Published in final edited form as:

J Neurosci Methods. 2014 April 15; 226: 57–65. doi:10.1016/j.jneumeth.2014.01.012.

The efficiency of fMRI region of interest analysis methods for detecting group differences

Joanna L. Hutchison^{1,2}, Nicholas A. Hubbard¹, Ryan M. Brigante¹, Monroe Turner¹, Traci I. Sandoval¹, G. Andrew J. Hillis³, Travis Weaver¹, and Bart Rypma^{1,2}

¹School of Behavioral and Brain Sciences, University of Texas at Dallas, Richardson, TX

²Department of Psychiatry, University of Texas Southwestern Medical Center, Dallas, TX

³Department of Psychology, University of Texas Southwestern Medical Center, Dallas, TX

Abstract

Background—Using a standard space brain template is an efficient way of determining region-of-interest (ROI) boundaries for functional magnetic resonance imaging (fMRI) data analyses. However, ROIs based on landmarks on subject-specific (i.e., native space) brain surfaces are anatomically accurate and probably best reflect the regional blood oxygen level dependent (BOLD) response for the individual. Unfortunately, accurate native space ROIs are often time-intensive to delineate even when using automated methods.

New Method—We compared analyses of group differences when using standard versus native space ROIs using both volume and surface-based analyses. Collegiate and military-veteran participants completed a button press task and a digit-symbol verification task during fMRI acquisition. Data were analyzed within ROIs representing left and right motor and prefrontal cortices, in native and standard space. Volume and surface-based analysis results were also compared using both functional (i.e., percent signal change) and structural (i.e., voxel or node count) approaches.

Results and Comparison with Existing Method(s)—Results suggest that transformation into standard space can affect the outcome of structural and functional analyses (inflating/minimizing differences, based on cortical geography), and these transformations can affect conclusions regarding group differences with volumetric data.

Conclusions—Caution is advised when applying standard space ROIs to volumetric fMRI data. However, volumetric analyses show group differences and are appropriate in circumstances when time is limited. Surface-based analyses using functional ROIs generated the greatest group

© 2014 Elsevier B.V. All rights reserved.

Corresponding Author: Joanna L. Hutchison, Center for Brain Health, University of Texas at Dallas, 2200 W. Mockingbird Lane, Dallas, TX 75235, Office Phone: 972-883-3253, Fax: 972-883-3231, joanna.hutchison@utdallas.edu.

This paper includes data that were accepted in a peer-reviewed abstract and presented via electronic poster at the Joint Annual Meeting of the International Society for Magnetic Resonance in Medicine and the European Society for Magnetic Resonance in Medicine and Biology, Stockholm, Sweden, May 2010.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

differences and were less susceptible to differences between native and standard space. We conclude that surface-based analyses are preferable with adequate time and computing resources.

Keywords

functional Magnetic Resonance Imaging (fMRI); native space; standard space; Region-of-Interest (ROI); group differences

1. Introduction

The focus of the present study was to ascertain whether significant group-level differences occur when a single data set is processed using different processing pipelines. We used both standardized anatomical brain space (i.e., standard space) analyses and subject-specific anatomical brain space (i.e., native space) analyses to define ROIs. We then performed group-level assessments using volume- and surface-based analyses. Using the sulci and gyri of a standardized atlas (e.g., Talairach and Tournoux, 1988), the standard space technique transforms an individual's data from native space to the specified atlas' three-dimensional coordinate space. A typical standard space technique using surface-based analyses creates a standardized mesh based on the cortical surfaces of the participants in the study (e.g., Saad and Reynolds, 2012). Alternatively, the individual's native space can be used to delineate structural ROIs by the definition of specific anatomical landmarks (e.g., sulci and gyri, such as the Calcarine Sulcus and the Pars Triangularis; cf. Fischl et al., 2008).

Both standard and native space methods ostensibly allow the researcher to ensure that functional activity is not erroneously assigned to extra-ROI regions due to individual differences in cortical architecture. Regrettably, standard space procedures have a high probability of incorrectly labeling voxels due to the heterogeneity of cortical folds across subjects. It has been assumed that matching an individual's functional data to that individual's own anatomy would allow for the most accurate analysis (e.g., Fischl et al., 2008; Siebert et al., 2011; Siebert et al., 2012). However, this method of analysis is often more labor- and time-intensive than using a standardized brain atlas, even with the availability of automated or semi-automated registration (e.g., Fischl et al., 2004; Desikan et al., 2006). ROIs are typically used to test hypotheses regarding the functions of various brain regions, but the functions of these regions can differ between groups (e.g., younger versus older, healthy versus clinical populations). Thus, accurate estimates of blood oxygen level dependent (BOLD) activity within ROIs are critical to inferences regarding functional group differences.

