Psychiatry Research: Neuroimaging 253 (2016) 7–14

Contents lists available at ScienceDirect

Psychiatry Research: Neuroimaging

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

The effects of white matter disease on the accuracy of automated segmentation

Helmet T. Karim [a], Carmen Andreescu [b], Rebecca L. MacCloud [b], Meryl A. Butters [b], Charles F. Reynolds III[b], Howard J. Aizenstein [a,b], Dana L. Tudorascu [b,c,d,*]

[a] Department of Bioengineering, University of Pittsburgh, Pittsburgh, PA, USA
[b] Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA
[c] Department of Internal Medicine, University of Pittsburgh, Pittsburgh, PA, USA
[d] Department of Biostatistics, Graduate School of Public Health, Pittsburgh, PA, USA

ARTICLE INFO

Article history:
Received 20 November 2015
Received in revised form
14 April 2016
Accepted 16 May 2016
Available online 19 May 2016

Keywords:
Segmentation
MRI
FMRI
White matter hyperintensities

ABSTRACT

Automated segmentation of the brain is challenging in the presence of brain pathologies such as white matter hyperintensities (WMH). A late-life depression population was used to demonstrate the effect of WMH on brain segmentation and normalization. We used an automated algorithm to detect WMH, and either filled them with normal-appearing white-matter (NAWM) intensities or performed a multi-spectral segmentation, and finally compared the standard approach to the WMH filling or multi-spectral segmentation approach using intra-class correlation coefficients (ICC). The presence of WMH affected segmentations for both approaches suggesting that studies investigating structural differences in populations with high WMH should account for WMH. We also investigated how functional data contrasts are affected using normalization between the standard compared to fill and multi-spectral approach. We found that the functional data was not affected. While replication with a larger sample is needed, this study shows that WMH can significantly affect the results of segmentation and these areas are not limited to those affected by WMH. It is clear that to study gray matter differences that some correction should be made to account for WMH. Future studies should investigate which methods for accounting for WMH are most effective.

1. Introduction

White matter hyperintensities (WMH) are areas in the white matter of the brain that have higher intensity values on a T2-weighted magnetic resonance (MR) scans compared to normal appearing white matter (NAWM). Usually appearing in mid-to-late-life, WMH are attributed to degenerative changes of long penetrating arteries (Breteler et al., 1994), resulting in demyelination, gliosis and axonal degeneration (Ovbiagele and Saver, 2006). WMH can have a variety of pathologies and are associated with several neurological disorders (e.g. multiple sclerosis). In older populations, the prevalence of white matter disease is highly variable, with reports indicating it affects from 5% to 90% of the population, depending on the study design and study population (Xiong and Mok, 2011). WMH are more common in populations 60 years and older and also are more common in women than in men (de Leeuw et al., 2001).

Although WMH appear as hyperintense on T2-weighted images, they are hypointense (dark) on the T1-weighted images. These dark spots on the T1-weighted images are not typically considered when using automated image processing methods. For instance, studies exploring structural brain changes rely on tissue segmentation algorithms. These algorithms can be affected by the presence of WMH. The process of tissue classification in presence of WMH is often a challenging task. Thus, the presence of WMH can result in tissue misclassification if it is not taken into account in the segmentation algorithm, as it has been reported in Levy-Cooperman et al. (2008). Most neuroimaging software platforms have been developed for populations with healthy brains and therefore there is only a low risk of tissue misclassification since the prevalence of WMH is very low or non-existent in younger populations. Therefore, incorrectly classifying the voxels as gray matter due to lower intensities on the T1 is not a significant issue in healthy subjects.

Further, it is often necessary to coregister all subject volumes into a common template space. This step can be affected by the presence of WMH (Eloyan et al., 2014) as atrophy and WMH can

* Correspondence to: Western Psychiatric Institute and Clinic, 3811 O'Hara St., Pittsburgh, PA 15213, USA.
E-mail address: dlt30@pitt.edu (D.L. Tudorascu).

http://dx.doi.org/10.1016/j.pscychresns.2016.05.003
0925-4927/

interfere with proper classification of the brain tissue. This interferes with proper coregistration to template space, especially in algorithms that rely on a unified segmentation/normalization algorithm.

WMH filling (Battaglini et al., 2012; Chard et al., 2010; Eloyan et al., 2014; Magon et al., 2014; Sdika and Pelletier, 2009; Valverde et al., 2014) is a previously established method used to alleviate the effect of WMH on segmentation/coregistration algorithms. After identifying WMH (using an automated method), they are filled with intensities of NAWM in the structural image. This allows for these regions to be segmented as white matter instead of being misclassified (usually as gray matter). Another approach utilizes a multi-spectral segmentation by using information from a T1-weighted and T2-weighted image. A WMH prevalence map can be input to identify regions most greatly affected by WMH. This allows for classification of these tissues and models the WMH as a separate tissue class.

