NeuroImage 101 (2014) 390–403

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Multisite longitudinal reliability of tract-based spatial statistics in diffusion tensor imaging of healthy elderly subjects

Jorge Jovicich [a,*,1], Moira Marizzoni [b,1], Beatriz Bosch [c], David Bartrés-Faz [d], Jennifer Arnold [e], Jens Benninghoff [e], Jens Wiltfang [e,ag], Luca Roccatagliata [f,g], Agnese Picco [h], Flavio Nobili [h], Oliver Blin [i], Stephanie Bombois [j], Renaud Lopes [k], Régis Bordet [ai], Valérie Chanoine [l], Jean-Philippe Ranjeva [l], Mira Didic [m,n], Hélène Gros-Dagnac [o,p], Pierre Payoux [o,p], Giada Zoccatelli [q], Franco Alessandrini [q], Alberto Beltramello [q], Núria Bargalló [r], Antonio Ferretti [s,t], Massimo Caulo [s,t], Marco Aiello [u], Monica Ragucci [u], Andrea Soricelli [u,v], Nicola Salvadori [w], Roberto Tarducci [x], Piero Floridi [ah], Magda Tsolaki [y], Manos Constantinidis [z], Antonios Drevelegas [z,aa], Paolo Maria Rossini [ab,ac], Camillo Marra [ad], Josephin Otto [ae], Martin Reiss-Zimmermann [ae], Karl-Titus Hoffmann [ae], Samantha Galluzzi [b], Giovanni B. Frisoni [b,af], The PharmaCog Consortium

[a] Center for Mind/Brain Sciences (CIMEC), University of Trento, Rovereto, Italy
[b] LENITEM Laboratory of Epidemiology, Neuroimaging, & Telemedicine — IRCCS San Giovanni di Dio-FBF, Brescia, Italy
[c] Alzheimer's Disease and Other Cognitive Disorders Unit, Department of Neurology, Hospital Clínic, and IDIBAPS, Barcelona, Spain
[d] Department of Psychiatry and Clinical Psychobiology, Universitat de Barcelona and IDIBAPS, Barcelona, Spain
[e] LVR-Clinic for Psychiatry and Psychotherapy, Institutes and Clinics of the University Duisburg-Essen, Essen, Germany
[f] Department of Neuroradiology, IRCSS San Martino University Hospital and IST, Genoa, Italy
[g] Department of Health Sciences, University of Genoa, Genoa, Italy
[h] Department of Neuroscience, Ophthalmology, Genetics and Mother–Child Health (DINOGMI), University of Genoa, Genoa, Italy
[i] Pharmacology, Assistance Publique — Hôpitaux de Marseille, Aix-Marseille University — CNRS, UMR 7289, Marseille, France
[j] Department of Neurology, EA1046, Lille University, Lille, France
[k] Department of Neuroradiology, EA1046, Lille University, Lille, France
[l] CRMBM–CEMEREM, UMR 7339, Aix Marseille Université — CNRS, Marseille, France
[m] APHM, CHU Timone, Service de Neurologie et Neuropsychologie, Marseille, France
[n] Aix-Marseille Université, INSERM U 1106, Marseille, France
[o] INSERM, Imagerie cérébrale et handicaps neurologiques, UMR 825, Toulouse, France
[p] Université de Toulouse, UPS, Imagerie cérébrale et handicaps neurologiques, UMR 825, CHU Purpan, Place du Dr Baylac, Toulouse Cedex 9, France
[q] Department of Neuroradiology, General Hospital, Verona, Italy
[r] Department of Neuroradiology and Magnetic Resonace Image core Facility, Hospital Clínic de Barcelona, IDIBAPS, Barcelona, Spain
[s] Department of Neuroscience Imaging and Clinical Sciences, University "G. d'Annunzio" of Chieti, Italy
[t] Institute for Advanced Biomedical Technologies (ITAB), University "G. d'Annunzio" of Chieti, Italy
[u] IRCCS SDN, Naples, Italy
[v] University of Naples Parthenope, Naples, Italy
[w] Section of Neurology, Centre for Memory Disturbances, University of Perugia, Perugia, Italy
[x] Medical Physics Unit, Perugia General Hospital, Perugia, Italy
[y] 3rd Department of Neurology, Aristotle University of Thessaloniki, Thessaloniki, Greece
[z] Interbalkan Medical Center of Thessaloniki, Thessaloniki, Greece
[aa] Department of Radiology, Aristotle University of Thessaloniki, Thessaloniki, Greece
[ab] Dept. Geriatrics, Neuroscience & Orthopaedics, Catholic University, Policlinic Gemelli, Rome, Italy
[ac] IRCSS S.Raffaele Pisana, Rome, Italy
[ad] Center for Neuropsychological Research, Catholic University, Rome, Italy
[ae] Department of Neuroradiology, University Hospital Leipzig, Leipzig, Germany
[af] Memory Clinic and LANVIE, Laboratory of Neuroimaging of Aging, University Hospitals and University of Geneva, Geneva, Switzerland
[ag] Department of Psychiatry and Psychotherapy, University of Göttingen, Göttingen, Germany
[ah] Neuroradiology Unit, Perugia General Hospital, Perugia, Italy
[ai] Department of Pharmacology, EA1046, Lille University, Lille, France

* Corresponding author at: Assistant Professor, Center for Mind/Brain Sciences, University of Trento, Italy. Fax: +39 0461 88 3066.
E-mail address: jorge.jovicich@unitn.it (J. Jovicich).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.neuroimage.2014.06.075

ARTICLE INFO

Article history:
Accepted 28 June 2014
Available online 12 July 2014

Keywords:
Brain diffusion tensor imaging
Reproducibility
Reliability
Tract-based spatial statistics
Multi-center
Multi-site MRI

ABSTRACT

Large-scale longitudinal neuroimaging studies with diffusion imaging techniques are necessary to test and validate models of white matter neurophysiological processes that change in time, both in healthy and diseased brains. The predictive power of such longitudinal models will always be limited by the reproducibility of repeated measures acquired during different sessions. At present, there is limited quantitative knowledge about the across-session reproducibility of standard diffusion metrics in 3 T multi-centric studies on subjects in stable conditions, in particular when using tract based spatial statistics and with elderly people. In this study we implemented a multi-site brain diffusion protocol in 10 clinical 3 T MRI sites distributed across 4 countries in Europe (Italy, Germany, France and Greece) using vendor provided sequences from Siemens (Allegra, Trio Tim, Verio, Skyra, Biograph mMR), Philips (Achieva) and GE (HDxt) scanners. We acquired DTI data ($2 \times 2 \times 2$ mm^3, b = 700 s/mm^2, 5 b0 and 30 diffusion weighted volumes) of a group of healthy stable elderly subjects (5 subjects per site) in two separate sessions at least a week apart. For each subject and session four scalar diffusion metrics were considered: fractional anisotropy (FA), mean diffusivity (MD), radial diffusivity (RD) and axial (AD) diffusivity. The diffusion metrics from multiple subjects and sessions at each site were aligned to their common white matter skeleton using tract-based spatial statistics. The reproducibility at each MRI site was examined by looking at group averages of absolute changes relative to the mean (%) on various parameters: i) reproducibility of the signal-to-noise ratio (SNR) of the b0 images in centrum semiovale, ii) full brain test–retest differences of the diffusion metric maps on the white matter skeleton, iii) reproducibility of the diffusion metrics on atlas-based white matter ROIs on the white matter skeleton. Despite the differences of MRI scanner configurations across sites (vendors, models, RF coils and acquisition sequences) we found good and consistent test–retest reproducibility. White matter b0 SNR reproducibility was on average 7 ± 1% with no significant MRI site effects. Whole brain analysis resulted in no significant test–retest differences at any of the sites with any of the DTI metrics. The atlas-based ROI analysis showed that the mean reproducibility errors largely remained in the 2–4% range for FA and AD and 2–6% for MD and RD, averaged across ROIs. Our results show reproducibility values comparable to those reported in studies using a smaller number of MRI scanners, slightly different DTI protocols and mostly younger populations. We therefore show that the acquisition and analysis protocols used are appropriate for multi-site experimental scenarios.

Introduction

Diffusion tensor imaging (DTI) is a quantitative MRI technique widely used for the in vivo characterization of white matter microstructural organization (Ciccarelli et al., 2008; Mori and Zhang, 2006). DTI can be applied to investigate both normal and pathological conditions, and in longitudinal studies it can measure changes of white matter tissue properties in normal aging (Lebel and Beaulieu, 2011; Sullivan and Pfefferbaum, 2007; Sullivan et al., 2010; Westlye et al., 2010) as well as in brain diseases like for example Alzheimer's Disease (Kantarci et al., 2010; Mielke et al., 2009; Scola et al., 2010; Teipel et al., 2010), Huntington's Disease (Magnotta et al., 2009; Sritharan et al., 2010; Weaver et al., 2009), multiple sclerosis (Calabrese et al., 2011; Harrison et al., 2011; Rashid et al., 2008; Sage et al., 2009), stroke recovery (Wang et al., 2006) and traumatic brain injury (Sidaros et al., 2008). Such longitudinal DTI studies can be used to test and develop DTI-based biomarker models of disease progression/recovery, which may be of great utility in better understanding physiopathology as well as for evaluating therapeutic effects.

DTI allows the description of tissue microstructures modeling the Gaussian diffusion properties of water and the detection of white matter lesions (Basser and Pierpaoli, 1996). The most commonly used DTI metrics in clinical studies are fractional anisotropy (FA) and mean diffusivity (MD). Complementary information about white matter structure can be obtained from axial (AD) and radial (RD) diffusivity which, with some limitations, are considered indices of axonal injury and demyelination, respectively (Song et al., 2005; Wheeler-Kingshott and Cercignani, 2009). In addition to these diffusion metrics, orientation information in white matter tracts can be obtained using more advanced DTI acquisition and analysis methods, for example with probabilistic tractography (Behrens et al., 2003; Parker et al., 2003), diffusion spectrum imaging (Wedeen et al., 2005) and high angular resolution methods (Wedeen et al., 2008). These methods, however, typically require longer acquisition times and/or specialized MRI sequences not always available on clinical scanners, and their implementations can therefore be challenging in large multi-centric longitudinal studies, particularly when involving elderly subjects. For these reasons this study focuses on standard DTI acquisitions and their scalar derived metrics (FA, MD, AD, RD).

Longitudinal multi-center MRI studies are becoming an increasingly common strategy to collect large datasets distributing the data acquisition load across multiple partners (Van Horn and Toga, 2009). Moreover, longitudinal studies reduce the between subject variability because each subject is his/her own control. One critical factor that limits the sensitivity to detect changes in any longitudinal study is the reproducibility of repeated measures. Obtaining reproducible quantitative results from DTI data is not trivial given that the final results are sensitive to a large number of acquisition and analysis factors (Jones and Cercignani, 2010). Various aspects of DTI reproducibility have been investigated, including basic reproducibility measures of different populations (Bonekamp et al., 2007; Ciccarelli et al., 2003; Heiervang et al., 2006; Marenco et al., 2006), evaluation of the effects of region of interest (ROI) drawing protocols (Wakana et al., 2007), effects of signal averaging (Farrell et al., 2007), head motion effects (Yendiki et al., 2013), as well as the effects of various acquisition parameters like for example b-value (Bisdas et al., 2008), diffusion weighting scheme (Landman et al., 2007; Vaessen et al., 2010), voxel size (Papinutto et al., 2013), and MRI scanner effects (Brander et al., 2010; Pagani et al., 2010; Pfefferbaum et al., 2003; Vollmar et al., 2010; White et al., 2011; Zhu et al., 2011).

However, despite the wide use of DTI as a tool to assess white matter integrity in 3 T MRI studies, across-session test–retest reliability of diffusion measures on subjects in stable conditions has not been thoroughly investigated using multiple MRI systems. Across-session reproducibility is useful to estimate the effective reproducibility errors that are part of a longitudinal study, since across-session acquisitions include additional sources of variance like MRI system instabilities, differences in head positioning and re-positioning within the RF coil, differences in automated acquisition procedures like auto shimming, as well as potential effects from how different operators follow instructions to execute the same acquisition protocol. These variability sources are negligible in within-session reproducibility studies. Table 1 outlines studies

Table 1
Summary of studies that evaluated within-scanner across-session test–retest reproducibility of 3 T diffusion results on healthy adult subjects. Abbreviations: WB = whole brain, cc_body = body of the corpus callosum, cc_genu = genu of the corpus callosum, cc_spl = splenium of the corpus callosum, crtsp = corticospinal tract, ILF = inferior lateral fasciculus, SLF = superior lateral fasciculus.

