NeuroImage 86 (2014) 81–90

Contents lists available at ScienceDirect

NeuroImage

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

Validity of modulation and optimal settings for advanced voxel-based morphometry

Joaquim Radua [a,b,c,*,1], Erick Jorge Canales-Rodríguez [a,c,1], Edith Pomarol-Clotet [a,c], Raymond Salvador [a,c]

[a] FIDMAG Research Unit, Barcelona, Spain
[b] Institute of Psychiatry, King's College London, London, UK
[c] Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Spain

ARTICLE INFO

Article history:
Accepted 30 July 2013
Available online 8 August 2013

Keywords:
Cluster-based statistics
Diffeomorphic registration
Modulation
Normalization
Threshold-free cluster enhancement
Voxel-based morphometry

ABSTRACT

Voxel-based morphometry (VBM) is a widely-used structural neuroimaging technique for comparing meso- and macroscopic regional brain volumes between patients and controls in vivo, but some of its steps, particularly the modulation, lack an experimental validation. The aims of this study were two-fold: a) to assess the effects of modulation to detect mesoscopic (i.e. between microscopic and macroscopic) abnormalities on published, classic VBM; and b) to suggest a set of potentially optimal settings for detecting mesoscopic abnormalities with new, advanced, high-resolution diffeomorphic VBM normalization algorithms. Sensitivity and false positive rate after modulating or not in classic VBM using different software packages and spatial statistics, and after setting a range of different parameters in advanced VBM (ANTS-SyN), were calculated in 10 VBM comparisons of 32 altered vs. 32 unaltered gray matter images from different healthy controls. Simulated brain abnormalities comprised mesoscopic volume differences mainly due to cortical thinning. In classic VBM, modulation was associated with a substantial decrease of the sensitivity to detect mesoscopic abnormalities ($p < 0.001$). Optimal settings for advanced VBM included the omission of modulation, the use of large smoothing kernels, and the application of voxel-based or threshold-free cluster enhancement (TFCE) spatial statistics. The modulation-related decrease in sensitivity was due to an increase in variance, and it was more severe in higher-resolution normalization algorithms. Findings from this study suggest the use of unmodulated VBM to detect mesoscopic abnormalities such as cortical thinning.

Introduction

One of the major applications of neuroimaging techniques into the investigation of neuropsychiatric disorders is the study of brain volumetric abnormalities using voxel-based morphometry (VBM) (Bora et al., 2011; Cooper et al., 2014; Fusar-Poli et al., 2012; Lansley et al., 2013; Nakao et al., 2011; Palaniyappan et al., 2012; Radua and Mataix-Cols, 2009; Radua et al., 2010, 2011, 2012a; Via et al., 2011).

VBM algorithms usually consist of the following steps: a) *segmentation* of the structural images into different brain tissues in order to obtain a map of the voxelwise probability of gray matter for each individual; b) macroscopic *normalization* of the individual gray matter images to a standard brain template so that the gray matter maps from all individuals can be approximately superimposed; c) *modulation*, in which normalized gray matter maps are scaled by the macroscopic transformations to preserve local volumes; d) *smoothing* of the modulated images so that each voxel contains information from its neighboring voxels; and e) voxel-based or cluster-based spatial statistics to compare patients with controls. Originally, the modulation step was not included in VBM pipelines, which only aimed to detect mesoscopic (i.e. between microscopic and macroscopic) regional abnormalities such as cortical thinning (Ashburner and Friston, 2001). Larger, macroscopic regional volumetric differences were removed during the normalization but could be detected with other techniques such as tensor-based morphometry (TBM) (Good et al., 2001a). Thus, volumetric abnormalities were partitioned into mesoscopic differences detectable with VBM and macroscopic differences detectable with TBM. With the subsequent incorporation of the modulation step into the VBM pipelines, this perspective changed and VBM was thought to detect both meso- and macroscopic regional abnormalities.

