.04.
755 014
- 756 Taquet, M., Scherrer, B., Boumal, N., Peters, J. M., Macq, B., and Warfield, S. K. (2015). Improved fidelity
757 of brain microstructure mapping from single-shell diffusion mri. Medical Image Analysis 26, 268–286.
758 doi:10.1016/j.media.2015.10.004
- 759 Tax, C. M., Grussu, F., Kaden, E., Ning, L., Rudrapatna, U., John Evans, C., et al. (2019). Cross-scanner
760 and cross-protocol diffusion mri data harmonisation: A benchmark database and evaluation of algorithms.
761 NeuroImage 195, 285–299. doi:10.1016/j.neuroimage.2019.01.077
- 762 Theaud, G., Houde, J.-C., Boré, A., Rheault, F., Morency, F., and Descoteaux, M. (2020). Tractoflow: A
763 robust, efficient and reproducible diffusion mri pipeline leveraging nextflow & singularity. NeuroImage
764 218, 116889. doi:10.1016/j.neuroimage.2020.116889
- 765 Tommaso, P. D., Chatzou, M., Floden, E. W., Barja, P. P., Palumbo, E., and Notredame, C. (2017). Nextflow
766 enables reproducible computational workflows. Nature Biotechnology 35, 316–319. doi:10.1038/nbt.
767 3820
- 768 Tournier, J.-D., Calamante, F., and Connelly, A. (2007). Robust determination of the fibre orientation
769 distribution in diffusion mri: Non-negativity constrained super-resolved spherical deconvolution.
770 NeuroImage 35, 1459–1472. doi:10.1016/j.neuroimage.2007.02.016
- 771 Tournier, J.-D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., et al. (2019). Mrtrix3: A
772 fast, flexible and open software framework for medical image processing and visualisation. NeuroImage
773 202, 116137. doi:10.1016/j.neuroimage.2019.116137
- 774 Tuch, D. S., Reese, T. G., Wiegell, M. R., Makris, N., Belliveau, J. W., and Wedeen, V. J. (2002). High
775 angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. Magnetic
776 Resonance in Medicine 44, 577–582. doi:10.1002/mrm.10268
- 777 Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., et al. (2010). N4itk:
778 Improved n3 bias correction. IEEE Transactions on Medical Imaging 29, 1310–1320. doi:10.1109/TMI.
779 2010.2046908
- 780 Veraart, J., Fieremans, E., and Novikov, D. S. (2016a). Diffusion mri noise mapping using random matrix
781 theory. Magnetic Resonance in Medicine 76, 1582–1593. doi:10.1002/mrm.26059

- 782 Veraart, J., Fieremans, E., and Novikov, D. S. (2019). On the scaling behavior of water diffusion in human
783 brain white matter. *NeuroImage* 185, 379–387. doi:10.1016/j.neuroimage.2018.09.075
- 784 Veraart, J., Novikov, D. S., Christiaens, D., Ades-aron, B., Sijbers, J., and Fieremans, E. (2016b). Denoising
785 of diffusion mri using random matrix theory. *NeuroImage* 142, 394–406. doi:10.1016/j.neuroimage.
786 2016.08.016
- 787 Vollmar, C., O’Muircheartaigh, J., Barker, G. J., Symms, M. R., Thompson, P., Kumari, V., et al. (2010).
788 Identical, but not the same: Intra-site and inter-site reproducibility of fractional anisotropy measures on
789 two 3.0t scanners. *NeuroImage* 51, 1384–1394. doi:10.1016/j.neuroimage.2010.03.046
- 790 Vos, S. B., Tax, C. M. W., Luijten, P. R., Ourselin, S., Leemans, A., and Froeling, M. (2017). The
791 importance of correcting for signal drift in diffusion mri. *Magnetic Resonance in Medicine* 77, 285–299.
792 doi:10.1002/mrm.26124
- 793 Wang, X., Li, X.-H., Cho, J. W., Russ, B. E., Rajamani, N., Omelchenko, A., et al. (2021). U-net model
794 for brain extraction: Trained on humans for transfer to non-human primates. *NeuroImage* 235, 118001.
795 doi:10.1016/j.neuroimage.2021.118001
- 796 Yang, J., Poot, D. H. J., van Vliet, L. J., and Vos, F. M. (2015). Estimation of diffusion properties
797 in three-way fiber crossings without overfitting. *Physics in Medicine & Biology* 60, 9123–9144.
798 doi:10.1088/0031-9155/60/23/9123
- 799 Zhang, H., Schneider, T., Wheeler-Kingshott, C. A., and Alexander, D. C. (2012). Noddi: Practical in vivo
800 neurite orientation dispersion and density imaging of the human brain. *NeuroImage* 61, 1000–1016.
801 doi:10.1016/j.neuroimage.2012.03.072
- 802 Özarslan, E., Koay, C. G., Shepherd, T. M., Komlosh, M. E., İrfanoğlu, M. O., Pierpaoli, C., et al.
803 (2013). Mean apparent propagator (map) mri: A novel diffusion imaging method for mapping tissue
804 microstructure. *NeuroImage* 78, 16–32. doi:10.1016/j.neuroimage.2013.04.016