Study	3 T MRI scanners (number)	Subjects (age, mean ± SD) across-session days	Reproducibility metrics	ROI
This study	Siemens Allegra (1), Verio (1), Skyra (1), Biograph mMR (1); GE HDxt (2); Philips Achieva (4)	40, 5 sub/scanner (63.2 ± 8.1) 14–31	Rep err, CV of FA, MD, RD, AD from ROIs & TBSS	Atlas based cc_body, cc_genu, cc_spl, crtsp, SFL, ILF WB tracts
Takao et al. (2012)	GE Signa (2)	224 (57 ± 10) 365	FA, RD, AD voxelwise	WB tracts
Huang et al. (2012)	Siemens TIM Trio (2)	6 (24 ± 6) 31	ICC of FA, MD from ROIs	Manual segmentation WB, cc_spl, SFL, ILF
Fox et al. (2012)	Siemens Trio (1) before and after upgrade	2 (35) 365 and 730	CV, Lin's concordance from ROIs	Manual segmentation cc, deep and periventricular WM, deep and cortical GM
Vollmar et al. (2010)	GE Signa (2)	9 (34 ± 8) 1–95	CV, ICC of FA from ROIs	Manual segmentation WB, cc_spl, SFL, ILF
Bisdas et al. (2008)	Philips Intera (1)	12 (34 ± 11) 14	CV, ICC of FA from ROIs	Manual segmentation cc_genu, cc_spl, int, crtsp
Jansen et al. (2007)	Philips Achieva (1)	10 (26 ± 2) 5–25	CV, ICC of FA, ADC from ROIs and voxel-wise	Atlas based frontal, temporal lobes, WB

that, to the best of our knowledge, have reported across-session test–retest reproducibility measures of diffusion data derived from adult healthy volunteers using 3 T systems. These studies are limited to few sites with identical 3 T scanners (Huang et al., 2012; Takao et al., 2012; Vollmar et al., 2010), focused mainly on young subjects (<40 years, except Takao et al., 2012), using DTI analysis mostly based on manual ROIs (Bisdas et al., 2008; Jansen et al., 2007) or aimed at evaluating MRI software upgrade effects (Fox et al., 2012). In other words, the impact of across-session reproducibility errors of DTI metrics derived from multi-site longitudinal 3 T studies is not clearly defined, in particular with the commonly used tract-based spatial statistics (TBSS) analysis (Smith et al., 2006). TBSS is particularly attractive for longitudinal voxel-wise analysis of DTI data given that individual diffusion parameter maps are projected onto a group-wise skeleton constructed from FA data to account for residual misalignments among individual white matter tracts in multiple measures and multiple subjects. These issues are relevant to the PharmaCog project, a new industry-academic European project aimed at identifying biomarkers sensitive to symptomatic and disease modifying effects of drugs for Alzheimer's disease (http://www.alzheimer-europe.org/FR/Research/PharmaCog).

The aims of the present study were the following: i) to implement a multi-site 3 T MRI data acquisition protocol for diffusion analysis (10 different MRI sites covering three common clinical MRI vendors in Europe), ii) to acquire across-session test–retest data (2 acquisitions at least one week apart) from a population of healthy stable elderly subjects (5 subjects per MRI site), and iii) to evaluate and compare the across-session reproducibility of FA, MD, AD and RD diffusivities within and across MRI sites using both voxel-based TBSS and an atlas-based ROI analyses. This multi-site DTI study is unique in that it characterizes brain diffusion reproducibility metrics particularly relevant to multi-center longitudinal studies of brain disease (within-site across-session reproducibility in healthy elderly subject) derived from both full brain voxel-based TBSS and atlas-based ROIs using a wide range of clinical 3 T scanners (Table 2). The test–retest raw DTI data from the 10 MRI sites is made publicly available (100 brain volumes).

Materials and methods

Several aspects of the subjects, study design and data preparation steps used in this diffusion study were already described in a recent morphometry study (Jovicich et al., 2013) but are here repeated for completeness and with the appropriate modifications.

Subjects

Twelve clinical sites (10 MRI sites) participated in this study across Italy (Brescia, Verona, Genoa, Rome, Chieti, Perugia and Naples), France (Marseille, Lille, and Toulouse), Germany (Essen) and Greece (Thessaloniki). The Brescia site was responsible for the coordination and analysis of the whole study and did not acquire MRI data. Each MRI site recruited 5 local volunteers within an age range of 50–80 years. The subject's age range corresponds to that of the clinical population that will be studied with the protocols tested in this reproducibility study. Each subject underwent two MRI sessions completed at least 7 days (but no more than 60 days) apart at the site. This short period between the test and retest sessions was chosen to minimize biological changes that could affect the reliability of the measures and to mimic the variability expected from separate sessions, as measured in longitudinal studies. Table 2 summarizes information about age, gender and test–retest interval times of the subjects recruited at each site. All participants were volunteers with no history of major psychiatric, neurological or cognitive impairment (referred to as healthy in this study), and provided written informed consent in accordance with the "classification" of the study in regard to the national regulations and laws in the different participating countries. An Appendix lists the participants' selection and exclusion criteria. Subjects were considered to be in stable conditions between the two MRI acquisition sessions. In France, the study received an authorization from the national drug regulatory agency (Agence Nationale de Sécurité du Médicament et des produits de santé) and an approval from the Comité de Protection des Personnes Sud-Méditerranée 1 (Marseille), for the three French sites (Marseille, Lille, and Toulouse). In Germany, Italy and Greece the study obtained authorization from one Ethics Committee relevant to each institution: Essen (Ethik-Kommission des Universitätsklinikums Essen), Verona (Comitato Etico Istituzioni Ospedaliere Cattoliche, CEIOC), Genoa (Comitato Etico IRCSS-Azienda Ospedaliera Universitaria San Martino-IST), Rome (Comitato Etico dell'Università Cattolica del Sacro Cuore, Facoltà di Medicina e Chirurgia "Agostino Gemelli"), Naples (Comitato Etico per la Sperimentazione clinica dell' IRCSS Fondazione SDN per la Ricerca e l'Alta Formazione in Diagnostica Nucleare), Chieti (Comitato di Etica per la Ricerca Biomedica dell'Università degli Studi "G. d'Annunzio" e dell'Azienda Sanitaria Locale di Lanciano-Vasto-Chieti), Perugia (Comitato Etico Aziende Sanitarie Umbria) and Thessaloniki (Greek association of Alzheimer's Disease and related disorders). All subjects signed informed consent.

MRI acquisitions

The ten 3 T MRI sites that participated in this study used different MRI system vendors and models (Siemens, GE, and Philips). Table 2 summarizes the main MRI system and DTI acquisition differences across sites. Only vendor-provided sequences were used. Each MRI scanning session consisted of several acquisitions, including: anatomical T2*, anatomical FLAIR, resting state fMRI, B0 map, DTI and two anatomical T1 scans, with a total acquisition time of approximately 40 min. The DTI acquisition for each subject was always the last acquisition of the session, regardless of MRI site.

Table 2
Summary of demographic, MRI system and DTI acquisition differences across MRI sites.

	Site 1 Verona	Site 2 Marseille	Site 3 Lille	Site 4 Toulouse	Site 5 Genoa	Site 6 Essen	Site 7 Chieti	Site 8 Naples	Site 9 Perugia	Site 10 Thessaloniki
Subjects' age: mean ± SD (range) years	67.8 ± 9.9 (26)	66.0 ± 8.3 (20)	64.2 ± 5.3 (13)	59.2 ± 4.5 (12)	58.2 ± 2.2 (5)	52.4 ± 1.5 (3)	68.8 ± 4.3 (11)	59.0 ± 3.5 (9)	60.8 ± 10.3(24)	56.6 ± 5.5 (11)
Test–retest interval (days)	28 ± 23	23 ± 22	15 ± 11	14 ± 10	24 ± 17	11 ± 5	11 ± 5	19 ± 15	10 ± 4	31 ± 8
Gender. (females/N)	2/5 (40%)	4/5 (80%)	3/5 (60%)	3/5 (60%)	2/5 (40%)	2/5 (40%)	5/5 (100%)	2/5 (40%)	3/5 (60%)	3/5 (60%)
3 T MRI scanner	Siemens Allegra	Siemens Verio	Philips Achieva	Philips Achieva	GE HDxt	Siemens Skyra	Philips Achieva	Siemens Biograph mMR	Philips Achieva	GE HDxt
MR system software version	VA25A	B17	3.2.2	3.2.2	15 M4A	D11	3.2.1	B18P	3.2.2	15 M4A
DTI sequence	Double refocused spin-echo	Double refocused spin-echo	Standard Stejskal–Tanner	Standard Stejskal–Tanner	Double refocused spin-echo	Double refocused spin-echo	Standard Stejskal–Tanner	Double refocused spin-echo	Standard Stejskal–Tanner	Double refocused spin-echo
TX/RX coil	Birdcage	Body/12-chan.	Body/8-chan.	Body/8-chan.	Body/8-chan.	Body/20-chan.	Body/8-chan.	Body/12-chan.	Body/8-chan.	Body/8-chan.
Parallel imaging: acceleration	None	GRAPPA 2	SENSE 1.5	SENSE 1.5	ASSET 2	GRAPPA 2	SENSE 2	GRAPPA 2	SENSE 2	ASSET 2
TR (ms)	9300	9300	8000	8000	14,974	9100	6564	9300	9517	15,150
TE (ms. shortest)	87	84	75	75	81	78	70	84	55	80.1
Acquisition (min:s)	5:53	6:04	5:27	5:27	8:59	5:23	5:59	7:54	10:02	10:12

Each subject had a total of two diffusion acquisitions, one from the test session and one from the retest session at least a week apart. In each session the following DTI acquisition parameters were common across sites: b-value = 700 s/mm^2, 5 b0 volumes, 30 gradient directions, voxel size 2 × 2 × 2 mm^3, 64 axial slices, zero slice gap. Siemens and GE scanners allowed saving the 5 b0 volumes separately, whereas the Philips scanners saved only one volume corresponding to the average of the 5 b0 acquisitions.

Other acquisition parameters including head RF coil, pulse sequence, TE (minimum value), TR, parallel acquisition and fat suppression methods were difficult to standardize due to the restrictions from each site's implementations. The choices for these parameters were made based on the optimal or possible options available at the different platforms (see Table 2). All images from multi-channel coils were reconstructed by the scanner as the sum of the squares across channels. When allowed by the MRI system, images were reconstructed and saved without additional filtering options that could differ across scanners introducing different degrees of smoothing.

Data preparation

Imaging data were initially anonymized at each site by replacing the subject name with a unique identifier using the free DicomBrowser tool (http://hg.xnat.org/dicombrowser). Anonymized dicom data were then compressed and uploaded onto a data sharing system accessible to all member sites, from where they were subsequently downloaded for analysis at central site (Brescia).

Downloaded anonymized dicom data were converted to nifti format using the free dcm2nii software (http://www.mccauslandcenter.sc.edu/mricro/mricron/dcm2nii.html, output format FSL — 4D NIFTI nii) from which the original dicom converted to nifty files were utilized.

DTIPrep was used for quality control, head motion and eddy current correction of the data prior to the estimation of diffusion metrics (Liu et al., 2010; Magnotta et al., 2012). DTIPrep is a public available software tool that performs automated quality control removing image volumes containing typical artifacts in DTI data, including signal drop out artifacts from head motion (http://www.nitrc.org/projects/dtiprep). The output of DTIPrep, here used with its default recommended settings, can be passed onto standard DTI analysis tools, having a single co-registered averaged b0 volume and updated information about the diffusion weighted image volumes kept for the tensor estimation.

White matter lesions evaluation

White matter lesions (WMLs) were assessed with the rating scale Age-Related White Matter Changes (ARWMC) on FLAIR MR images (Wahlund et al., 2001). WMLs were evaluated by a single rater separately in the right and left hemispheres. Scores of 0, 1, 2 and 3 were assigned in frontal, parieto-occipital, temporal and infratentorial areas to: no WMLs, focal lesions, beginning confluence of lesion, and diffuse involvement of the entire region, respectively. Scores of 0, 1, 2 and 3 were assigned in basal ganglia to: no WMLs, 1 focal lesion, more than 1 focal lesion, and confluent lesions, respectively. For each area, a score was given by the sum of the right and left subscores, and the total WML score was computed as the sum of all subscores. Intra-rater reliability was assessed on a random sample of 20 subjects. The intraclass correlation coefficient was 0.99, indicating a very high reliability.

