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