Different VBM algorithms have been reported to yield different results, urging the establishment of a standardized set of optimal VBM settings in order to increase methodological uniformity and validity (Borgwardt et al., 2012; Henley et al., 2010). Some of the differences between different VBM methods have been already investigated, especially in segmentation and normalization, (Acosta-Cabronero et al., 2008; Ashburner, 2007; Ashburner and Friston, 2011; Avants et al., 2008, 2010; Fein et al., 2006), leading to an improvement of the VBM

* Corresponding author at: FIDMAG Research Unit, Benito Menni CASM, Carrer d'Antoni Pujadas 38, 08830 Sant Boi de Llobregat, Barcelona, Spain. Fax: +34 93 640 02 68.
E-mail address: joaquim.radua@kcl.ac.uk (J. Radua).
[1] The first two authors contributed equally to this work.

1053-8119/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2013.07.084

pipelines included in standard neuroimaging packages such as SPM (Wellcome Trust Centre for Neuroimaging, London) or FSL (Centre for Functional MRI of the Brain, Oxford). Also, the optimal full width at half maximum (FWHM) of the smoothing kernel in classic VBM has been reported to be larger than 8 mm (Salmond et al., 2002), and cluster-based spatial statistics (Bullmore et al., 1999) have been shown to bias the results towards some regions of the brain (Good et al., 2001b; Hayasaka et al., 2004; Mechelli et al., 2005; Worsley et al., 1999).

In contrast, a proper experimental validation of modulation is still lacking. Subjective impressions have been published, pointing to an adequate effect in terms of well localized and anatomically relevant findings in a set of exploratory studies. But as far as we know, there exists no validation based on a true knowledge of the location of the abnormalities and on statistical analyses which allow inferring and predicting performance parameters such as sensitivity or false positive rate under different conditions. This lack of experimental validation may be related to the apparent straightforwardness of its logic: if a brain region is artificially enlarged during the normalization to the standard brain template, the value of its voxels should be proportionally reduced to ensure that the overall volume of the region in the normalized image is the same as in the original image. The gray matter probability of a given region in the original image may be considered a measure of volume, e.g. a value of 0.6 would indicate that 60% of the voxel volume is gray matter, so that if each voxel is 1 mm^3 (i.e., 1 μl), the voxel would contain 0.6 μl of gray matter. During normalization, some regions are expanded and others are contracted, with these volumetric changes being captured by the Jacobian determinants of the transformation. As exemplified in Fig. 1, a Jacobian determinant of 0.86 would mean that the brown voxel in Montreal Neurological Institute (MNI) space corresponds to 1/0.86 = 1.16 blue voxels in native space. *Normalization* would have reduced these bits of the brain by 1.16, so that if they overall contained 4.8 μl of gray matter, they would now only contain 4.1 μl. Subsequent *modulation* would consist in multiplying the gray matter value of the MNI voxel by 1.16 in order to restore the original volume (4.8 μl). This correction would be independent of the accuracy and precision of the normalization, i.e. it would not matter whether the normalization was anatomically correct or not, or whether it was low- or high-resolution, as far as the expansion or shrinkage applied to each voxel was completely captured by the corresponding Jacobian determinant.

There may be other hints to explain why modulation has been so widely accepted by neuroimaging researchers despite its lack of experimental validation. First, analyses of modulated were thought to test for regional *volume* differences, while analyses of unmodulated data to *only* test for regional *concentration* differences (Ashburner and Friston, 2000; Good et al., 2001b; Mechelli et al., 2005). Second, modulation was included as one of the main steps of the so-called 'optimized VBM' (Good et al., 2001b). And last but not least, modulated analyses yielded a set of neurologically interesting results which were not detected in the unmodulated analyses (Good et al., 2001b; Keller et al., 2004).