Estimation of diffusion metrics

After the DTIPrep quality control the rest of the analysis was carried out using standard FSL's Diffusion Toolbox tools (FSL4; http://www.fmrib.ox.ac.uk/fsl). Following skull and non-brain tissue removal (BET), each diffusion-weighted series was used to compute the six independent components of the diffusion tensor. The three eigenvalues (λ_1, λ_2, λ_3, in decreasing magnitude order) and eigenvectors of the resulting

tensor map were derived by matrix diagonalization (Basser and Pierpaoli, 1996). The maps for the following four diffusion metrics were obtained with FSL's Diffusion Toolbox using standard definitions:

$$MD = (\lambda_1 + \lambda_2 + \lambda_3)/3$$

$$FA = \sqrt{\frac{1}{2}}\sqrt{\frac{(\lambda_1-\lambda_2)^2 + (\lambda_2-\lambda_3)^2 + (\lambda_1-\lambda_3)^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}$$

$$AD = \lambda_1$$
$$RD = (\lambda_2 + \lambda_3)/2$$

TBSS and atlas-based ROI analyses

We investigated test–retest reproducibility/variability using both full-brain voxel-based and ROI based analyses. The full-brain TBSS analysis (Smith et al., 2006) was computed at each MRI site separately to create a FA template by registration of the individual FA images, test and retest, to the most representative of the site followed by FA map projection onto the common reference space of the MNI152. All non-linear registrations were done with the default nonlinear registration tool FNIRT (Andersson et al., 2007a,b), which use a b-spline representation of the registration warp field (Rueckert et al., 1999). The FMRIB58_FA atlas was not chosen as a common template for all sites because this atlas has been created with subjects aged 20–50 years, which may have differences relative to the population here studied (age range 50–80). The mean of all FA maps in a site was then computed and used to create a mean FA skeleton representing the common tracts. The skeleton of each MRI site was then thresholded with an FA value of 0.2 to exclude the non skeleton voxels and each individual FA map was projected into it, one for the test and one for the retest session. Finally, the non-linear warps and the skeleton projection achieved with FA images were applied to the MD, AD and RD maps to bring them into standard space at each MRI site. Group pair-wise analyses of test–retest differences of these maps allowed evaluating the spatial distribution of significant differences at each site.

A complementary atlas-based ROI analysis was done to estimate the averaged reproducibility errors within areas of interest. The ROI analysis was focused on the corpus callosum (body, genu and splenium), corticospinal tract (left/right), inferior and superior lateral fasciculi (left/right), which are of relevance in MCI and AD studies (Drago et al., 2011). These white matter ROIs are pre-defined in the JHU-ICBM-FA-1mm atlas and were back-projected with a non-linear co-registration to each subject's test and retest session FA maps in the MNI152 space obtained from TBSS. In each MRI site, each ROI label was overlapped with the corresponding site's FA TBSS skeleton space to remove any CSF and gray matter voxels. These new ROIs were then used in each subject's FA, MD, AD and RD maps (MNI152 space) to compute the absolute percent test–retest variability of the various metrics.

General evaluation of image quality across sites: SNR and head motion

Signal-to-noise ratio (SNR) estimations were used for the general assessment of image quality across the MRI sites. Quantitative estimation of numerical SNR values is challenging for DTI because of the different signal properties of the b0 and diffusion weighted images. Typically the b0 images are used and there are several possibilities for computing from them the noise (Laganà et al., 2010). In this study we used the motion corrected, co-registered and averaged b0 volume output by DTIPrep for each subject. The centrum semiovale was chosen as a representative white matter ROI for SNR assessment (Farrell et al., 2007). This ROI was pre-defined in the JHU-ICBM-T2-1mm atlas and back-projected with a non-linear co-registration to each subject's test and retest session averaged b0 volume. For each site and subject we computed the SNR as the mean signal divided by the standard deviation of the signal within the ROI (Dietrich et al., 2007). The global SNR mean for each site was the SNR average across subjects including test and retest sessions. The test–retest reproducibility of the SNR was estimated by averaging across subjects, in each site, the percent absolute SNR differences between sessions relative to the mean SNR across sessions.

Head motion during the DTI acquisition may introduce biases in the derived metrics so it is important to report it (Alhamud et al., 2012; Ling et al., 2012; Yendiki et al., 2013). The rotation and translation parameters from each DTI acquisition were obtained using FSL's linear registration tool FLIRT of each brain volume to the averaged b0 volume (Jenkinson et al., 2002). To characterize potential head motion differences across sites we evaluated the means and the maximum/minimum motion correction parameters (rotations and translations), separately for the test and retest sessions at each site.

Evaluation of reliability of DTI metrics in atlas-based ROIs

One of the goals of this study was to quantify the across-session test–retest reliability of diffusion parameters obtained from a multi-site acquisition protocol made as similar as possible across sites. To evaluate the reliability of the brain diffusion metrics we analyzed their variability, or reproducibility error, across the test–retest sessions for each ROI and MRI site. We considered the following measures of variability: absolute error relative to the mean and coefficient of variance. These metrics were chosen because at least one of them tends to be reported in reproducibility studies (Table 1), thus allowing for wider comparisons across studies.

The absolute error relative to the mean (ε) was computed for each subject, ROI and diffusion metric (DM) as follows:

$$\varepsilon_{DM} = 100 \times \frac{|DM_{retest} - DM_{test}|}{(DM_{retest} + DM_{test})/2}$$

where DM refers to any of the four scalar metrics evaluated in this study: FA, MD, AD and RD. The group error for every MRI site and ROI was then averaged across subjects. The measure was chosen because it is intuitive and because the estimation of the means is more robust than the estimation of the variance from the signed differences, in particular for low number of subjects.

The coefficient of variation (CV) was computed for each subject, ROI and diffusion metric as the ratio between the standard deviation and the mean of the test–retest diffusion index. The group CV for every MRI site and ROI was then averaged across subjects.

Statistical analysis

The following statistical analyses tests were done, using SPSS (v.13.0):

- To test for MRI site effects on subject's age, averaged b0 SNR, mean FA, MD, AD, RD and their corresponding across-session reproducibility errors, we used one-way Kruskall–Wallis tests (non-parametric version of ANOVA) with MRI site as factor. Significance threshold was set at $p < 0.05$.
- To test for head rotation and translation differences between test and retest sessions, for each site, the two-tailed Wilcoxon test was used (non-parametric version of the paired Student's t-test).
- To test for voxel-wise test–retest TBSS group differences at each site, two contrast (positive and negative) paired t-tests were used with 5000 random permutations. Significance threshold was set to $p < 0.05$ corrected for multiple comparisons using threshold-free cluster enhancement.

Table 3
Summary of automated DTIPRep quality control results. Number of gradients eliminated at each MRI site with corresponding data percentage.

MRI sites	Slice-wise check	Venetian blind check
Site 1	11 (3.1%)	0 (0%)
Site 2	8 (2.2%)	1 (0.3%)
Site 3	17 (5.5%)	0 (0%)
Site 4	17 (5.5%)	0 (0%)
Site 5	23 (6.6%)	0 (0%)
Site 6	2 (0.6%)	0 (0%)
Site 7	8 (2.3%)	1 (0.3%)
Site 8	3 (0.9%)	0 (0%)
Site 9	17 (4.6%)	0 (0%)
Site 10	14 (5.0%)	8 (2.9%)

Results

In this multi-site 3 T study (10 clinical scanners from different vendors) we estimated the test–retest reliability (within-site, across two separate sessions at least a week apart) of diffusion measures derived from healthy elderly DTI data (FA, MD, radial and axial diffusivities). Both full brain (TBSS) and ROI (atlas-based) approaches were used.

The 50 subjects enrolled (Table 2) resulted to be all Caucasian with similar age distributions except for site 6 (younger group, mean age 52.4 ± 1.5 years, significantly different from sites 1 and 7, Kruskall–Wallis, $p < 0.05$). There were no age distribution differences between the other MRI sites. The time interval between test and retest scans ranged from 7 to a maximum of 55 days, with a mean and standard deviation of 17 ± 13 days. A Kruskal–Wallis test showed that there were no significant MRI site effects of the scan–rescan time intervals ($p = 0.174$). The overall mean ARWMC score for white matter lesions was 3.1 ± 3.0, without significant MR site effects (Kruskall–wallis, $p = 0.68$).

General image data quality assurance: DTIPrep, SNR and head motion

The overall acquired DTI dataset consisted of 100 brains: 10 MRI sites, 5 subjects per site, 2 acquisitions per subject (one test, one retest) obtained on different dates. A quick visual inspection was done on all data to ensure that there were no gross partial brain coverage errors. The dataset was run through DTIPrep for further automatic quality assurance (Liu et al., 2010; Magnotta et al., 2012). Table 3 shows the DTIPrep summary results of the number of volumes eliminated at each site due to intensity-related artifacts (slice-wise check) and motion-related artifacts (venetian blind check). The number of volumes discarded are also expressed as percentage of the total number of 3D volumes per site (each b0 and diffusion weighted image was considered a volume). On average DTIPrep eliminated 3.4% 3D volumes per site with the slice-wise check (range of: 0.6–6.6% across sites). The venetian blind artifact check discarded no data in 7 sites and removed very few gradient directions in the other 3 sites (1.3% per site in average, with a range of 0.3–2.9% across sites). Visual inspection of the resulting FA maps showed a high qualitative similarity across sites (Fig. 1).

Signal-to-noise ratio within the centrum semiovale was evaluated on the averaged b0 volume of each subject to evaluate SNR and test–retest SNR reproducibility at each site (Fig. 2). The mean ROI size used for the SNR analysis was (2.5 ± 0.4) cm^3 across the group of subjects (Fig. 2A). Significant MRI site effects were found for SNR (Fig. 2B, Kruskall–Wallis test, $p < 0.05$). However, the test–retest reproducibility errors of the SNR showed no significant MRI site effects (Fig. 2C, Kruskall–Wallis test, $p > 0.05$), with an average value of (7 ± 1) % across sites.

Head motion summary results by site and by acquisition session are shown in Table 4. Average, maximum and minimum values are shown for volume-by-volume rotations and translations during data acquisition. No significant MRI site effects were observed for the mean rotation and mean translation parameters (Kruskall–Wallis test, $p > 0.05$). None of the MRI sites showed significant differences between test and retest sessions for head rotations and translations (two tailed Wilcoxon test, $p > 0.05$). The maximum rotation was observed in MRI site 9 (2.44°) and the maximum translation in MRI site 10 (2.69 mm). These two sites are the ones with the longer DTI acquisition times in the group (about 10 min, Table 2).

Full brain voxel-based results on the white matter TBSS skeleton

Voxel-based test–retest FA reproducibility maps were computed at each site on the white matter TBSS skeleton. Fig. 3 shows the reproducibility maps overlaid on the FA map of a representative slice and subject for each MRI site. No significant group differences were found between test and retest at any of the sites (t-test, $p > 0.05$). The color coded maps show that the FA variability pattern is similar across sites, with reproducibility errors mostly <10% and increasing as the white matter skeleton approaches the cortex. Similar results were found for the other scalar metrics (MD, AD and RD). To better characterize test–retest variability in specific brain areas we performed an ROI analysis.

Atlas-based ROI results

Fig. 4 shows an example of the atlas-based ROIs co-registered to a sample volunteer. Fig. 4 also lists, in decreasing size order, the average

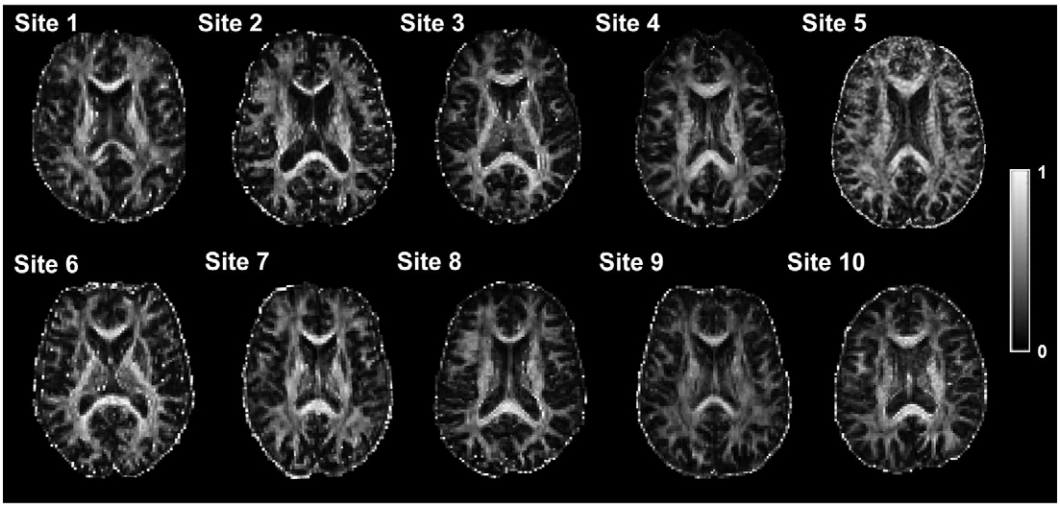

Fig. 1. Sample single-subject FA maps across different 3 T MRI sites for qualitative comparison. See Table 2 for MRI sites characteristics.