Further, it is possible that native and standard space approaches are differentially affected by the use of volume and surface-based analyses. It has recently been demonstrated (e.g., Tucholka et al., 2012) that surfaced-based group analyses are more sensitive than analogous volume-based analyses. Such surfaced-based analyses have been shown to improve reliability over volume-based analyses, particularly for suprathreshold data. However, surface-based analyses involve a burden in terms of time and resources, and there is some risk that the data are not properly projected to the surface (cf. Fischl et al., 1999; Operto et al., 2008; Greve and Fischl, 2009). Thus we sought to compare group differences in BOLD

activity estimates in standard and native space using both volume and surface-based analyses.

Numerous functional magnetic resonance imaging (fMRI) studies have used Brodmann Area (BA; Brodmann, 1909/2006) ROIs in their analyses (see Rypma, 2006; e.g., Fischl et al., 2008). Yet a fundamental problem in defining ROIs on the individual's cortical surface is that of adjudicating a voxel to the appropriate ROI in the absence of cytoarchitectural information. One common way to delineate areas of cortex is to rely on sulcal and gyral folding. Fischl et al. (2008) demonstrated that folding is accurate, stable, and provides histologically appropriate BA demarcations. Hence, a folding-based alignment in native space has the potential to account for significant anatomical variability. Implementing native space folding alignment could therefore yield improved accuracy in determining ROIs across the cortex compared to a standard space alignment method that does not account for individual variability in folding.

In the present study, we evaluated native and standard space data using both volume- and surface-based approaches (see Figure 2). Anatomical motor cortex (MC) and prefrontal cortex (PFC) BA ROIs were defined based on anatomical landmarks in each individual's native space and also in standard space. Functional ROIs were generated by using only suprathreshold voxels within the anatomical ROIs. Results were compared using functional (i.e., BOLD percent signal change) and structural (i.e., voxel or node count) approaches for young, healthy collegiate and older military-veteran participants.

2. Materials and methods

2.1. Participants

Twenty-three healthy participants were recruited from the University of Texas at Dallas (UTD; ages 20-39, $M=24\pm4$ years, 11 female). Forty-seven participants (ages 40-73, $M=58\pm8$ years, 0 female) were recruited from a group of military-veterans formerly deployed to the 1991 Persian Gulf War; military-veterans' participation was included as part of a larger study. Approximately half of the military-veterans were diagnosed with Gulf War Illnesses (GWI; $n=23$; for neural sequelae of GWI see Hubbard et al., 2013), whereas the remainder of the military-veterans reported being healthy ($n=24$). All participants gave written informed consent. All procedures were approved by the UTD and the University of Texas Southwestern Medical Center (UTSWMC) Institutional Review Boards.

2.2. The fMRI tasks

Participants performed a button press task (BPT), in which they were required to press left and right hand buttons simultaneously when a black and white annulus flashed on the screen (i.e., every 18 seconds). This task reliably elicits BOLD activity within motor cortex (e.g., Rypma et al., 2005). The task sequence began with an 18 s period to reach steady-state magnetization followed by a single run of 20 trials (360 s).

Participants also performed the digit-symbol verification task (DSVT; Rypma et al., 2006). Each trial of the DSVT contained a key consisting of number-symbol pairings toward the top of the screen and a single number-symbol probe toward the bottom of the screen.

Participants were allotted 3.5 s to determine whether the number-symbol probe matched the corresponding number-symbol pair in the key. They were instructed to press the button located in their right hand if there was a match, or the button located in their left hand if there was not a match. 'Left=No' and 'Right=Yes' were continuously presented at the lower left and right hand corners of the screen, respectively. The task sequence began with an 18 s period to reach steady-state magnetization followed by 4 s trial events with a variable inter-trial interval (0, 4, 8, or 12 s). There were 3 runs of 52 trials (6 min 18 s per run), for a total of 156 trials.

2.3. Scanning parameters

All magnetic resonance imaging data were acquired on a Siemens 3 Tesla magnet equipped with a 12-channel receiver array head coil and body transmit coil. An fMRI BOLD echo planar imaging pulse sequence was used to acquire BPT and DSVT functional data (TE= 20 ms, TR=2000 ms, flip angle=90 degrees, trans-axial plane, head-to-foot acquisition, matrix 74×74, slices=44, voxels=3×3×3.5 mm³; BPT volumes=189; DSVT volumes=159). A high-resolution magnetization-prepared rapid acquisition with gradient echo (MPRAGE) 3D volume image was also acquired for each participant (TE=3.31 ms, TR= 2250 ms, flip angle=12 degrees, sagittal plane, left to right acquisition, matrix=256×256, slices=160, slice thickness=1 mm, voxels=1 mm³).