Proper classification of brain tissue is required in order for the statistical analysis to be valid in populations with age-specific diseases or other diseases that can severely affect brain tissue. In this study, we used a group of older adults enrolled in a study of late-life depression (LLD). As depression in older adults is associated with the presence of small vessel ischemic disease (Alexopoulos, 2006; Taylor et al., 2013; Tudorascu et al., 2014) this group is likely to have a significant burden of WMH, and thus is well-suited for testing how WMH burden interfere with tissue segmentation and registration. In general, aging populations tend to have higher WMH burden, but patients with LLD often have an even higher WMH burden compared to age-matched healthy controls (Herrmann et al., 2008). The main purpose of our work was to describe the effect of WMH on segmentation. We however did not investigate the efficacy of the methods used and only investigated the effect of some correction on the results. We also aimed to look at the effect of WMH on functional data coregistration/normalization to a standard (Montreal Neurological Institute, MNI) space. Thus, we examined (via intraclass correlation coefficients, ICC) the differences between the standard approach, which does not account for WMH, and the WMH filling approach as well as the multi-spectral segmentation approach. This will indicate the effect each method has on the previously uncorrected approach and does not indicate accuracy/efficacy.

2. Methods

2.1. Study design and subjects

The subjects participated in a five-year multi-site study of treatment of Late Life Depression (LLD) with an imaging component added to the Pittsburgh site (Lenze et al., 2015). The University of Pittsburgh IRB approved this study. Subjects were included in the LLD study if they were older than 65 years of age, had major depressive disorder that met DSM-IV criteria (First et al., 1997), and had Montgomery-Asberg Depression Rating Scale (MADRS) score greater than 15. Subjects had a mean (std.) adapted Edinburgh handedness index of 17 (6) (right-handed, 1 missing subject with self-reported right-handedness) (Oldfield, 1971). Subjects were excluded if they had a history of mania or psychosis, alcohol/substance abuse, current use of antidepressant or involvement in mental health treatment, dementia or any neurodegenerative disease, prescribed medications used in Alzheimer's disease, and medical conditions with known significant effects on mood (e.g. clinical stroke) and neurologic disorders with known effects on the brain (e.g., multiple sclerosis). After subjects completed informed consent, a baseline MRI scan was acquired. As part of the multi-site study, subjects underwent detailed neurocognitive assessment, a clinical interview, mood and anxiety assessment, and laboratory testing.

A total of 37 participants signed consent, but four were not included for the following reasons: withdrawal due to failure to communicate/failure to comply, or they were deemed (early on) to no longer have Major Depressive Disorder (MDD). Six were not included because they either did not have a structural MPRAGE or FLAIR (see data collection section). A total of 27 subjects were included in this analysis (mean age 67 (SD=6) years).

2.2. MRI/fMRI data collection

All scanning was conducted using a 3T Siemens Trio TIM scanner (VB17) with a 12-channel head coil (parallel imaging) located at the Magnetic Resonance Research Center at the University of Pittsburgh. The head was immobilized using vacuum pads to minimize motion artifacts. Multiple imaging sequences were acquired. A map used to identify WMH was collected, specifically an axial T2-weighted fluid-attenuated inversion recovery (FLAIR, TR=9160 ms, TE=90 ms, TI=2500 ms, FA=150°, FOV=256 × 212, slice thickness=3 mm, slices=48, matrix size=256 × 212, no gap). An axial T1-weighted 3D sequence was collected (TR =2300 ms, TE=3.43 ms, TI=900 ms, FA=9 deg, FOV=256 × 224, slice thickness=1 mm, slices=176, matrix size=256 × 224, no gap). An axial T2*-weighted BOLD interleaved acquisition using gradient-echo echoplanar imaging (EPI) was also collected (TR=2000 ms, TE=34 ms, FOV=256 × 256, slices=28, slice thickness=4 mm, FA=90°, matrix size=128 × 128, no gap, 117 time points) to measure changes in blood oxygen-level dependent (BOLD) response during the task (see next section). Total scan time was one hour. Structural scans covered from the bottom of the cerebellum to the top of the head, whereas functional scans only covered from midway of the cerebellum to the top of the head.

2.3. Faces/shapes task

The study used the faces-shapes task to study the effect of emotional responses to faces (Hariri et al., 2003). We have further used it to study the effect of WMH on functional coregistration due to its activation of some subcortical structures that may be affected by WMH. During the task, subjects are required to select one of two facial expressions (angry/fearful and neutral) that match that of a simultaneously presented target expression. As a control task, subjects matched one of two geometric shapes with a control shape. The matching task (5 blocks) was interleaved with 4 blocks of the experimental faces task. Each block lasted 24 s and contains 6 trials lasting 4 s each.

Before the beginning of each block, a brief instruction ("match emotion" or "match form") is presented for 2 s. On the faces trials, 12 different images are used, six per block, three of each gender, all derived from a standard set of pictures of facial affect (Ekman et al., 1975). In the control condition, six different sets of geometric forms are used, each set including two geometric shapes. During imaging, subjects responded with button presses, allowing for determination of accuracy and reaction time. Stimulus presentation and response recording were controlled using the E-Prime software package (Psychology Software Tools, Inc., Pittsburgh, 2002).

2.4. Data processing

All of our preprocessing used Statistical Parametric Mapping software (SPM 12) (http://www.fil.ion.ucl.ac.uk/spm/) as well as Python and Matlab programming platforms. The following sections are ordered temporally, thus each section precedes the following in terms of order of analysis.