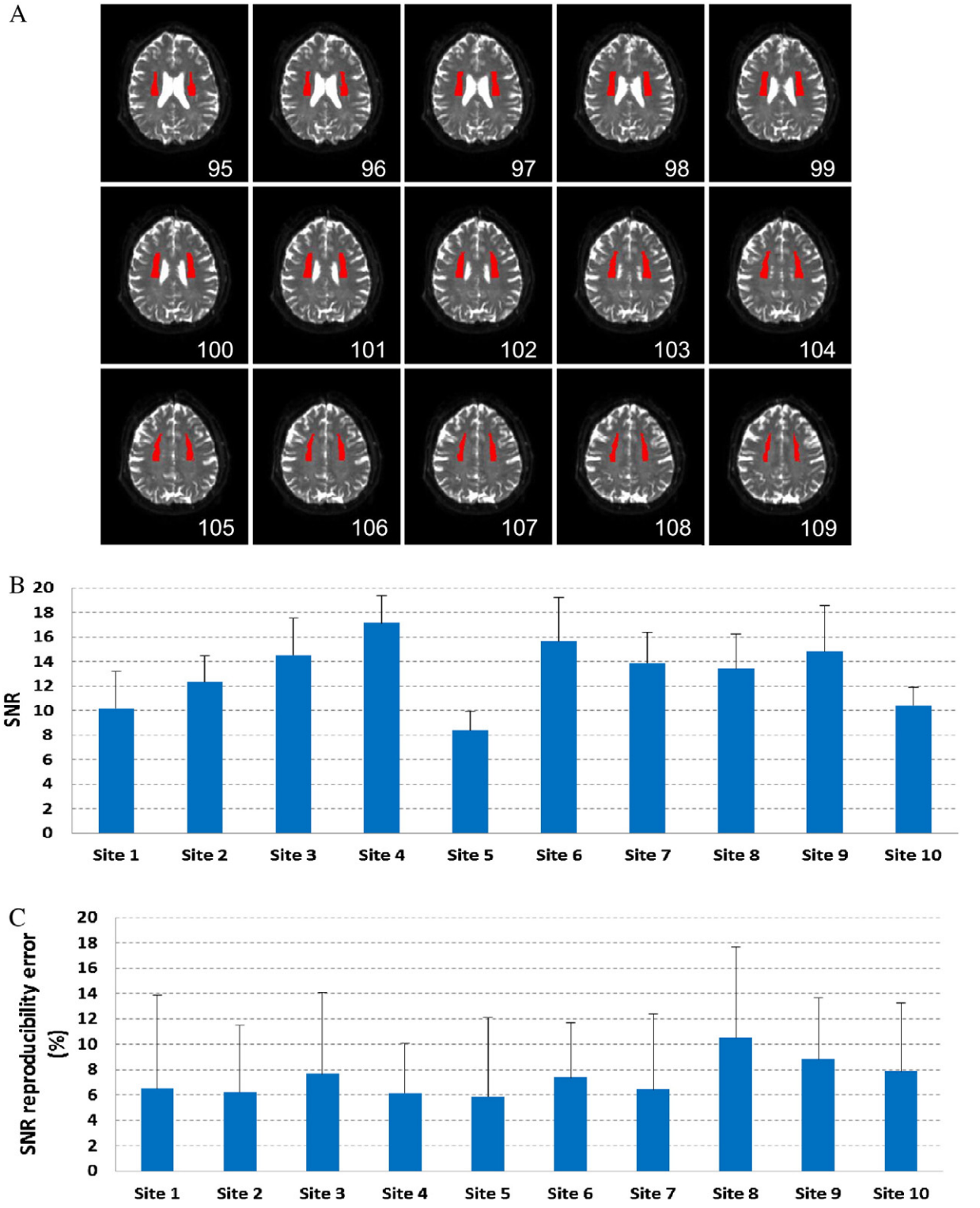

Fig. 2. SNR comparison across sites from b0 volume. Atlas-based white matter ROI (centrum semiovale, red) superimposed on the mean b0 image of a sample volunteer for automatic quantification of SNR (A). The JHU-ICBM-T2-1mm atlas coordinates (Z) are displayed in the lower right corner. Within-site group SNR mean and standard deviation from the b0 averaged volume ($n = 5$) (B) and its test–retest percent variability relative to the mean (C). A significant MRI site effect was evident for SNR (Kruskall–Wallis test, $p < 0.05$) but not for SNR reproducibility (Kruskall–Wallis test, $p > 0.05$). The means and standard deviations are computed across subjects, hemisphere and sessions (test–retest). See Table 2 for MRI sites characteristics.

ROI size for each structure when grouping all subjects across sites. The ROIs are shown co-registered to the averaged b0 volume. Good qualitative agreement was found between the ROI definition on the original atlas and the co-registration to each subject, across all sites, also between test and retest sessions (Fig. 5).

Fig. 6 shows the group average diffusion metrics (FA, MD, RD and AD) for each ROI at each site, averaging the test and retest sessions. We found that for every ROI and diffusion metrics there were significant site effects (Kruskall–Wallis, $p < 0.05$).

Table 4
Head motion estimates across MRI sites and acquisition sessions. Rotations and translations are characterized by their average volume-by-volume values across all volumes in the acquisition as well as by their maximum and minimum values.

		DTI motion parameters									
		Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7	Site 8	Site 9	Site 10
Rotation (°): mean ± SD	Test	0.17 ± 0.15	0.22 ± 0.16	0.12 ± 0.11	0.11 ± 0.16	0.15 ± 0.12	0.27 ± 0.36	0.20 ± 0.16	0.14 ± 0.12	0.19 ± 0.15	0.22 ± 0.20
	Retest	0.16 ± 0.14	0.23 ± 0.20	0.16 ± 0.17	0.10 ± 0.11	0.17 ± 0.17	0.16 ± 0.11	0.20 ± 0.19	0.18 ± 0.22	0.30 ± 0.32	0.20 ± 0.24
Translation (mm): mean ± SD	Test	0.41 ± 0.42	0.25 ± 0.17	0.17 ± 0.17	0.23 ± 0.23	0.21 ± 0.18	0.36 ± 0.46	0.35 ± 0.24	0.39 ± 0.30	0.20 ± 0.16	0.35 ± 0.44
	Retest	0.28 ± 0.19	0.35 ± 0.18	0.20 ± 0.15	0.20 ± 0.25	0.10 ± 0.09	0.18 ± 0.14	0.24 ± 0.18	0.25 ± 0.20	0.25 ± 0.22	0.21 ± 0.22
Max rotation (°)	Test	0.78	0.82	1.40	0.93	0.62	0.76	0.83	0.86	0.92	1.35
	Retest	0.90	1.02	0.91	0.69	1.40	1.71	0.96	1.06	2.44	1.98
Max translation (mm)	Test	1.13	0.81	0.87	1.17	0.86	1.00	1.01	0.91	1.09	1.30
	Retest	0.91	0.91	0.87	1.27	0.48	2.22	1.46	0.83	0.78	2.69

Fig. 7 shows the group test–retest absolute % reproducibility errors for each of the diffusion metrics and ROIs (five measures per site, ROI and DTI metrics). Right and left hemisphere results were similar and therefore averaged for the corticospinal tracts, the superior and inferior lateral fasciculi. Consistently with the voxel-based analysis we found that overall the test–retest reproducibility errors were in general below 10%, with the smaller ROIs (inferior lateral fasciculi and cortical spinal tracts) having a tendency for higher reproducibility errors. Significant MRI site effects were found only for MD in the superior lateral fasciculi (Kruskall–Wallis test, $p < 0.05$). The sites with the peak head motion estimates (9 and 10) tended to give higher reproducibility errors relative to the other sites, especially for MD and RD estimates in corpus callosum and superior/inferior lateral fasciculi. Removing sites 9 and 10 from the site-effect analysis gave no significant effects in the reproducibility, regardless of DTI metric or ROI. Similar findings were obtained when evaluating the coefficient of variance (Supplementary Fig. S1).

Fig. 8 shows the reproducibility errors grouping all ROIs for each site (45 estimates per site given that there are 5 subjects and 9 ROIs per site). Errors are mostly within the range of 2–6%. Significant MRI site effects were found for RD and MD (Kruskall–Wallis, $p = 0.005$ and $p = 0.025$, respectively). These effects were largely driven by sites 9 and 10 since they became non-significant when these sites were eliminated from the site-effect analysis. It was of interest to investigate which diffusion metrics offered the lowest and highest reproducibility errors. Paired t-tests between the reproducibility of the various diffusion metrics (paired by site) showed that FA and AD were the metrics with the lowest reproducibility errors (in average across sites 2.8 ± 0.1% and 3.0 ± 0.4%, respectively, not significantly different between themselves, $p < 0.05$). MD followed with a significantly higher reproducibility error relative to FA and AD (3.2 ± 0.4%) and the highest errors were observed on RD reproducibility (4.5 ± 1.1%, significantly higher errors than all the other diffusion metrics, $p < 0.05$). Removing sites 9 and 10 from these evaluations, given that they are the ones with the highest sensitivity to head motion effects, shows that RD remains as the only metric with significantly higher reproducibility errors relative to those from any of the others metrics, which show no significant differences in reproducibility between themselves (FA, AD and MD).

Given that there was a wide range of scan–rescan intervals (7–55 days) we investigated if this affected reproducibility errors. A Pearson's correlation analysis between scan–rescan time interval of each subject and their corresponding reproducibility errors of FA, MD, RD and AD showed that the correlations were all positive, very small (<0.15) and non-significant ($p > 0.05$).

To further investigate the effects of our limited sample size per site on random reproducibility errors we performed two additional analyses. In one analysis we increased the number of subjects by pooling subjects across common MRI system vendors (Siemens: 20 subjects, Philips 20: subjects, GE: 10 subjects) and tested for vendor differences in diffusion metric reproducibility errors. A Kruskal–Wallis test across vendors showed that there were no significant vendor effects for any of the error metrics (Table 5). The other analysis consisted on estimating the Pearson's correlations between each subject's white matter SNR (measured in the centrum semiovale of the averaged 5 b0 volumes) and the subjects' corresponding reproducibility errors for the various metrics. Correlating reproducibility errors with SNR may be considered as grouping data by SNR regardless of site. We found no significant Pearson's correlation for any of the metrics ($p > 0.05$).

Discussion

In this Pharmacog Consortium study, we found that the test–retest reliability/variability of DTI metrics estimated with TBSS in a 3 T

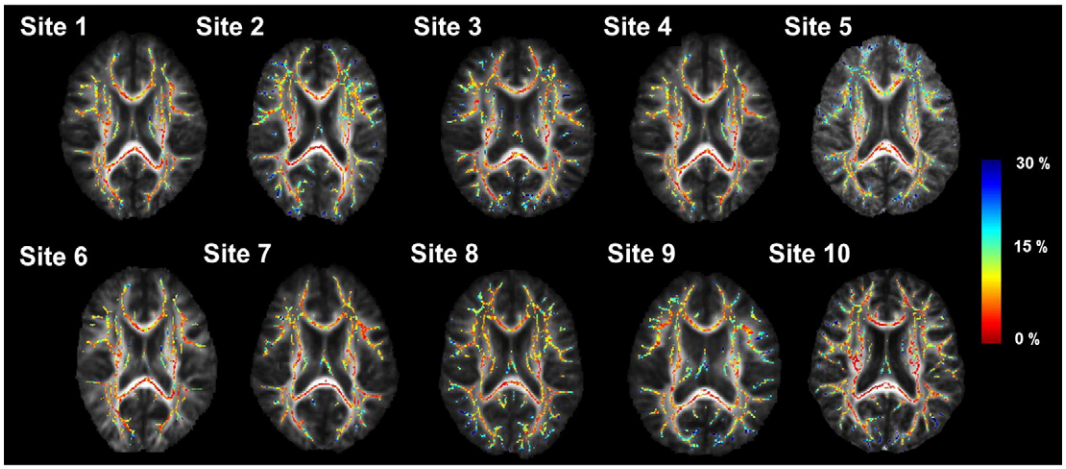

Fig. 3. Voxel-based tract-based spatial statistical (TBSS) analysis across test–retest acquisitions. Within-site group reproducibility error (across subjects) of FA computed on each site-specific TBSS skeleton and overlaid on a representative mean FA slice. Color scale shows the magnitude of the across-session reproducibility error map from 0% (red) to 30% (blue). See Table 2 for MRI sites characteristics.