2.4. Functional processing

AFNI software (Cox, 1996) was used to preprocess the functional data. Data for individual participants were corrected for slice-timing offset and motion, and were spatially filtered with an 8 mm full width half maximum Gaussian kernel. The data for each voxel and run were scaled by the mean for that voxel and run, yielding parameter estimates expressed in terms of percent signal change (i.e., $100 \times y_t/M_y$, t = time point). A regressor was constructed for each task by convolving a hemodynamic response model (a gamma-variate function using Cohen parameters; $b = 8.6$, $c = .547$, maximum amplitude=1.0) with each trial onset in a task-reference function (Cohen et al., 1997). For each run, regressors for motion correction estimates and linear, quadratic, and cubic trends were included in the baseline regression model. Data were exported into the SAS® software program (Version 9.1, SAS Institute, Cary, NC) for further statistical analysis.

2.5. Individual BA ROI generation

All image data were processed in Neuroimaging Informatics Technology Initiative 1.1 format (NIFTI-1.1) for portability across platforms (<http://nifti.nimh.nih.gov/nifti-1/>). Pial and white matter surface reconstructions for each hemisphere were obtained from the MPRAGE using Freesurfer's (<http://surfer.nmr.mgh.harvard.edu/>) "recon-all" (i.e., reconstruction) shell script in a Fedora Unix environment. Inflated surfaces were visually inspected for white matter deletions, pial deletions, and abnormalities, and control points were added to include white or gray matter not automatically included in the surfaces; dura and skull were manually deleted from the pial surface as necessary. The Freesurfer reconstruction algorithm was re-applied and surfaces were visually re-inspected. After two of the authors (T.S. and G.A.J.H.) concurred that each surface matched its respective

MPRAGE and that all white matter and pial layers were intact, these surfaces were used to create a mid-thickness (i.e., midpoint between pial and white matter) surface in Caret (Van Essen et al., 2001) using the Caret “surface-average” algorithm.

Caret software automatically generated three landmarks prior to flattening: Medial Wall Dorsal, Medial Wall Ventral, and Calcarine Fissure. These landmarks were manually adjusted to better reflect each individual’s anatomy. These landmarks were then used as cuts to flatten the brain and became the outer edges of the flattened surface. Three additional Caret-recommended landmarks were manually drawn on the flattened surface: the Central Sulcus, the Sylvian Fissure, and the anterior half of the Superior Temporal Gyrus (labeled aSTG; see Figure 1; Van Essen, 2005). In order to correct for inconsistencies in the locations of BAs and their alignment with gyral and sulcal landmarks, we manually added five landmarks (3 sulci and 2 gyri) that we determined to be consistently identifiable and accurate in determining BAs across participants (see Figure 1): Inferior Rostral Sulcus, Pars Triangularis, Pars Orbitalis, Inferior Frontal Sulcus, and Superior Temporal Sulcus. These additional landmarks were particularly helpful in localizing PFC BAs (cf. Fischl et al., 2008, regarding variability in frontal folding).

Surfaces were then spherically registered to the standard surface template using Caret software, which includes algorithms to account for variations in sulcal depth. The surfaces were expanded into a sphere and the landmarks from the individuals were aligned with the landmarks defined on the standard surface sphere. BAs were derived for each hemisphere by employing the Caret spherical registration process and manual adjustments were accomplished using the Caret paint tool. Two authors (T.S. and G.A.J.H.) systematically verified and agreed that surfaces and BAs were created without flaws or abnormalities. Volume maps were created in Caret and were then processed using in-house MATLAB® (Version 7.4, Mathworks, Natick, MA) code to reassign arbitrarily-assigned Caret-generated mask values to reflect standard BA numerical designations. ROI masks for BAs in left and right motor cortex (MC; as defined by area 4) and left and right prefrontal cortex (PFC; as defined by areas 8, 9, 10, 11, 44, 45, 46, and 47) were generated using AFNI by down-sampling the data to functional space and then applying the mask to the native space functional data. Analyses for this paper were restricted to data within left and right MC and PFC ROIs.

Time-intensiveness and required labor are concerns in analysis pipeline development. In our pipeline, the entire process of generating individual ROIs was typically completed within two working days if Freesurfer and CARET processes did not encounter problems. Complex issues were often encountered; for older military-veteran participants, the process often took up to four working days. The most pervasive problem that contributed to extended processing time was poor separation of older participants’ (i.e., > 50 years) white matter and pial surfaces. Computing power available versus the amount of computing power required to process each brain was also a factor limiting the rate at which data could be analyzed. Our Fedora Linux-based operating system with an Intel® Xenon® (W5590 @ 3.33 GHz) processor permitted us to process about three or four brains at the same time on a single Linux drive. Because of the iterative nature of the processing pipeline when problems were encountered, several brains required multiple processing attempts.