The present study was designed to: a) assess the effects of modulation on published, classic VBM analyses; and b) suggest a set of potentially optimal settings for new, advanced VBM algorithms such as 'advanced normalization tools' (ANTS) (Avants et al., 2008, 2010; Klein et al., 2009) or 'diffeomorphic anatomical registration through exponentiated Lie algebra' (DARTEL) (Ashburner, 2007; Ashburner and Friston, 2011) in order to provide hints for future studies. From an initial pool of 128 images from healthy individuals, a re-sampling scheme was used to generate 10 independent pairs of samples. For each pair, images from one of the samples (*simulated patients*) were artificially modified to simulate mesoscopic abnormalities (mainly cortical thinning), while the other sample (*controls*) was left unmodified. Next, comparisons between both samples were conducted using classic and advanced VBM with different sets of settings (e.g. different software packages, modulating or not, and etcetera). Finally, sensitivity, false positive rate and other performance outcomes were used to assess the effects of modulation on classic VBM, as well as to suggest optimal settings for advanced VBM.

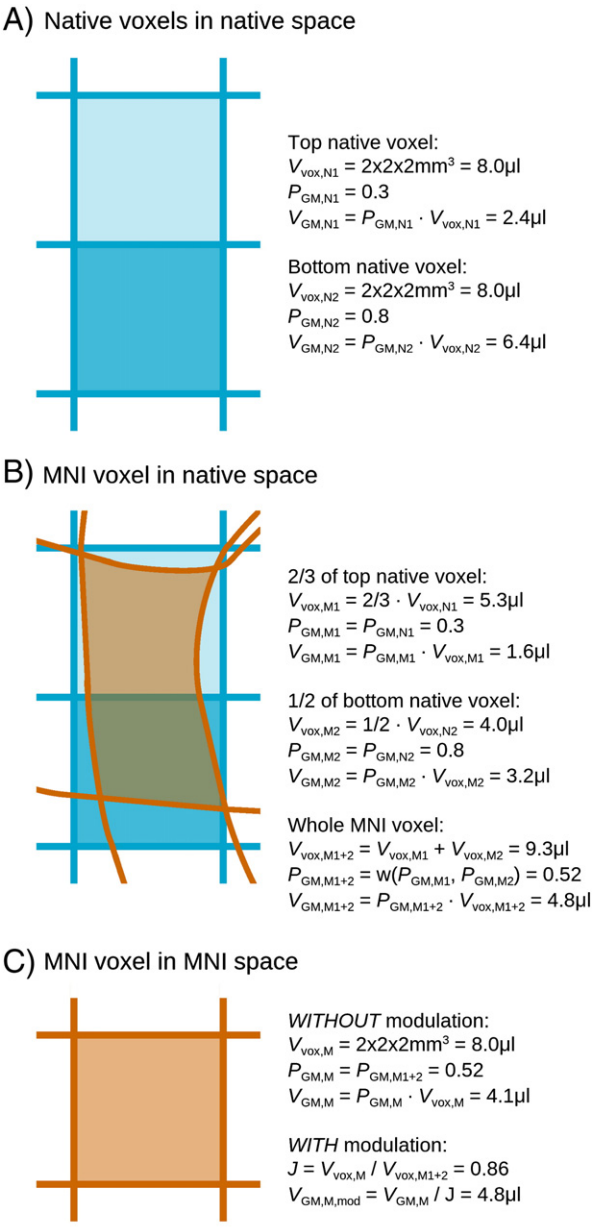

Fig. 1. Distortion and restoration of regional volumes during normalization and modulation. The voxels in native space have been colored blue. The voxel in Montreal Neurological Institute (MNI) space has been colored brown.

Material and methods

Acquisition of original gray matter maps

Structural brain images for this study were acquired from 128 right-handed healthy adults (64 women, 37.8 ± 11.3 years) who did not report a personal or family history of mental illness, alcohol or drug abuse, and/or treatment with psychotropic medication. Participants were scanned in a 1.5-T GE Signa scanner (General Electric Medical Systems, Milwaukee, WI, USA) with the following acquisition parameters: T1-weighted sequence, 180 axial slices, 1 mm slice thickness with no gap, 512 × 512 matrix size, 0.5 × 0.5 × 1 mm^3 voxel