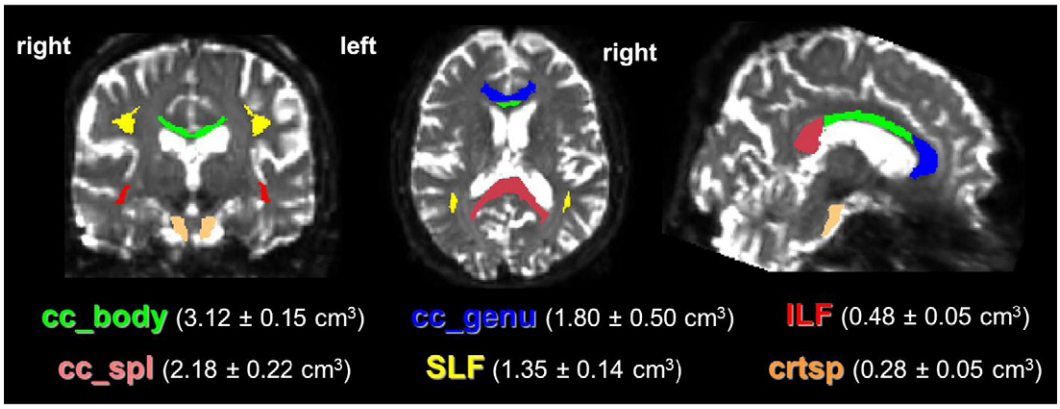

Fig. 4. Atlas-based ROIs co-registered to the b0 image of a sample volunteer for automatic quantification of diffusion metrics. Atlas defined ROIs back-projected onto the subject space were intersected with the TBSS skeleton. For each structure the mean ROI volume across all sites, subjects and sessions is shown. Abbreviations: cc_body = body of the corpus callosum, cc_genu = genu of the corpus callosum, cc_spl = splenium of the corpus callosum, crtsp = corticospinal tract, ILF = inferior lateral fasciculus, SLF = superior lateral fasciculus.

consortium using vendor provided sequences is consistent across sites despite the heterogeneity of MRI scanner configurations. This suggests that pooling DTI data from longitudinal multi-site studies has a potential for accelerating the evaluation of biomarkers related to water diffusion changes. We had three main findings: (1) in a group of healthy stable elderly subjects, the across-session test–retest reproducibility errors were largely consistent across 10 MRI acquisition sites and in average within the range of 2–6% for all diffusion metrics, despite substantial differences across MRI systems used; (2) overall the most reproducible DTI indicators were fractional anisotropy (FA) and axial diffusivity (AD), followed by mean (MD) and finally radial diffusivity (RD); (3) increased head motion sensitivity in the sites with longest acquisitions was consistent with the measurements of the largest motion effects and a tendency for higher reproducibility errors in DTI metrics.

The elderly population studied had low white matter lesion scores that were both consistent across sites and with low vascular burden, with values within norms for healthy elderly populations (Galluzzi et al., 2009). In agreement with previous single-site 3 T studies, we found that FA was one of the most reliable diffusion parameters (Fox et al., 2012; Vollmar et al., 2010). We extended this observation by finding that in our 3 T consortium (10 acquisition sites, 3 vendors and 6 different models) the most reliable white matter diffusion metrics were FA and AD (2.8% and 3.0%, ROI averaged reliability errors respectively, not significantly different). The reproducibility error of FA within the corpus callosum, across sites, was 1.7 ± 0.7%, which is consistent with previous reports from single and 2-site studies (Bisdas et al., 2008; Vollmar et al., 2010). Also CV in the corpus callosum was in good agreement with previous 3 T study (Bisdas et al., 2008; Vollmar et al., 2010) and lower relative to 1.5 T studies (Bonekamp et al., 2007; Ciccarelli et al., 2003). Besides the corpus callosum, the CV of the other structures, excluding MD, AD and RD of corticospinal tracts, were below 10% which is desirable for biological variables related to imaging (Marenco et al., 2006). Consistent with a previous study we also found that test–retest reproducibility gets worse as ROI size gets smaller (Vollmar et al., 2010). This variability may be partially related to imperfect registration during the DTI analysis. Finally, comparisons among the four diffusion indices showed that the highest reproducibility errors were found in average in RD followed then by MD.

In an attempt to characterize sources of noise, recent studies have highlighted the importance of head motion effects during the acquisition of DTI data for diffusion metrics (Alhamud et al., 2012; Ling et al., 2012), in particular heterogeneous increases of test–retest reproducibility errors related to higher head motion in healthy children (Yendiki et al., 2013). Our study supports and extends these results by showing

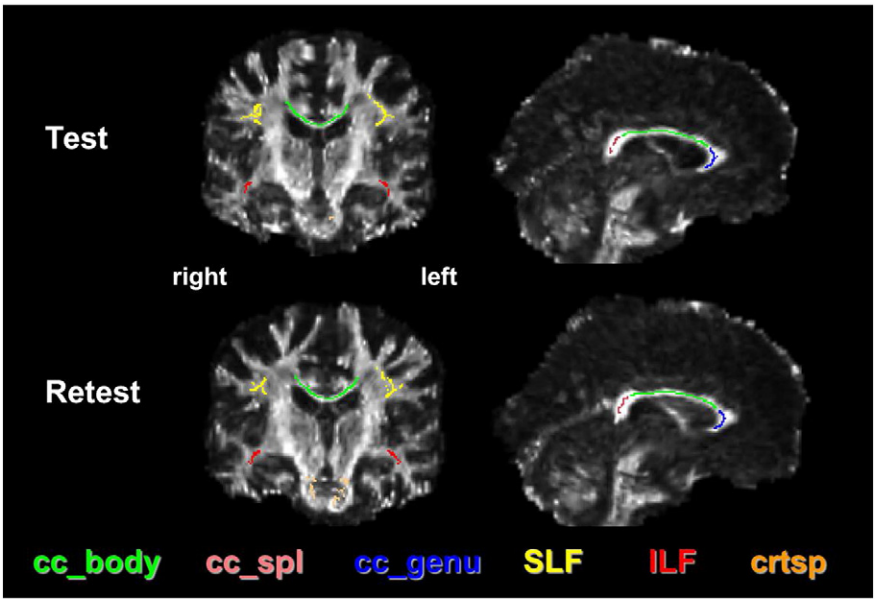

Fig. 5. Example of anatomical correspondence between atlas-based ROI and single subject's FA map derived from site-specific TBSS analysis. Corpus callosum body (cc-body), splenium (cc-spl) and genu (cc-genu) ROIs with superior/inferior lateral fasciculi (SLF/ILF) and corticospinal tract (crtsp) on FA maps from test and retest sessions (sample subject from Site 2).

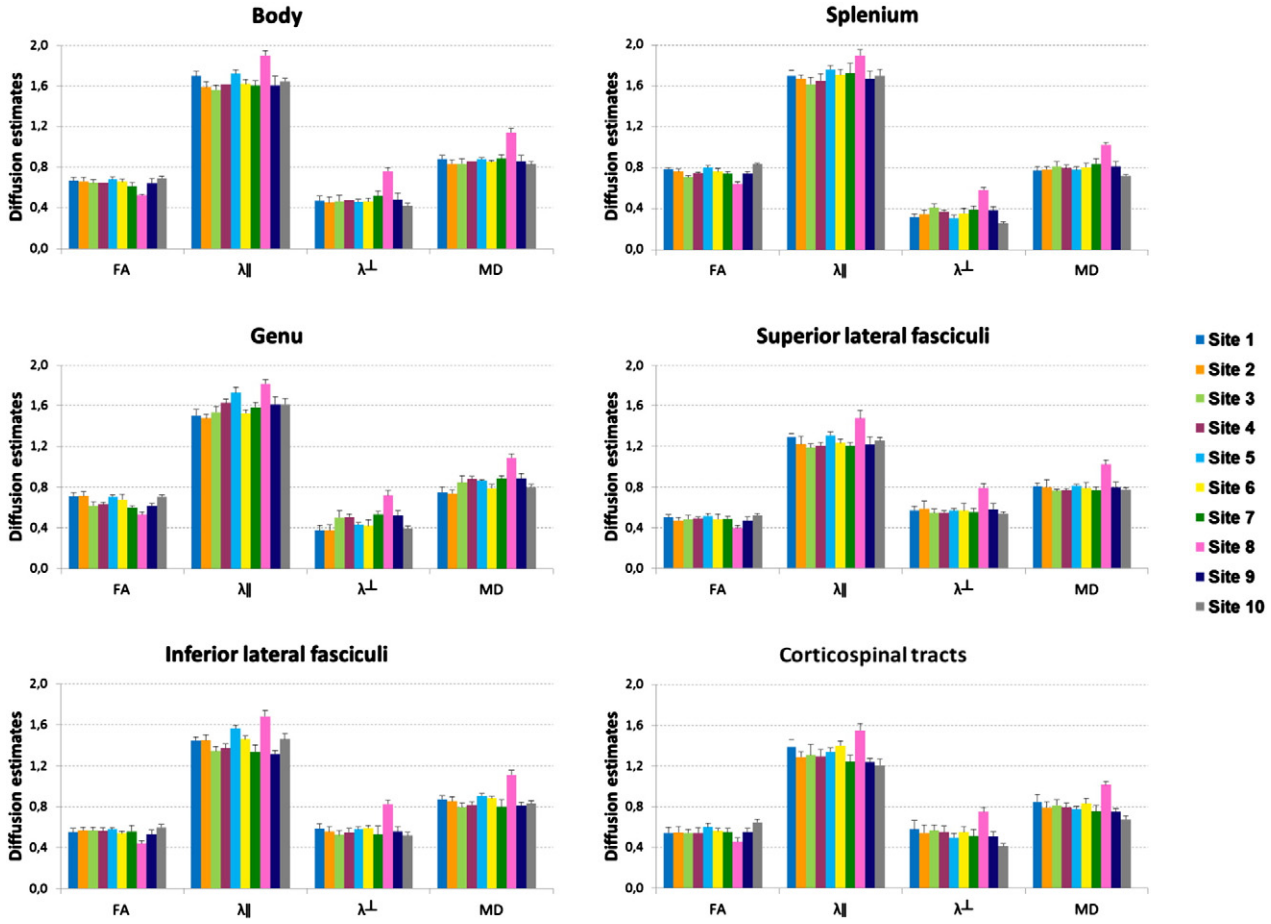

Fig. 6. Diffusion parameter estimates across sites and ROIs. Within-site group means and standard deviation (across subjects, scanner sessions and hemispheres) of fractional anisotropy (FA), mean, axial and radial diffusivity (MD, AD and RD, respectively). Significant MRI site effects (Kruskall–Wallis test, $p < 0.05$) were found for each structure and metric. See Table 2 for MRI sites characteristics.

that in a multi-site study MRI site with elderly people the sites with the longest DTI data acquisition protocols (i.e., the highest sensitivity to head motion effects) gave the maximum head translations and rotations as well as a tendency to higher reproducibility errors in some regions (corpus callosum body/splenium, superior lateral fasciculi). Variable head motion sensitivity across MRI sites could also depend on variability of the moments at which the DTI data were acquired within the session. This is relevant since there might be different stages of muscle relaxation during the acquisition (Leemans and Jones, 2009). In this study such effects were minimized by having the DTI acquisitions systematically be the last acquisition. Overall these results stress the importance of two separate but related aspects: i) the DTI acquisition duration in multi-site studies should be kept as homogeneous and as short as possible to minimize acquisition-related motion sensitivity biases, ii) regardless of the DTI acquisition time, it is important to monitor motion parameters carefully and potentially consider correction approaches to minimize motion related biases (Yendiki et al., 2013). The sites with longer acquisition times did not show significantly higher b0 SNR values relative to the other sites. This is most likely due to the fact that the longer acquisition times (which could not be modified in our clinical sites) is mostly driven by the longer protocol TRs (>9 s), which is more than an order of magnitude larger than the typical white matter T1 at 3 T (Wansapura et al., 1999; Wright et al., 2008). The SNR measured on the b0 images is expected to be minimally sensitive to head motion relative to the diffusion metrics because of the lack of diffusion sensitizing gradients.