Identification of anatomical landmarks for Freesurfer processing was also a rate limiting step. Due to anatomic variability, there were individual brains for which landmarks were difficult to discern, often resulting in two to three extra hours of labor. Uninterrupted, this portion of the process required two to three hours for Freesurfer to complete—that is, if the brain was being processed alone on the Linux machine—and then an additional two to three hours were required for Caret to complete processing.

2.6. Standard space volume BA ROI generation

In a procedure parallel to the individual BA ROI generation, the AFNI (Cox, 1996) Colin TTN27 template was spherically registered to the Caret Colin template (Van Essen, 2002) to create volume maps of BAs for standard space. ROI masks for BAs in left and right MC and PFC were generated using AFNI, down-sampled to functional space, and applied to the standard space-transformed functional data.

2.7. Processing for surface-based analyses

A cortical surface was constructed for each subject using the same recon-all reconstruction algorithm in Freesurfer; however, in this instance we used the surfaces as generated by Freesurfer, without additional corrections. These surfaces consisted of meshes with edges and nodes based on the geometry and topology of each individual's anatomical brain images. The time required to run the full processing (i.e., recon-all) was approximately 20 hours per subject using a Fedora Linux-based operating system with an Intel® Xenon® (W5590 @ 3.33 GHz) processor. This yielded a total time of about 1400 hours to create these surfaces, which we were not able to create concurrently due to the computer set-up in our laboratory. These surfaces were then aligned with re-sampled anatomical images in SUMA (Surface Mapping, a program that interacts with AFNI; Saad and Reynolds, 2012). Left and right anatomical and functional masks were generated for MC and PFC. These masks were applied to functional images first, and then they were mapped onto white matter and pial surfaces in SUMA.

A spherical template was used to create a standardized mesh for the standard surface model in SUMA. This standard model coordinate system was then used to recreate each participant's standardized cortical surface such that coordinates across participants shared the same localization across the standard space (cf. Saad and Reynolds, 2012). Node-mapping using this methodology has been demonstrated as 99.9% accurate in terms of alignment to the native surface (Saad et al., 2004).

2.8. Statistical analyses

Data were exported out of their respective ROIs and were analyzed using SAS software. Using a functional approach, beta values from the fMRI regression analyses were scaled to reflect percent signal change as outlined in section 2.4. Mixed model analyses (cf. Chen et al., 2013) were then used to determine the effects of our data treatments (analysis method: volume/surface, space: native/standard, ROI: anatomical/functional, and two-way interactions of these variables) on percent signal change values. Group (collegiate/military-veteran), health (healthy/GWI) and hemispheric (left/right) differences were assessed in light of the various data treatment approaches. Using a structural approach, the total number

of voxels or nodes within left and right MC and PFC regions were likewise compared over the ROIs. Age was also included in the models as a covariate. Hemispheric differences were very limited in scope and those results are therefore not presented.

3. Results

3.1. Percent Signal Change Comparisons—A Functional Approach

Comparisons of percent signal change are the most common assessment when comparing fMRI data between groups. Differences in percent signal change are often assumed to reflect differences in neural activity between groups, even though this is not always the case. Factors such as the processing of the data, as assessed in this paper, and other factors beyond the scope of this paper such as perfusion (e.g., Ances et al., 2008; Hutchison et al., 2013a,b) can potentially impact percent signal change values.

As seen in Figure 3, an interaction between data type and ROI type indicated greater percent signal change with data from the functional ROI compared to data from the anatomical ROI. This effect was greater on the surface but was also significant in the volume analyses (Figure 3a,b; PFC $F(1, 1016) = 1434.19, p < .0001$; MC $F(1, 959) = 740.86, p < .0001$). The greatest percent signal change within both PFC (Figure 3a) and MC (Figure 3b) was obtained using surface analyses with a functional ROI. Within PFC, standard and native space analyses were equivalent on the surface. However, standard space analysis yielded greater percent signal change relative to native space analysis for volume analyses (Figure 3c). In terms of main effects, surface percent signal change was greater than volume percent signal change (Figure 3d) and percent signal change within the functional ROI was greater than percent signal change within the anatomical ROI (Figure 3e).

