Brain Research 1669 (2017) 131–140

Contents lists available at ScienceDirect

Brain Research

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

Research report

Statistical differences in the white matter tracts in subjects with depression by using different skeletonized voxel-wise analysis approaches and DTI fitting procedures

Maurizio Bergamino [a,*], Madison Farmer [b,1], Hung-wen Yeh [a,2], Elisabeth Paul [c,3], J. Paul Hamilton [c,3]

[a] Laureate Institute for Brain Research, Tulsa, OK, USA
[b] Roosevelt University, Department of Industrial and Organizational Psychology, Chicago, IL, USA
[c] Center for Social and Affective Neuroscience, Department of Clinical and Experimental Medicine, Linköping, Sweden

ARTICLE INFO

Article history:
Received 7 February 2017
Received in revised form 9 June 2017
Accepted 12 June 2017
Available online 16 June 2017

Keywords:
Diffusion tensor imaging
Major depressive disorder
TBSS
ANTs
DTI fitting algorithms

ABSTRACT

Major depressive disorder (MDD) is one of the most significant contributors to the global burden of illness. Diffusion tensor imaging (DTI) is a procedure that has been used in several studies to characterize abnormalities in white matter (WM) microstructural integrity in MDD. These studies, however, have provided divergent findings, potentially due to the large variety of methodological alternatives available in conducting DTI research. In order to determine the importance of different approaches to coregistration of DTI-derived metrics to a standard space, we compared results from two different skeletonized voxel-wise analysis approaches: the standard TBBS pipeline and the Advanced Normalization Tools (ANTs) approach incorporating a symmetric image normalization (SyN) algorithm and a group-wise template (ANTs TBSS). We also assessed effects of applying twelve different fitting procedures for the diffusion tensor. For our dataset, lower fractional anisotropy (FA) and axial diffusivity (AD) in depressed subjects compared with healthy controls were found for both methods and for all fitting procedures. No group differences were found for radial and mean diffusivity indices. Importantly, for the AD metric, the normalization methods and fitting procedures showed reliable differences, both in the volume and in the number of significant between-groups difference clusters detected. Additionally, a significant voxel-based correlation, in the left inferior fronto-occipital fasciculus, between AD and self-reported stress was found only for one of the normalization procedure (ANTs TBSS). In conclusion, the sensitivity to detect group-level effects on DTI metrics might depend on the DTI normalization and/or tensor fitting procedures used.

1. Introduction

Major depressive disorder (MDD) is a chronic and debilitating disease that currently accounts for a significant portion of the global disability burden (Lyness et al., 2002). Given that disturbances in the functional and structural integrity of the brain have been implicated in the pathophysiology of depression, investigating the neural substrates of MDD will continue to be vital until brain-based interventions for depression are improved.

In the last decade, white matter (WM) integrity—as derived from magnetic resonance imaging (MRI) assessments—has been explored in MDD. Diffusion tensor imaging (DTI) is an MR-based procedure used to characterize WM tracts in the living brain. This procedure renders indices—most prominently, fractional anisotropy (FA), radial diffusivity (RD), axial diffusivity (AD), and mean diffusivity (MD)—that can be used to operationalize the integrity of WM tracts. Several studies have reported lower FA values in depressed versus healthy control (HC) samples, suggesting WM fiber disruption or disorganization in this disorder.

Low FA values have been reported in a variety of WM tracts including the inferior longitudinal fasciculus (ILF) (Versace et al., 2010; Zou et al., 2008), superior longitudinal fasciculus (SLF) (Murphy et al., 2012; Wu et al., 2011), inferior fronto-occipital fasciculus (IFOF) (Bergamino et al., 2016; Cullen et al., 2010), internal

* Corresponding author at: Laureate Institute for Brain Research, 6655 South Yale Ave., Tulsa, OK 74136, USA.

E-mail addresses: maurizio.bergamino@gmail.com (M. Bergamino), madisonc-farmer@gmail.com (M. Farmer), hyeh@laureateinstitute.org (H.-w. Yeh), elisabeth.paul@student.uva.nl (E. Paul), paul.hamilton@liu.se (J.P. Hamilton).

[1] Roosevelt University, Department of Industrial and Organizational Psychology, 430 S Michigan Ave, AUD 827, Chicago, IL 60605, USA.

[2] Laureate Institute for Brain Research, 6655 South Yale Ave., Tulsa, OK 74136, USA.

[3] Center for Social and Affective Neuroscience, Department of Clinical and Experimental Medicine, Linköping, Sweden.

http://dx.doi.org/10.1016/j.brainres.2017.06.013
0006-8993/

and external capsule (Guo et al., 2012; Zou et al., 2008), and in the corpus callosum (Guo et al., 2012; Murphy et al., 2012). In addition to FA, differences in other DTI-derived metrics have been observed in depressed relative to healthy people. Lower values of AD have been found in depressed subjects inside IFOF (Bergamino et al., 2016) and in the anterior thalamic radiation (ATR). Elevated values of RD (Lai and Wu, 2014), apparent diffusion coefficient (ADC) (Wu et al., 2011), and MD (Benedetti et al., 2011) have been identified in the SLF as well as in the majority of the WM fiber bundles connecting structures of the anterior limbic network. These findings, however, have not been consistently replicated (Abe et al., 2010; Choi et al., 2014; Olvet et al., 2016).

The variability of findings from DTI-based studies of MDD could be accorded to variation in the gender and age composition of samples used and/or the relatively low number of the subjects involved in the studies. Importantly, though, methodological heterogeneity could also account for these discrepant findings. Indeed, several different analysis techniques can be applied in order to obtain DTI-derived metrics; these include region of interest (ROI) analysis (Li et al., 2007), the histogram method (Tha et al., 2013), voxel-based analysis (VBA) (Srivastava et al., 2016), and the tract-based spatial statistics (TBSS) approach (Han et al., 2014; Zuo et al., 2012).