This study has several general limitations; many of them were already discussed in a recent multi-site morphometry study from the same consortium (Jovicich et al., 2013). Specifically, each site scanned different subjects (just five), only twice while using an unbalanced distribution of MRI scanners (Siemens: 4, Philips: 4, GE: 2). The fact that subjects were different across sites means that our experimental design does not allow us to evaluate systematic differences across sites, but only compare the random errors in the reproducibility variability across sites. Systematic differences across sites could arise due to various reasons, including differences in the details of the acquisition sequence implementations. For example, despite the fact of having identical nominal acquisition parameters like voxel size and b-value, there is no guarantee that the corresponding effective values will be consistent across sites. Such parameters are known to affect the magnitude of the diffusion metrics (Jones and Cercignani, 2010; Oouchi et al., 2007), but it has also been seen that they do not affect test–retest reproducibility (Papinutto et al., 2013). We in fact measured significant MRI site effects in b0 SNR and diffusion metrics. These differences are most likely due to a combination of systematic and anatomical differences across sites. Despite these differences, the DTI-derived metrics obtained from our acquisitions and analysis protocol were found to be largely consistent across MRI sites and also in agreement with previously reported 3 T DTI studies on healthy elderly subjects (Likitjaroen et al., 2012; Mielke et al., 2009). The low sample size per site is an important limitation, so we explored how this could affect the reproducibility errors of the diffusion metrics. Two approaches were considered for increasing the population sample by pooling over different aspects of the data: a test for vendor effects and a test for correlation with b0 SNR. Both tests confirmed the main findings of reproducibility consistency across sites. Another limitation of this study is that there

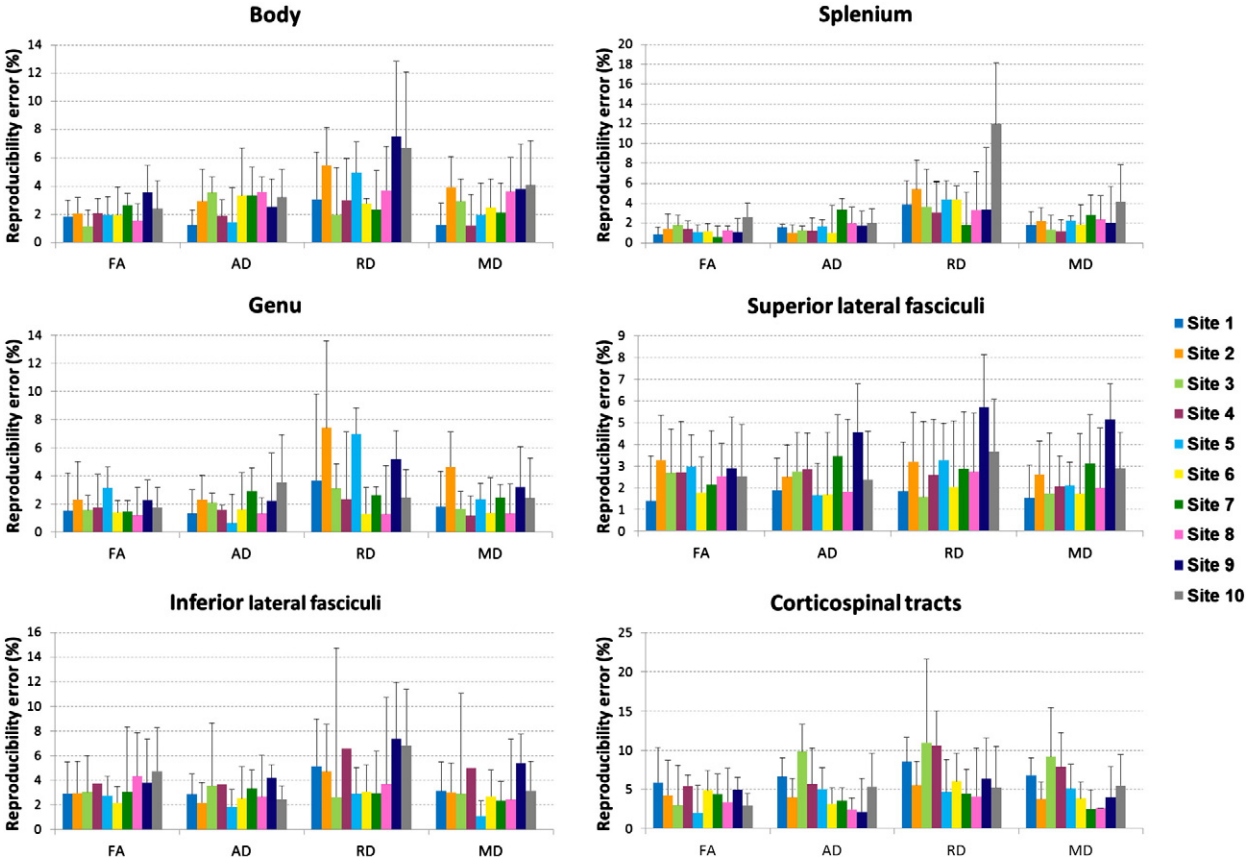

Fig. 7. Test–retest reproducibility of diffusion metrics across sites and ROIs. Within-site group mean reproducibility error and standard deviation of body, splenium, genu of the corpus callosum, superior lateral fasciculus, inferior lateral fasciculus, corticospinal tract (across subjects, sessions and hemispheres where possible). Test–retest-errors are shown on fractional anisotropy (FA), mean, axial and radial diffusivity (MD, AD and RD, respectively). A significant site effect was evident only for MD in the superior lateral fasciculi (Kruskall–Wallis test, $p < 0.05$). See Table 2 for MRI site characteristics.

is a wide range in scan–rescan days (7–55 days) across sites. However, two observations suggest that this range did not significantly bias our findings: first, the intervals did not give significant site effects, and secondly they did not correlate with any of the reproducibility errors (FA, MD, RD, AD). One additional limitation of this study is that the acquisition time for DTI data collection was limited (6–10 min) given that the overall acquisition included several other images and that the "clinical" condition to which the protocol is intended to be applied deals with patients with a relatively low level of collaboration (i.e. early stages of dementia). This constraint across all MRI sites made it difficult to use acquisition protocols with multiple b-values which can model also non-Gaussian diffusion properties of water. Such protocols are of interest not only because of their potentially higher reproducibility (Correia et al., 2009) but also because of their potential sensitivity to better characterize white matter lesions by means of diffusional kurtosis imaging metrics (Coutu et al., 2014; Fieremans et al., 2013). Finally, since this is not a random effects study our results cannot be generalized to acquisition protocols or populations other than those used and described here. Extension of this work includes further investigating possible sources of variability and ways to reduce them. One approach that

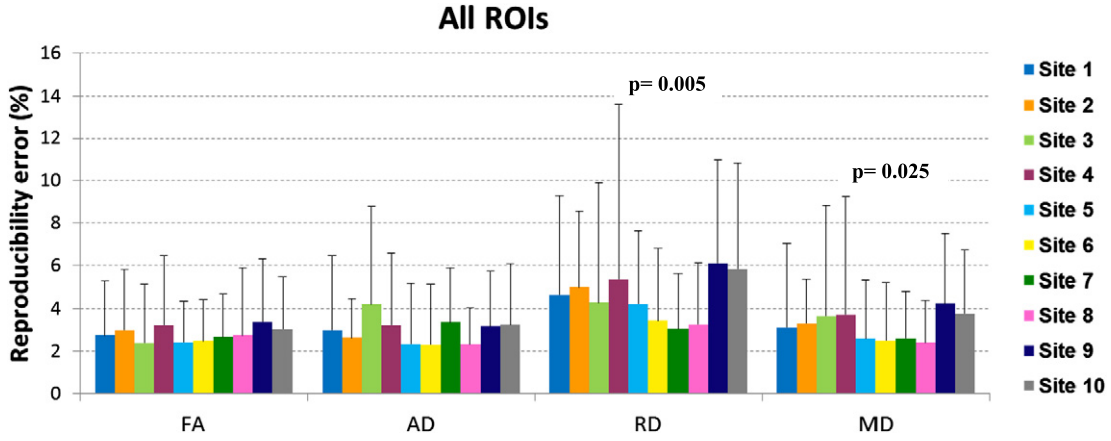

Fig. 8. Summary of MRI site effects on diffusion index reproducibility. Within-site group mean reproducibility error and standard deviation (across subjects, structures and sessions) in fractional anisotropy (FA), mean, axial and radial diffusivity (MD, AD and RD, respectively). A significant site effect was evident for radial and mean diffusivity (Kruskall–Wallis test, $p < 0.05$). See Table 2 for MRI sites characterization.

Table 5
Mean test–retest reproducibility errors values averaging across ROIs and sites with common MRI system vendors (see Table 2 for sites' descriptions). The last row indicates for each DTI-derived metric the Kruskal–Wallis results when testing for vendor effects.

Vendor group (number of subjects)	Test–retest reproducibility errors (%) grouping sites for each vendor			
	FA	MD	AD	RD
Siemens (n = 20)	2.62 ± 0.91	2.77 ± 1.11	2.67 ± 0.94	4.19 ± 1.52
Philips (n = 20)	2.63 ± 0.75	2.24 ± 1.44	2.99 ± 1.51	3.48 ± 1.50
GE (n = 10)	2.81 ± 1.18	2.85 ± 1.09	2.00 ± 1.00	4.97 ± 1.31
Kruskal–Wallis test for vendor effect	p = 0.88	p = 0.84	p = 0.20	p = 0.25

may potentially reduce test–retest variability could be that of adapting the TBSS analysis pipeline to use recently proposed methods that account for residual spatial variance from multiple time points (Engvig et al., 2012) as well as using improved registration methods for TBSS (de Groot et al., 2013). In addition, it has been recently shown that adding water free modeling to the diffusion tensor improves the specificity in DTI derived metrics by reducing partial volume effects from the CSF component, which might increase with aging or neurodegeneration (Pasternak et al., 2009, 2012). An intriguing possibility that needs to be evaluated is whether water free modeling can lead to improvements in the test–retest reliability, both within and across MRI sites.

To summarize, our study provides several estimations for the across-session reproducibility errors of the various diffusion metrics: i) per site over white matter ROIs (n = 5 subjects/site); ii) per site grouping across ROIs (n = 5 subjects/site), and iii) per vendor grouping across ROIs (n = 20 subjects for Siemens and Philips, n = 10 subjects for GE). Each of these results might help in designing future studies depending on the specific vendors used, the specific brain areas of interest, and particular target effect sizes. The multi-site anonymous diffusion metric imaging data generated in this study (100 brain volumes for each metric, FA, MD, axial (AD) and radial (RD) diffusivity) will be made publicly available to promote the development and evaluation of brain diffusion analysis tools (https://neugrid4you.eu/datasets).

Conclusions

Longitudinal multisite neuroimaging designs are typically used to identify differential tissue property changes associated with normal development, plasticity or disease progression/regression. The reliability of neuroanatomical measurements over time and across MRI sites is crucial for the statistical power of longitudinal studies. The main result of this multi-site study with ten 3 T MRI sites is that the across-session test–retest reproducibility/variability obtained with the protocol used was comparable to that reported by other studies which used either a lower number of MRI sites, hand-drawn ROI or longer acquisitions. This suggests that the protocol is suitable for multi-center longitudinal DTI studies. To the best of our knowledge this is the first study that confirms the multi-site DTI study feasibility in a 3 T consortium with various MRI vendors using vendor provided sequences for estimating across session test–retest reliability on a population of healthy elderly subjects in stable conditions. In addition, within the limitations of the sample size and MRI sites tested, our study provides preliminary reference values for absolute percent test–retest variability errors for a variety of white matter structures and diffusion metrics. An additional result suggested the importance of trying to keep the DTI acquisition time as short and homogeneous as possible across the consortium. This should help in reducing acquisition-related biases that may introduce higher reproducibility errors due to higher head motion sensitivity in longer acquisitions. Finally, the anonymous diffusion maps (FA, MD, RD, AD) will be made publicly available so that they can be used to further evaluate other diffusion analysis tools.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2014.06.075.