As seen in Figure 4, group differences in percent signal change were affected by brain region. Within PFC, collegiate participants showed greater percent signal change than military-veterans (Figure 4a). Within MC, main effects of group (collegiate/military-veteran) and health (healthy/GWI) were only marginally significant (MC collegiate least squared mean (LSM) = $0.5389 \pm$ standard error 0.0696 , MC military-veteran $LSM = 0.3744 \pm 0.0372$, difference estimate (de) = $0.1645 \pm 0.0864, t(62) = 1.90, p = .062$; MC GWI $LSM = 0.5288 \pm 0.0583$, MC healthy $LSM = 0.3845 \pm 0.0432, de = 0.1443 \pm 0.0745, t(45) = 1.94, p = .059$).

Within both PFC and MC, there were significant interactions between ROI type and group (PFC $F(1, 783) = 7.87, p = .005$; MC $F(1, 514) = 6.71, p = .010$). Specifically, within PFC, percent signal change using the anatomical ROI was positive for collegiate participants, whereas it was negative for military-veterans (Figure 4b). However, percent signal change using the functional ROI was only marginally greater for collegiate participants than for military-veterans (Figure 4b). In contrast, within MC, the anatomical ROIs yielded no differences between collegiate and military-veteran groups (Figure 4c). However, using the functional ROI, collegiate participants evidenced greater percent signal change than military-veteran groups (Figure 4c).

Group and health also interacted with data type in MC such that differences were more pronounced for these variables with surface analyses than with volume analyses (Group $F(1,882) = 14.83, p = .0001$; Health $F(1,692) = 11.17, p = .0009$). Specifically, collegiate participants had greater percent signal change than military-veteran participants with surface analyses, but this difference was not significant with volume analyses (Figure 4d). Likewise, GWI participants had greater percent signal change than healthy participants with surface analyses, but this difference was also not significant with volume analyses (Figure 4e).

3.2. Voxel and Node Count Comparisons—A Structural Approach

Significantly different voxel or node counts using standard versus native space and volume-versus surface-based analyses could lead to disparate interpretations regarding group differences. Specifically, counts of voxels or nodes within anatomical ROIs index brain volume within the given brain region; in contrast, counts of voxels or nodes within functional ROIs index the spatial extent of activation within the given brain region. This structural approach has been successfully used as a metric of the diffusion of functional activation across cortex (e.g., Kami et al., 1995; Di et al., 2013), and it is potentially beneficial in teasing apart differences between groups when signal-to-noise ratios might differ in one group compared to another (cf. Saad et al., 2003). To simplify interpretation, we conducted volume (voxel) and surface (node) data analyses separately.

As seen in Figure 5, native space ROIs subsumed greater spatial extent (i.e., more voxels or nodes) than standard space ROIs (Figure 5a,b). Because we defined our functional ROIs as subsets of the anatomical ROIs containing suprathreshold voxels, it was not surprising that anatomical ROIs contained significantly more voxels or nodes than functional ROIs (all $ps < .0001$). More interestingly, there was an interaction between space and ROI type such that larger differences occurred between anatomical and functional spatial extent in native space than occurred between anatomical and functional spatial extent in standard space (Figure 5c,d,e,f; PFC $F(1,1003) = 52.97, p < .0001$; MC $F(1,475) = 56.23, p < .0001$). In fact, when using nodes to assess spatial extent, there were significantly more nodes within the functional mask in native space, but not in standard space (Figure 5e,f).

As seen in Figure 6, group differences in the extent of activation within the volume analyses were affected by brain region and method of assessment (voxels or nodes). Within PFC, collegiate participants had a greater extent of activation than military-veterans using either voxels or nodes (Figure 6a). With nodes in particular, this effect was greater with standard space compared to native space data (Figure 6b); this effect was also greater with functional compared to anatomical ROIs (Figure 6c).

Results were more varied within MC. Standard space voxel analyses indicated that military-veterans and collegiate participants had a similar extent of activation, whereas native space voxel analyses indicated that military-veterans had a greater extent of activation than collegiate participants (Figure 6d). Using voxel analyses, military-veteran and collegiate participants had equivalent extent of activation with functional ROIs, but military-veterans had larger anatomical ROIs than their collegiate counterparts (Figure 6e). Using node analyses, military-veteran and collegiate participants had similar extents of activation with functional ROIs, as well as similarly sized anatomical ROIs (Figure 6g). Additionally, when

using node analyses there was a significant interaction between space and group ($F(1, 446) = 6.46, p = .011$; using difference estimates as means from the following two contrasts yields Cohen's $d = .436$). That is, within native space, military-veterans exhibited marginally greater spatial extent than collegiate participants; however, within standard space, collegiate participants exhibited marginally greater spatial extent than military-veterans (Figure 6f).

