Magnetic Resonance Imaging xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Magnetic Resonance Imaging

journal homepage: www.mrijournal.com

Reproducibility and biases in high field brain diffusion MRI: An evaluation of acquisition and analysis variables☆

Nico Dario Papinutto [a,b,*], Francesca Maule [b], Jorge Jovicich [b,c]

[a] Department of Neurology, University of California, San Francisco, CA, USA
[b] Center for Mind/Brain Sciences (CIMeC), University of Trento, Italy
[c] Department of Cognitive and Education Sciences, University of Trento, Italy

ARTICLE INFO

Article history:
Received 24 September 2012
Revised 29 January 2013
Accepted 8 March 2013
Available online xxxx

Keywords:
Diffusion tensor imaging
Reproducibility
Biases
Fractional anisotropy
Mean diffusivity
Principal eigenvalues
High field MRI

ABSTRACT

Diffusion tensor imaging (DTI) of *in-vivo* human brain provides insights into white matter anatomical connectivity, but little is known about measurement difference biases and reliability of data obtained with last generation high field scanners (>3 T) as function of MRI acquisition and analyses variables. Here we assess the impact of acquisition (voxel size: $1.8 \times 1.8 \times 1.8$, $2 \times 2 \times 2$ and $2.5 \times 2.5 \times 2.5$ mm^3, b-value: 700, 1000 and 1300 s/mm^2) and analysis variables (within-session averaging and co-registration methods) on biases and test-retest reproducibility of some common tensor derived quantities like fractional anisotropy (FA), mean diffusivity (MD), axial and radial diffusivity in a group of healthy subjects at 4 T in three regions: arcuate fasciculus, corpus callosum and cingulum. Averaging effects are also evaluated on a full-brain voxel based approach. The main results are: i) group FA and MD reproducibility errors across scan sessions are on average double of those found in within-session repetitions ($\approx$1.3 %), regardless of acquisition protocol and region; ii) within-session averaging of two DTI acquisitions does not improve reproducibility of any of the quantities across sessions at the group level, regardless of acquisition protocol; iii) increasing voxel size biased MD, axial and radial diffusivities to higher values and FA to lower values; iv) increasing b-value biased all quantities to lower values, axial diffusivity showing the strongest effects; v) the two co-registration methods evaluated gave similar bias and reproducibility results. Altogether these results show that reproducibility of FA and MD is comparable to that found at lower fields, not significantly dependent on pre-processing and acquisition protocol manipulations, but that the specific choice of acquisition parameters can significantly bias the group measures of FA, MD, axial and radial diffusivities.

1. Introduction

Diffusion tensor imaging (DTI) has become one of the most common MRI tools for *in vivo* characterization of normal and abnormal human white matter anatomical connectivity [1]. Changes in white matter tissue properties, particularly as seen in longitudinal studies, is manifested in normal ageing [2] as well as in other conditions like for example stroke recovery [3], Huntington's Disease [4], Alzheimer's Disease [5] and multiple sclerosis [6]. Such longitudinal studies offer the promise of establishing imaging-based biomarkers of disease, which may be of great utility in better understanding brain disorders and in evaluating therapeutic treatments.

Two of the most commonly used quantities derived from DTI data are fractional anisotropy (FA) and mean diffusivity (MD), scalar quantities that provide information on the white matter architecture at a voxel level [7]. Further information about white matter tissue architecture can be also obtained from the axial or parallel diffusivity (largest eigenvalue, often indicated as $\lambda||$) and the radial or perpendicular diffusivity ($\lambda\perp$, average of the two smallest eigenvalues) [8], being aware that their microscopic interpretation has to be done carefully, particularly when comparing patient populations in areas known to contain heterogeneous fiber orientations [9].