Acknowledgments

Pharmacog is funded by the EU-FP7 for the Innovative Medicine Initiative (grant no. 115009). All members of the Pharmacog project deserve sincere acknowledgement for their significant efforts, but unfortunately, they are too numerous to mention. The authors would like to especially thank the people who contributed to the early phases of this study, including Luca Venturi, Genoveffa Borsci, Thomas Günther, and Aurélien Monnet, as well as Alberto Redolfi for his support with the Intellimaker database and Neugrid portal. We also thank the reviewers for all their comments which helped improve the clarity of the paper.

Conflict of interest

The authors have no conflict of interests to declare.

Appendix A. Selection and exclusion criteria of study subjects

Selection criteria:

Participants will be (i) healthy volunteers (between 50 and 80 years old) and/or (ii) subjects (between 50 and 80 years old), who will perform a 3 T-MRI for reasons such as migraine, headache, auditory or visual symptoms, paresthesias, and whose scan will be negative (see exclusion criteria below).

Exclusion criteria:

- Based on the medical history:
  - Ischaemic lesions already detected in a previous scan
  - Head injury with loss of consciousness > 24 h
  - Current substance abuse
  - Current therapy with steroids or current chemotherapy
  - Loss of weight > 5 kg in the last 6 months
  - Systemic disease with frequent involvement of the CNS (lupus, HIV, rheumatoid arthritis)
  - CNS disease diagnosed by a specialist or in treatment (such as epilepsy, ictus).
- Based on the MRI scan:
  - Cerebral metastasis or CNS primary tumor still benign (except for pituitary microadenoma)
  - Suspected multiple sclerosis + MRI evidence of white matter lesions
  - Suspected recent stroke + MRI evidence of infarct
  - Aneurysm > 10 mm and arteriovenous malformations (except for venous angioma)
  - Dysgenesia of central nervous system.

The method (ii) of enrolment has been used to enroll a large group of cognitively intact persons in the Italian Brain Normative Archive of structural MR scans (the IBNA study) aimed to build an Italian archive of high resolution MR images of normal subjects to be used as normative references for patients with suspected neurodegenerative disorders (Galluzzi et al., 2009).

References

Alhamud, A., Tisdall, M.D., Hess, A.T., Hasan, K.M., Meintjes, E.M., van der Kouwe, A.J., 2012. Volumetric navigators for real-time motion correction in diffusion tensor imaging. Magn. Reson. Med. 68, 1097–1108.

Andersson, J.L.R., Jenkinson, M., Smith, S., 2007a. Non-linear optimisation. FMRIB technical reportTR07JA1 from www.fmrib.ox.ac.uk/analysis/techrep.

Andersson, J.L.R., Jenkinson, M., Smith, S., 2007b. Non-linear registration, aka spatial normalisation. FMRIB technical reportTR07JA2 from www.fmrib.ox.ac.uk/analysis/techrep.

Basser, P., Pierpaoli, C., 1996. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. J. Magn. Reson. Ser. B 111, 209–219.
Behrens, T.E., Woolrich, M.W., Jenkinson, M., Johansen-Berg, H., Nunes, R.G., Clare, S., Matthews, P.M., Brady, J.M., Smith, S.M., 2003. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. Magn. Reson. Med. 50, 1077–1088.
Bisdas, S., Bohning, D.E., Besenski, N., Nicholas, J.S., Rumboldt, Z., 2008. Reproducibility, interrater agreement, and age-related changes of fractional anisotropy measures at 3 T in healthy subjects: effect of the applied b-value. AJNR Am. J. Neuroradiol. 29, 1128–1133.
Bonekamp, D., Nagae, L.M., Degaonkar, M., Matson, M., Abdalla, W.M., Barker, P.B., Mori, S., Horská, A., 2007. Diffusion tensor imaging in children and adolescents: reproducibility, hemispheric, and age-related differences. NeuroImage 34, 733–742.
Brander, A., Kataja, A., Saastamoinen, A., Ryymin, P., Huhtala, H., Ohman, J., Soimakallio, S., Dastidar, P., 2010. Diffusion tensor imaging of the brain in a healthy adult population: Normative values and measurement reproducibility at 3 T and 1.5 T. Acta Radiol. 51, 800–807.
Calabrese, M., Rinaldi, F., Seppi, D., Favaretto, A., Squarcina, L., Mattisi, I., Perini, P., Bertoldo, A., Gallo, P., 2011. Cortical diffusion-tensor imaging abnormalities in multiple sclerosis: a 3-year longitudinal study. Radiology 261, 891–898.
Ciccarelli, O., Parker, G.J., Toosy, A.T., Wheeler-Kingshott, C.A., Barker, G.J., Boulby, P.A., Miller, D.H., Thompson, A.J., 2003. From diffusion tractography to quantitative white matter tract measures: a reproducibility study. NeuroImage 18, 348–359.
Ciccarelli, O., Catani, M., Johansen-Berg, H., Clark, C., Thompson, A., 2008. Diffusion-based tractography in neurological disorders: concepts, applications, and future developments. Lancet Neurol. 7, 715–727.
Correia, M.M., Carpenter, T.A., Williams, G.B., 2009. Looking for the optimal DTI acquisition scheme given a maximum scan time: are more b-values a waste of time? Magn. Reson. Imaging 27, 163–175.
Coutu, J.P., Chen, J.J., Rosas, H.D., Salat, D.H., 2014 Jun. Non-Gaussian water diffusion in aging white matter. Neurobiol Aging 35 (6), 1412–1421.
de Groot, M., Vernooij, M.W., Klein, S., Ikram, M.A., Vos, F.M., Smith, S.M., Niessen, W.J., Andersson, J.L., 2013. Improving alignment in Tract-based spatial statistics: evaluation and optimization of image registration. NeuroImage 76, 400–411.
Dietrich, O., Raya, J.G., Reeder, S.B., Reiser, M.F., Schoenberg, S.O., 2007. Measurement of signal-to-noise ratios in MR images: influence of multichannel coils, parallel imaging, and reconstruction filters. J. Magn. Reson. Imaging 26, 375–385.
Drago, V., Babiloni, C., Bartrés-Faz, D., Caroli, A., Bosch, B., Hensch, T., Didic, M., Klafki, H.W., Pievani, M., Jovicich, J., Venturi, L., Spitzer, P., Vecchio, F., Schoenknecht, P., Wiltfang, J., Redolfi, A., Forloni, G., Blin, O., Irving, E., Davis, C., Hårdemark, H.G., Frisoni, G.B., 2011. Disease tracking markers for Alzheimer's disease at the prodromal (MCI) stage. J. Alzheimers Dis. 26 (Suppl. 3), 159–199.
Engvig, A., Fjell, A.M., Westlye, L.T., Moberget, T., Sundseth, Ø., Larsen, V.A., Walhovd, K.B., 2012. Memory training impacts short-term changes in aging white matter: a longitudinal diffusion tensor imaging study. Hum. Brain Mapp. 33, 2390–2406.
Farrell, J.A., Landman, B.A., Jones, C.K., Smith, S.A., Prince, J.L., van Zijl, P.C., Mori, S., 2007. Effects of signal-to-noise ratio on the accuracy and reproducibility of diffusion tensor imaging-derived fractional anisotropy, mean diffusivity, and principal eigenvector measurements at 1.5 T. J. Magn. Reson. Imaging 26, 756–767.
Fieremans, E., Benitez, A., Jensen, J.H., Falangola, M.F., Tabesh, A., Deardorff, R.L., Spampinato, M.V., Babb, J.S., Novikov, D.S., Ferris, S.H., Helpern, J.A., 2013 Nov-Dec. Novel white matter tract integrity metrics sensitive to Alzheimer disease progression. AJNR Am J Neuroradiol 34 (11), 2105–2112.
Fox, R.J., Sakaie, K., Lee, J.C., Debbins, J.P., Liu, Y., Arnold, D.L., Melhem, E.R., Smith, C.H., Philips, M.D., Lowe, M., Fisher, E., 2012. A validation study of multicenter diffusion tensor imaging: reliability of fractional anisotropy and diffusivity values. AJNR Am. J. Neuroradiol. 33, 695–700.
Galluzzi, S., Testa, C., Boccardi, M., Bresciani, L., Benussi, L., Ghidoni, R., Beltramello, A., Bonetti, M., Bono, G., Falini, A., Magnani, G., Minonzio, G., Piovan, E., Binetti, G., Frisoni, G.B., 2009. The Italian Brain Normative Archive of structural MR scans: norms for medial temporal atrophy and white matter lesions. Aging Clin. Exp. Res. 21, 266–276.
Harrison, D.M., Caffo, B.S., Shiee, N., Farrell, J.A., Bazin, P.L., Farrell, S.K., Ratchford, J.N., Calabresi, P.A., Reich, D.S., 2011. Longitudinal changes in diffusion tensor-based quantitative MRI in multiple sclerosis. Neurology 76, 179–186.
Heiervang, E., Behrens, T.E., Mackay, C.E., Robson, M.D., Johansen-Berg, H., 2006. Between session reproducibility and between subject variability of diffusion MR and tractography measures. NeuroImage 33, 867–877.
Huang, L., Wang, X., Baliki, M.N., Wang, L., Apkarian, A.V., Parrish, T.B., 2012. Reproducibility of structural, resting-state BOLD and DTI Data between identical scanners. PLoS ONE 7, e47684.
Jansen, J.F., Kooi, M.E., Kessels, A.G., Nicolay, K., Backes, W.H., 2007. Reproducibility of quantitative cerebral T2 relaxometry, diffusion tensor imaging, and 1H magnetic resonance spectroscopy at 3.0 Tesla. Investig. Radiol. 42, 327–337.
Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. NeuroImage 17, 825–841.
Jones, D.K., Cercignani, M., 2010. Twenty-five pitfalls in the analysis of diffusion MRI data. NMR Biomed. 23, 803–820.
Jovicich, J., Marizzoni, M., Sala-Llonch, R., Bosch, B., Bartrés-Faz, D., Arnold, J., Benninghoff, J., Wiltfang, J., Roccatagliata, L., Nobili, F., Hensch, T., Tränkner, A., Schönknecht, P., Leroy, M., Lopes, R., Bordet, R., Chanoine, V., Ranjeva, J.P., Didic, M., Gros-Dagnac, H., Payoux, P., Zoccatelli, G., Alessandrini, F, Beltramello, A., Bargalló, N., Blin, O., Frisoni, G.B., 2013 Dec. PharmaCog Consortium. Brain morphometry reproducibility in multi-center 3T MRI studies: a comparison of cross-sectional and longitudinal segmentations. Neuroimage 83, :472–484.
Kantarci, K., Avula, R., Senjem, M.L., Samikoglu, A.R., Zhang, B., Weigand, S.D., Przybelski, S.A., Edmonson, H.A., Vemuri, P., Knopman, D.S., Ferman, T.J., Boeve, B.F., Petersen, R.C., Jack, C.R., 2010. Dementia with Lewy bodies and Alzheimer disease: neurodegenerative patterns characterized by DTI. Neurology 74, 1814–1821.
Laganà, M., Rovaris, M., Ceccarelli, A., Venturelli, C., Marini, S., Baselli, G., 2010. DTI parameter optimisation for acquisition at 1.5 T: SNR analysis and clinical application. Comput. Intell. Neurosci. 254032.
Landman, B.A., Farrell, J.A., Jones, C.K., Smith, S.A., Prince, J.L., Mori, S., 2007. Effects of diffusion weighting schemes on the reproducibility of DTI-derived fractional anisotropy, mean diffusivity, and principal eigenvector measurements at 1.5 T. NeuroImage 36, 1123–1138.
Lebel, C., Beaulieu, C., 2011. Longitudinal development of human brain wiring continues from childhood into adulthood. J. Neurosci. 31, 10937–10947.
Leemans, A., Jones, D.K., 2009. The B-matrix must be rotated when correcting for subject motion in DTI data. Magn. Reson. Med. 61, 1336–1349.
Likitjaroen, Y., Meindl, T., Friese, U., Wagner, M., Buerger, K., Hampel, H., Teipel, S.J., 2012. Longitudinal changes of fractional anisotropy in Alzheimer's disease patients treated with galantamine: a 12-month randomized, placebo-controlled, double-blinded study. Eur. Arch. Psychiatry Clin. Neurosci. 262, 341–350.
Ling, J., Merideth, F., Caprihan, A., Pena, A., Teshiba, T., Mayer, A.R., 2012. Head injury or head motion? Assessment and quantification of motion artifacts in diffusion tensor imaging studies. Hum. Brain Mapp. 33, 50–62.
Liu, Z., Wang, Y., Gerig, G., Gouttard, S., Tao, R., Fletcher, T., Styner, M., 2010. Quality control of diffusion weighted images. Proc. Soc. Photo. Opt. Instrum. Eng. 7628.
Magnotta, V.A., Kim, J., Koscik, T., Beglinger, L.J., Espinso, D., Langbehn, D., Nopoulos, P., Paulsen, J.S., 2009. Diffusion tensor imaging in preclinical Huntington's disease. Brain Imaging Behav. 3, 77–84.
Magnotta, V.A., Matsui, J.T., Liu, D., Johnson, H.J., Long, J.D., Bolster, B.D., Mueller, B.A., Lim, K., Mori, S., Helmer, K.G., Turner, J.A., Reading, S., Lowe, M.J., Aylward, E., Flashman, L.A., Bonett, G., Paulsen, J.S., 2012. Multicenter reliability of diffusion tensor imaging. Brain Connect. 2, 345–355.
Marenco, S., Rawlings, R., Rohde, G.K., Barnett, A.S., Honea, R.A., Pierpaoli, C., Weinberger, D.R., 2006. Regional distribution of measurement error in diffusion tensor imaging. Psychiatry Res. 147, 69–78.
Mielke, M.M., Kozauer, N.A., Chan, K.C., George, M., Toroney, J., Zerrate, M., Bandeen-Roche, K., Wang, M.C., Vanzijl, P., Pekar, J.J., Mori, S., Lyketsos, C.G., Albert, M., 2009. Regionally-specific diffusion tensor imaging in mild cognitive impairment and Alzheimer's disease. NeuroImage 46, 47–55.
Mori, S., Zhang, J., 2006. Principles of diffusion tensor imaging and its applications to basic neuroscience research. Neuron 51, 527–539.
Oouchi, H., Yamada, K., Sakai, K., Kizu, O., Kubota, T., Ito, H., Nishimura, T., 2007. Diffusion anisotropy measurement of brain white matter is affected by voxel size: underestimation occurs in areas with crossing fibers. AJNR Am. J. Neuroradiol. 28, 1102–1106.
Pagani, E., Hirsch, J.G., Pouwels, P.J., Horsfield, M.A., Perego, E., Gass, A., Roosendaal, S.D., Barkhof, F., Agosta, F., Rovaris, M., Caputo, D., Giorgio, A., Palace, J., Marino, S., De Stefano, N., Ropele, S., Fazekas, F., Filippi, M., 2010. Intercenter differences in diffusion tensor MRI acquisition. J. Magn. Reson. Imaging 31, 1458–1468.
Papinutto, N.D., Maule, F., Jovicich, J., 2013. Reproducibility and biases in high field brain diffusion MRI: an evaluation of acquisition and analysis variables. Magn. Reson. Imaging 31, 827–839.
Parker, G.J., Haroon, H.A., Wheeler-Kingshott, C.A., 2003. A framework for a streamline-based probabilistic index of connectivity (PICo) using a structural interpretation of MRI diffusion measurements. J. Magn. Reson. Imaging 18, 242–254.
Pasternak, O., Sochen, N., Gur, Y., Intrator, N., Assaf, Y., 2009. Free water elimination and mapping from diffusion MRI. Magn. Reson. Med. 62, 717–730.
Pasternak, O., Westin, C.F., Bouix, S., Seidman, L.J., Goldstein, J.M., Woo, T.U., Petryshen, T.L., Mesholam-Gately, R.I., McCarley, R.W., Kikinis, R., Shenton, M.E., Kubicki, M., 2012. Excessive extracellular volume reveals a neurodegenerative pattern in schizophrenia onset. J. Neurosci. 32, 17365–17372.
Pfefferbaum, A., Adalsteinsson, E., Sullivan, E.V., 2003. Replicability of diffusion tensor imaging measurements of fractional anisotropy and trace in brain. J. Magn. Reson. Imaging 18, 427–433.
Rashid, W., Hadjiprocopis, A., Davies, G., Griffin, C., Chard, D., Tiberio, M., Altmann, D., Wheeler-Kingshott, C., Tozer, D., Thompson, A., Miller, D.H., 2008. Longitudinal evaluation of clinically early relapsing–remitting multiple sclerosis with diffusion tensor imaging. J. Neurol. 255, 390–397.
Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L., Leach, M.O., Hawkes, D.J., 1999. Nonrigid registration using free-form deformations: application to breast MR images. IEEE Trans. Med. Imaging 18, 712–721.
Sage, C.A., Van Hecke, W., Peeters, R., Sijbers, J., Robberecht, W., Parizel, P., Marchal, G., Leemans, A., Sunaert, S., 2009. Quantitative diffusion tensor imaging in amyotrophic lateral sclerosis: revisited. Hum. Brain Mapp. 30, 3657–3675.
Scola, E., Bozzali, M., Agosta, F., Magnani, G., Franceschi, M., Sormani, M.P., Cercignani, M., Pagani, E., Falautano, M., Filippi, M., Falini, A., 2010. A diffusion tensor MRI study of patients with MCI and AD with a 2-year clinical follow-up. J. Neurol. Neurosurg. Psychiatry 81, 798–805.
Sidaros, A., Engberg, A.W., Sidaros, K., Liptrot, M.G., Herning, M., Petersen, P., Paulson, O.B., Jernigan, T.L., Rostrup, E., 2008. Diffusion tensor imaging during recovery from severe traumatic brain injury and relation to clinical outcome: a longitudinal study. Brain 131, 559–572.
Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E., Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M.Z., Matthews, P.M., Behrens, T.E., 2006. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. NeuroImage 31, 1487–1505.