4. Discussion

In the present study, participants completed a BPT and a DSVT during fMRI acquisition. Functional data were analyzed in native space and transformed into standard space for comparative analysis using both volume- and surface-based analyses. BA ROIs were used to define left and right ROIs within MC and PFC. Mixed effects models evaluated whether the various processing pipeline choices affected signal/extent in the assessment of group differences (collegiate and military-veteran, healthy and GWI). Our results suggest that the method of analysis can be critical in the assessment of group differences.

Using a functional approach, signal intensity was greater on the surface and when using functional ROIs. This processing combination maximized the chance of detecting group differences. The use of native or standard space in this scenario resulted in similar outcomes.

Using a structural approach, we found differences between brain regions. Anatomical PFC ROIs for the collegiate group were larger (i.e., contained more voxels or nodes) than those of the military-veteran group, indicating that atrophy might be a factor within PFC for the military-veteran group (a correction to minimize spatial variance would be appropriate in this case, but was beyond the scope of the present study; see Di et al., 2013). In contrast, anatomical MC ROIs for the military-veteran group were larger than those of the collegiate group, ruling out atrophy as a cause for group differences in functional extent within MC. In PFC, standard space analyses detected more group differences, whereas in MC, native space analyses detected more group differences. In both PFC and MC, surface-based analyses detected more group differences than did volume-based analyses, and collegiate participants evidenced a greater extent of functional activation than their military-veteran counterparts.

One possible cause for the observed differential effects of volume-based versus surface-based analyses is that much of the activity that is detected by fMRI technology comes from the surface, thereby rendering surface-based approaches more sensitive than their volume-based counterparts (cf. Tucholka et al., 2012). Using volume-based analyses yielded different results within different cortical regions. Our volume analyses suggested that we were more successful in detecting group differences when using a standard space approach within cortical regions exhibiting greater levels of anatomical variability, such as PFC (cf. Fischl et al., 2008). However, we were more successful in detecting group differences when using a native space approach within regions exhibiting less variability, such as MC (cf. Fischl et al., 2008). It might be that optimal decisions regarding analysis technique involve consideration of the brain regions under investigation.

Consistent with Tucholka et al. (2012), our overarching pattern of results suggested that surface-based analyses using functional ROIs are preferable for detecting group differences

in percent signal change, assuming that an appropriate projection to the surface can be obtained. However, surface generation can be time intensive, even when using automated methods. When using volume analyses, the use of native space was more important. Although this, too, requires greater processing time than simply using an automated standard space transformation. Our results converge with other studies indicating that the most time efficient, automated methods might not be the most accurate. For example, Kennedy et al. (2009) compared voxel-based morphometry (VBM) to manual volumetry and found VBM to provide realistic estimates in regional gray matter of anatomically defined areas in older individuals, but overestimates were found for peak volume. Kennedy et al. (2009) recommended using VBM as a first-pass strategy followed by manual measurement of anatomical areas for greater accuracy; they suggested that the VBM method, as often practiced by investigators, is less than ideal. Likewise, our results indicate that volumetric standard space analyses, as often implemented by researchers, are weaker than other analysis procedures in determining group differences.

Our results suggest that it is reasonable to use a volumetric standard space atlas for initial analyses of fMRI data. However, our results and those of others (e.g., Kennedy et al., 1998) suggest that although it is more time-efficient to transform subjects' data into standard space, this method of normalization and ROI circumscription can affect the outcome of statistical analyses, from which group differences could be minimized or inflated. For instance, we observed greater percent signal change in standard compared to native space volume analyses within PFC (Figure 3c). However, we found greater group differences in extent with native space compared to standard space volume analyses within MC (Figure 6d).

Caution should be exercised if using standard space ROIs for analyzing volumetric fMRI data due to potential distortion of the results. We suggest that this is particularly important if special populations are being examined using group comparisons. More research is certainly needed investigating methods that permit simultaneous volume and surface-based analyses incorporating more accurate ROI parcellation (e.g., deReus and van de Heuvel, 2013; Glasser et al., 2013; Wig et al, 2013). The use of semi-automated methods in native space (or more ideally, on the surface) is suggested as a judicious approach in analyzing fMRI data. This method is not as labor-intensive as drawing ROIs manually on individual brain surfaces, but it offers greater opportunity to detect group differences compared to the transformation of functional data into standard space.

Acknowledgments

This work was supported by VA IDIQ contract number VA549-P-0027 awarded and administered by the Dallas, TX VA Medical Center, by NIH (NCRR) Grant Number UL1RR024982, and by DOD grant no. DAMD 17-01-1-0741 from the U.S. Army Medical Research and Materiel Command. The content of this paper does not necessarily reflect the position or the policy of the U.S. government, and no official endorsement should be inferred. This research was also funded in part by NIH grant 1R01AG029523 (BR).