Methods based on probabilistic tractography have been recently introduced to also characterize the projection of nerve fiber bundles between brain areas [10,11]. Diffusion spectrum imaging (DSI) [12] and a variety of high angular resolution diffusion imaging (HARDI) techniques have been shown to better model white matter microstructure where fibers cross [13]. Although very promising,

☆ Author Contributions: Conceived and designed the experiments: NDP JJ. Performed the experiments: NDP. Analyzed the data: NDP FM. Contributed reagents/materials/analysis tools: NDP FM. Wrote the paper: NDP JJ.

* Corresponding author. Department of Neurology, University of California, San Francisco, 675 Nelson Rising Lane, San Francisco, CA 94158, USA. Tel.: +1 415 502 7253.
E-mail address: nico.papinutto@ucsf.edu (N.D. Papinutto).

0730-725X/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.mri.2013.03.004

these methods currently require much longer acquisition times than standard DTI so their implementation for longitudinal studies is more challenging, particularly for studies involving very young or aged subjects.

Performing reliable repeated measures of diffusion MRI-derived quantities is essential in longitudinal studies, since the reproducibility error sets a limit to the minimal changes that can be associated to the dynamics of the neurobiological processes under investigation. However, obtaining reproducible diffusion MRI-derived data is not trivial because the estimations depend on the combination of multiple data acquisition parameters and complex data analysis steps [14]. Several aspects of DTI reproducibility have been evaluated in repeated measures studies. These studies can be grouped in two broad categories, depending on whether the DTI acquisition parameters were manipulated or not. Studies with fixed acquisition parameters have focused on the reproducibility of the basic measuring method (typically FA and MD) by keeping the analysis method fixed [15–19] or by also manipulating the analysis, for example to look at region of interest (ROI) drawing protocols [20]. The other category of DTI reproducibility studies corresponds to the case in which some acquisition parameters were manipulated to evaluate their effect on the reproducibility of DTI-derived quantities. These studies evaluated how reproducibility was affected by several parameters, including signal-to-noise ratio (SNR) from data averaging at 1.5 T [21], b-values at 3 T [22], diffusion weighting schemes at 1.5 T [23] and 3 T [24], and scanner (vendor/field) [25–29]. A very recent study has also compared the reproducibility of DTI and DSI results at 3 T [30]. Most of these reproducibility studies manipulated one acquisition variable, did not evaluate the effects of spatial resolution, and the only study that investigated systematically the effects of averaging was done at 1.5 T with a single subject [21]. With the more common higher field systems, it is of interest to know whether the time investment of multiple acquisitions for within session averaging has a return in improved reproducibility at a group level. Similarly, the higher SNR available in the newer high field systems equipped with parallel imaging might allow improved reproducibility at higher spatial resolution, and these effects might relate to the choice of b-value. Choices of acquisition parameters might affect not only reproducibility but also the actual values of the FA and MD estimates [14,31]. In this paper we refer to such relative measurement differences as biases. Such effects may be important to know when multi-site studies cannot perfectly match acquisition protocols across all sites (e.g. when pooling data from retrospective studies). In addition, if voxel size is to be studied in an evaluation of biases and reproducibility then it is important to consider spatial smoothing effects introduced by data processing methods, in particular data co-registration.

The main goal of the present study was to evaluate the effect of data acquisition factors in the estimation of repeated measures of fractional anisotropy (FA), mean diffusivity (MD), axial and radial diffusivity as derived from DTI in a group of healthy volunteers using high-field MRI (4 T). Specifically, a subject-based ROI analysis was used to evaluate how the estimation of the chosen scalar quantities, as well as their test-retest reproducibility are affected by voxel size (1.8^3, 2.0^3, 2.5^3 mm^3) and by b-value (700, 1000, 1300 s/mm^2). The choice of voxel sizes and b-values was made to be within a range of values very often used at 1.5 T and 3 T (for example, [22,29,32]).

In addition (see Appendix), several analysis variables were also investigated to evaluate their effects on the biases and reproducibility of the diffusion quantities, namely: within-session data averaging (comparison of four different averaging methods and evaluation of whether averaging improves across-session reproducibility), and image co-registration (comparison of a subject-based ROI method that preserves each native resolution with atlas-based ROI and atlas-based full brain voxel analyses).