Song, S.K., Yoshino, J., Le, T.Q., Lin, S.J., Sun, S.W., Cross, A.H., Armstrong, R.C., 2005. Demyelination increases radial diffusivity in corpus callosum of mouse brain. NeuroImage 26, 132–140.
Sritharan, A., Egan, G.F., Johnston, L., Horne, M., Bradshaw, J.L., Bohanna, I., Asadi, H., Cunnington, R., Churchyard, A.J., Chua, P., Farrow, M., Georgiou-Karistianis, N., 2010. A longitudinal diffusion tensor imaging study in symptomatic Huntington's disease. J. Neurol. Neurosurg. Psychiatry 81, 257–262.
Sullivan, E.V., Pfefferbaum, A., 2007. Neuroradiological characterization of normal adult ageing. Br. J. Radiol. 80 (Spec No 2), S99–S108.
Sullivan, E.V., Rohlfing, T., Pfefferbaum, A., 2010. Longitudinal study of callosal microstructure in the normal adult aging brain using quantitative DTI fiber tracking. Dev. Neuropsychol. 35, 233–256.
Takao, H., Hayashi, N., Kabasawa, H., Ohtomo, K., 2012. Effect of scanner in longitudinal diffusion tensor imaging studies. Hum. Brain Mapp. 33, 466–477.
Teipel, S.J., Meindl, T., Wagner, M., Stieltjes, B., Reuter, S., Hauenstein, K.H., Filippi, M., Ernemann, U., Reiser, M.F., Hampel, H., 2010. Longitudinal changes in fiber tract integrity in healthy aging and mild cognitive impairment: a DTI follow-up study. J. Alzheimers Dis. 22, 507–522.
Vaessen, M.J., Hofman, P.A., Tijssen, H.N., Aldenkamp, A.P., Jansen, J.F., Backes, W.H., 2010. The effect and reproducibility of different clinical DTI gradient sets on small world brain connectivity measures. NeuroImage 51, 1106–1116.
Van Horn, J.D., Toga, A.W., 2009. Multisite neuroimaging trials. Curr. Opin. Neurol. 22, 370–378.
Vollmar, C., O'Muircheartaigh, J., Barker, G.J., Symms, M.R., Thompson, P., Kumari, V., Duncan, J.S., Richardson, M.P., Koepp, M.J., 2010. Identical, but not the same: intra-site and inter-site reproducibility of fractional anisotropy measures on two 3.0 T scanners. NeuroImage 51, 1384–1394.
Wahlund, L.O., Barkhof, F., Fazekas, F., Bronge, L., Augustin, M., Sjögren, M., Wallin, A., Ader, H., Leys, D., Pantoni, L., Pasquier, F., Erkinjuntti, T., Scheltens, P., Changes, E.T. F.o.A.-R.W.M., 2001. A new rating scale for age-related white matter changes applicable to MRI and CT. Stroke 32, 1318–1322.
Wakana, S., Caprihan, A., Panzenboeck, M.M., Fallon, J.H., Perry, M., Gollub, R.L., Hua, K., Zhang, J., Jiang, H., Dubey, P., Blitz, A., van Zijl, P., Mori, S., 2007. Reproducibility of quantitative tractography methods applied to cerebral white matter. NeuroImage 36, 630–644.
Wang, C., Stebbins, G.T., Nyenhuis, D.L., deToledo-Morrell, L., Freels, S., Gencheva, E., Pedelty, L., Sripathirathan, K., Moseley, M.E., Turner, D.A., Gabrieli, J.D., Gorelick, P.B., 2006. Longitudinal changes in white matter following ischemic stroke: a three-year follow-up study. Neurobiol. Aging 27, 1827–1833.
Wansapura, J.P., Holland, S.K., Dunn, R.S., Ball, W.S., 1999. NMR relaxation times in the human brain at 3.0 Tesla. J. Magn. Reson. Imaging 9, 531–538.
Weaver, K.E., Richards, T.L., Liang, O., Laurino, M.Y., Samii, A., Aylward, E.H., 2009. Longitudinal diffusion tensor imaging in Huntington's Disease. Exp. Neurol. 216, 525–529.
Wedeen, V.J., Hagmann, P., Tseng, W.Y., Reese, T.G., Weisskoff, R.M., 2005. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. Magn. Reson. Med. 54, 1377–1386.
Wedeen, V.J., Wang, R.P., Schmahmann, J.D., Benner, T., Tseng, W.Y., Dai, G., Pandya, D.N., Hagmann, P., D'Arceuil, H., de Crespigny, A.J., 2008. Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. NeuroImage 41, 1267–1277.
Westlye, L.T., Walhovd, K.B., Dale, A.M., Bjørnerud, A., Due-Tønnessen, P., Engvig, A., Grydeland, H., Tamnes, C.K., Ostby, Y., Fjell, A.M., 2010. Life-span changes of the human brain white matter: diffusion tensor imaging (DTI) and volumetry. Cereb. Cortex 20, 2055–2068.
Wheeler-Kingshott, C.A., Cercignani, M., 2009. About "axial" and "radial" diffusivities. Magn. Reson. Med. 61, 1255–1260.
White, T., Magnotta, V.A., Bockholt, H.J., Williams, S., Wallace, S., Ehrlich, S., Mueller, B.A., Ho, B.C., Jung, R.E., Clark, V.P., Lauriello, J., Bustillo, J.R., Schulz, S.C., Gollub, R.L., Andreasen, N.C., Calhoun, V.D., Lim, K.O., 2011. Global white matter abnormalities in schizophrenia: a multisite diffusion tensor imaging study. Schizophr. Bull. 37, 222–232.
Wright, P.J., Mougin, O.E., Totman, J.J., Peters, A.M., Brookes, M.J., Coxon, R., Morris, P.E., Clemence, M., Francis, S.T., Bowtell, R.W., Gowland, P.A., 2008. Water proton T1 measurements in brain tissue at 7, 3, and 1.5 T using IR-EPI, IR-TSE, and MPRAGE: results and optimization. MAGMA 21, 121–130.
Yendiki, A., Koldewyn, K., Kakunoori, S., Kanwisher, N., Fischl, B., 2013 Nov 21. Spurious group differences due to head motion in a diffusion MRI study. NeuroImage 88c, 79–90. http://dx.doi.org/10.1016/j.neuroimage.2013.11.027.
Zhu, T., Hu, R., Qiu, X., Taylor, M., Tso, Y., Yiannoutsos, C., Navia, B., Mori, S., Ekholm, S., Schifitto, G., Zhong, J., 2011. Quantification of accuracy and precision of multi-center DTI measurements: a diffusion phantom and human brain study. NeuroImage 56, 1398–1411.