References

- Ances BM, Leontiev O, Perthen JE, Liang C, Lansing AE, Buxton RB. Regional differences in the coupling of cerebral blood flow and oxygen metabolism changes in the response to activation: implications for BOLD-fMRI. *Neuroimage*. 2008; 39:1510–1521. [PubMed: 18164629]
- Brodmann K. Brodmann's localization in the cerebral cortex: The principles of comparative localization in the cerebral cortex based on cytoarchitectonics. Garey, LJ., translator. New York: Springer; 2006. Original work published 1909
- Chen G, Saad ZS, Britton JC, Pine DS, Cox RW. Linear mixed-effects modeling approach to FMRI group analysis. *Neuroimage*. 2013; 73:176–190. [PubMed: 23376789]
- Cohen JD, Perlstein WM, Braver TS, Nystrom LE, Noll DC, Jonides J, Smith EE. Temporal dynamics of brain activation during a working memory task. *Nature*. 1997; 386:604–608. [PubMed: 9121583]
- Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res*. 1996; 29:162–173. [PubMed: 8812068]
- de Reus MA, van de Heuvel MP. The parcellation-based connectome: limitations and extensions. *Neuroimage*. 2013; 80:397–404. [PubMed: 23558097]
- Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany RJ. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*. 2006; 31:968–980. [PubMed: 16530430]
- Di X, Kannurpatti SS, Rypma B, Biswal BB. Calibrating BOLD fMRI activations with neurovascular and anatomical constraints. *Cer Cortex*. 2013; 23:255–263.
- Fischl B, Rajendran N, Busa E, Augustinack J, Hinds O, Yeo BTT, Mohlberg H, Amunts K, Zilles K. Cortical folding patterns and predicting cytoarchitecture. *Cerebral Cortex*. 2008; 18:1973–1980. [PubMed: 18079129]
- Fischl B, Sereno MI, Tootell RBH, Dale AM. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum Brain Mapp*. 1999; 8:272–284. [PubMed: 10619420]
- Fischl B, van der Kouwe A, Destrieux C, Halgren E, Ségonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy D, Caviness V, Makris N, Rosen B, Dale AM. Automatically parcellating the human cerebral cortex. *Cer Cortex*. 2004; 14:11–22.
- Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL, Xu J, Jbabdi S, Webster M, Polimeni JR, Van Essen DC, Jenkinsen M. The minimal processing pipelines for the Human Connectome Project. *Neuroimage*. 2013; 80:105–124. [PubMed: 23668970]
- Greve DN, Fischl B. Accurate and robust brain image alignment using boundary-based registration. *Neuroimage*. 2009; 48:63–72. [PubMed: 19573611]
- Holmes CJ, Hoge R, Collins L, Woods R, Toga AW, Evans AC. Enhancement of MR images using registration for signal averaging. *J Comput Assist Tomogr*. 1998; 22:324–333. [PubMed: 9530404]
- Hubbard NA, Hutchison JL, Motes MA, Shokri-Kojori E, Bennett IJ, Brigante RM, Haley RW, Rypma B. Central executive dysfunction and deferred prefrontal processing in veterans with Gulf War illness. *Clinical Psychological Science*. 2013 Advanced Access. 10.1177/2167702613506580
- Hutchison JL, Lu H, Rypma B. Neural mechanisms of age-related slowing: the CBF/ CMRO2 ratio mediates age-differences in BOLD signal and human performance. *Cer Cortex*. 2013a; 23:2337–2346. 10.1093/cercor/bhs233
- Hutchison JL, Shokri-Kojori E, Lu H, Rypma B. A BOLD perspective on age-related neurometabolic-flow coupling and neural efficiency changes in human visual cortex. *Front Psychol*. 2013b; 4:1–13. 10.3389/fpsyg.2013.00244 [PubMed: 23382719]
- Karni A, Meyer G, Jezard P, Adams MM, Turner R, Ungerleider LG. Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature*. 1995; 377:155–158. [PubMed: 7675082]
- Kennedy KM, Erickson KI, Rodrigue KM, Voss MW, Colcombe SJ, Kramer AF, Acker JD, Raz N. Age-related differences in regional brain volumes: A comparison of optimized voxel-based morphometry to manual volumetry. *Neurobiol Aging*. 2009; 30:1657–1676. [PubMed: 18276037]