2. Materials and methods

2.1. Data acquisition and subjects

Ten subjects participated in this study (4 males, 6 females, age between 22 and 58 years; mean 32 ± 11 years). All subjects were healthy volunteers with no history of psychiatric, neurological or cognitive impairment, and provided written informed consent to participate in this study, which was approved by the Ethical Committee of the University of Trento.

Magnetic resonance images were acquired with a 4 T Bruker Medspec scanner (Bruker Medical, Ettlingen, Germany) using a birdcage-transmit, eight-channel receive head coil (USA Instruments, Inc., Ohio, USA). Each subject underwent 2 scan sessions on two separate dates at least a week apart. Each session included a T1-weighted structural image (3D MPRAGE, $1 \times 1 \times 1$ mm^3, TE = 4 ms; flip angle = 7°; iPAT parallel acceleration factor 2, TI = 1020 ms, bandwidth = 150 Hz/pixel, TA = 5 min) optimized for maximal contrast to noise ratio between grey and white matter at 4 T. In each session seven different diffusion weighted image protocols (Table 1) were acquired to later assess test-retest reproducibility effects on the following acquisition parameters: cubic voxel (voxel size, 1.8^3, 2^3, and 2.5^3 mm^3), b-value (700, 1000, 1300 s/mm^2) and number of acquisitions per session (1 or 2). The DTI acquisition protocol with voxel size $2 \times 2 \times 2mm^3$ was limited to only one b-value (1000 s/mm^2) to maintain the overall acquisition protocol under at total time of 90 minutes. The following acquisition parameters were kept fixed across all protocols and sessions: twice refocused 2D SE-EPI sequence [33], GRAPPA iPAT = 2, 5 images without any sensitizing diffusion gradient applied (hereafter called b0), 30 diffusion weighted images with diffusion gradients applied along unique directions that were defined by an electrostatic repulsion algorithm [34] axial slice acquisition along the x-y plane of the static magnetic field reference frame, two separate acquisitions per protocol, per session. The other parameters varied slightly across subjects and across the different resolutions, in the limits imposed by the relative dependence of acquisition parameters, by the specific absorption rate (SAR) and the nerve stimulation safety restrictions, with the goal of having a TE = 90 ms, an analogous brain coverage and a total scan time about 4:30 min per acquisition. The number of slices varied from 45 to 60 and the in-plane square FOV changed slightly across the various spatial resolutions: 230 mm^2, 256 mm^2 and 240 mm^2 for voxel sizes 1.8 mm^3, 2.0 mm^3, and 2.5 mm^3, respectively. Partial-Fourier (PF) factors of 7/8 or 1 (Full-Fourier) were used to reduce cardiac pulsation artifacts [35]. Overall, each of the ten subjects had 42 diffusion datasets for the analysis (7 acquisition protocols, 2 within session repetitions that could be treated separately or averaged (giving 2 + 1 datasets) and 2 sessions). Subject's brain in the second visit scan was positioned approximately at the center of the RF coil, following the same general guidelines used in the first visit.

Table 1
Acquisition protocols description. Each subject underwent two MRI scanning sessions on different days, and in each session the following diffusion weighted image protocols were acquired twice.

Protocol	b-value [s/mm^2]	Voxel size [mm^3]
P_1	700	$1.8 \times 1.8 \times 1.8$
P_2	1000	$1.8 \times 1.8 \times 1.8$
P_3	1300	$1.8 \times 1.8 \times 1.8$
P_4	1000	$2.0 \times 2.0 \times 2.0$
P_5	700	$2.5 \times 2.5 \times 2.5$
P_6	1000	$2.5 \times 2.5 \times 2.5$
P_7	1300	$2.5 \times 2.5 \times 2.5$

Please cite this article as: Papinutto ND, et al, Reproducibility and biases in high field brain diffusion MRI: An evaluation of acquisition and analysis variables, Magn Reson Imaging (2013), http://dx.doi.org/10.1016/j.mri.2013.03.004