The current industry standard in voxel-based DTI analysis is TBSS. The standard TBSS pipeline normalizes FA maps—and applies to the other DTI indices this same normalization—to a common standard space, and then projects these maps onto a "skeleton" which is obtained by "thinning" the non maximum-suppression perpendicular to the local tract structure in the group-mean FA image. Subsequent statistical analyses are then implemented directly in these DTI-skeletonized maps. Even though several investigators (Zalesky, 2011) have questioned the reliability and interpretability of TBSS, especially in its coregistration algorithms, it remains the most popular software pipeline to coregister DTI-derived metrics for performing voxel-wise comparisons. However, other coregistration algorithms can be used to normalize DTI maps to a common standard space. Advanced Normalization Tools (ANTs) using a symmetric image normalization (SyN) algorithm (Avants et al., 2008), for example, has already been used in different DTI studies. Indeed, Klein and colleagues found that this algorithm provides superior registration performance in T1-weighted MR registration tasks and metrics when compared to FNIRT and other coregistration tools (Klein et al., 2009). Furthermore, Tustison et. al. found that the ANTs algorithm is superior to FNIRT specifically for FA coregistration (Tustison et al., 2014).

In addition to normalization procedures, algorithms used for tensor fitting of diffusion-weighted data can have important effects on obtained results. The linear least square (LSQ) model, for example, is among the most basic models used in diffusion MRI for the estimation of diffusion parameters. The fitting procedure that employs this algorithm is very fast, but it incorrectly assumes that data outliers are homogeneously distributed and therefore fails to appropriately de-weight their contributions. On the other hand, the weighted linear least square (WLSQ) approach assigns a weight according to how much the original noise variation is affected by the logarithmic transform of the data. This fitting algorithm is slightly slower than the LSQ, but is more precise. For non-linear least squares (NLS) algorithms, the estimation of the tensor is performed directly from the signal equation $S = S_0 \exp(-b\bar{D})$. Due to their non-linear nature, however, relatively slow and computationally intensive, iterative regression algorithms are needed to minimize the error between the predicted signal and observed signal intensity in the DTI acquisition. Other more robust fitting algorithms can be found in literature, such as robust estimation of tensors by outlier rejection (RESTORE) (Chang et al., 2005), or the informed RESTORE (iRESTORE) (Chang et al., 2012). These last two algorithms are not typically used in DTI studies, in spite of the fact that they have been shown to provide more robust estimates of diffusion parameters than other fitting procedures. Importantly, too, complex interactions between coregistration methods and fitting procedures have been identified. Maximov et al., for example, demonstrated that the agreement and reliability of TBSS results depend on the DTI data fitting procedures applied (Maximov et al., 2015).

In the present methodological study, we investigated WM microstructural integrity by using different DTI-derived metrics, in individuals with depression in comparison with age-and gender-matched healthy controls (HC). We assessed WM integrity by analyzing DTI-derived skeletonized maps through two different methods: the standard TBSS pipeline and another skeletonized approach, where the coregistration algorithms used in TBSS (FLIRT and FNIRT) were replaced with the ANTs SyN coregistration and with a group template (ANTs TBSS). In addition, we combined these two coregistration approaches with twelve different algorithms for fitting the DTI data. We hypothesized that different fitting procedures would lead to different results, and that more robust fitting algorithms would improve the final statistical analysis. Moreover, we predicted that the skeletonized analysis procedure incorporated in ANTs TBSS would improve the DTI normalization to a standard space, and therefore, increase sensitivity to detect between-groups differences in DTI metrics.

Relatively recently, we have acquired methodological tools for understanding the biological substrates—and effects—of stress in the pathophysiology of depression (Sapolsky, 1996). Indeed, atrophy in brain structures, such as the hippocampus, particularly vulnerable to the effects of stress hormones have been observed reliably in this disorder (Campbell et al., 2004). Moreover, we recently observed that stress-related effects in depression can occur in white-matter tracts as well (Bergamino et al., 2016). Given this, in the current investigation we examined effects on the sensitivity to detect stress-related white-matter anomalies in MDD of different processing pipelines and tensor fitting procedures.

2. Results

Neuroimaging data from two depressed participants and one HC were removed from the study due to excessive motion; therefore, data from twenty-four depressed and twenty-three HC subjects were included in the final analysis. Depressed and HC groups did not differ significantly in age ($p = 0.24$). The groups did differ with respect to PSS score (mean (SD) for MDD 21.29 (3.20), for HC 16.09 (3.22); $p < 0.001$) and PSWQ score (mean (SD) for MDD 61.83 (15.73), for HC 35.59 (10.01); $p < 0.001$).

We present in Fig. 1 the FA/MD scatter plot within skeletonized WM for HC and depressed samples, where more dispersive MD values were observed when the standard TBSS pipeline was used. In order to calculate the MD dispersion values, we used the ratio of the SD resulting from applying the TBSS pipeline and the SD associated with the ANTs pipeline. Values greater than one mean that the TBSS standard pipeline had more dispersive MD values relative to the ANTs TBSS method. Moreover, because the FA and MD values followed Gaussian distributions; we applied a F-test to compare if the variance ratio was greater than one. Only the TORTOISE *i*RESTORE tensor fitting algorithm rendered similar dispersive MD values for the ANTs TBSS analysis (ratio = 1.10, $p = 0.26$). This indicates that TBSS pipeline yielded more differences between HC and MDD in MD metrics than the ANTs TBSS pipeline.

Whole-WM-skeleton corrected voxel-wise analyses, shown in Fig. 2 and reported in Table 3, revealed three WM tracts with reduced FA in depressed patients as compared to HCs: left inferior