- Kennedy DN, Lange N, Makris N, Bates J, Meyer J, Caviness JS. Gyri of the human neocortex: an MRI-based analysis of volume and variance. *Cer Cortex*. 1998; 8:372–384.
- Operto G, Bulot R, Anton J-L, Coulon O. Projection of fMRI data onto the cortical surface using anatomically-informed convolution kernels. *Neuroimage*. 2008; 39:127–135. [PubMed: 17931891]
- Rypma B. Factors controlling neural activity during delayed-response task performance: testing a memory organization hypothesis of prefrontal function. *Neuroscience*. 2006; 139:223–235. [PubMed: 16343777]
- Rypma B, Berger JS, Genova HM, Rebbeschi D, D’Esposito M. Dissociating age-related changes in cognitive strategy and neural efficiency using event-related fMRI. *Cortex*. 2005; 41:582–594. [PubMed: 16042034]
- Rypma B, Berger JS, Prabhakaran V, Bly BM, Kimberg DY, Biswal BB, D’Esposito M. Neural correlates of cognitive efficiency. *Neuroimage*. 2006; 33:969–979. [PubMed: 17010646]
- Saad ZS, Reynolds RC. SUMA. *Neuroimage*. 2012; 62:768–773. [PubMed: 21945692]
- Saad ZS, Ropella KM, DeYoe EA, Bandettini PA. The spatial extent of the BOLD response. *Neuroimage*. 2003; 19:132–144. [PubMed: 12781733]
- Saad, ZS.; Reynolds, RC.; Argall, B.; Japee, S.; Cox, RW. SUMA: an interface for surface-based intra- and inter-subject analysis with AFNI. 2004 IEEE International Symposium on Biomedical Imaging: From Nano to Macro; Arlington, VA. IEEE; 2004. p. 1510-1513.
- Siebert TM, Brewer JB. Default network correlations analyzed on native surfaces. *J Neurosci Meth*. 2011; 198:301–311.
- Siebert TM, Majid TSA, Aron AR, Corey-Bloom J, Brewer JB. Stability of resting fMRI interregional correlations analyzed in subject-native space: a one-year longitudinal study in healthy adults and premanifest Huntington’s disease. *Neuroimage*. 2012; 59:2452–2463. [PubMed: 21945695]
- Talairach, J.; Tournoux, P. *Coplanar Stereotaxic Atlas of the Human Brain*. Thieme Medical; New York: 1988.
- Tucholka A, Fritsch V, Poline J-B, Thirion B. An empirical comparison of surface-based and volume-based group studies in neuroimaging. *Neuroimage*. 2012; 63:1443–1453. [PubMed: 22732555]
- Van Essen DC. Windows on the brain. The emerging role of atlases and databases in neuroscience. *Curr Opin Neurobiol*. 2002; 12:574–579. [PubMed: 12367638]
- Van Essen DC. A Population-Average, Landmark- and Surface-based (PALS) atlas of human cerebral cortex. *Neuroimage*. 2005; 28:635–662. [PubMed: 16172003]
- Van Essen DC, Drury HA, Dickson J, Harwell J, Hanlon D, Anderson CH. An integrated software suite for surface-based analyses of cerebral cortex. *J Am Med Inform Assoc*. 2001; 8:443–459. [PubMed: 11522765]
- Wig GS, Laumann TO, Petersen SE. An approach for parcellating human cortical areas using resting-state correlations. *Neuroimage*. 2013 in press.

Highlights

- Transformation into standard space affected structural and functional analysis results and group difference determinations.
- Surface-based analyses using functional ROIs yielded the most group differences.
- Native-space volumetric analyses generated significant group differences and would be appropriate when time is limited.

Figure 1.

The eight landmarks drawn after flattening in Caret. Three of these landmarks were Caret-recommended (cf. Van Essen, 2005): Central Sulcus, Sylvian Fissure, anterior half of the Superior Temporal Gyrus (labeled aSTG). An additional five landmarks (3 sulci and 2 gyri) were added that we found to be consistent across participants: Inferior Rostral Sulcus, Pars Triangularis, Pars Orbitalis, Inferior Frontal Sulcus, and Superior Temporal Sulcus.

Figure 2.

An overview of data processing decisions commonly made by researchers. Depending on the processing pathway chosen, statistical significance (e.g., $p < .05$) might or might not be attained. The present paper seeks to elucidate the most appropriate processing pipelines for fMRI data.

Figure 3.

Effects of data processing on percent signal change. (a) Data Type (surface, volume) $\times$ ROI (functional, anatomical) within PFC. (b) Data Type (surface, volume) $\times$ ROI (functional, anatomical) within MC. (c) Data Type (surface, volume) $\times$ Brain Space (native, standard). (d) Brain Region (PFC, MC) $\times$ Brain Space (native, standard). (e) Brain Region (PFC, MC) $\times$ ROI (functional, anatomical). For simple contrasts: *** p .001, ** p .01, * p .05. For interactions: ††† p .001, †† p .01.

Figure 4.

Effects of data processing on group differences in percent signal change. (a) Brain Region (PFC, MC) $\times$ Group (collegiate, military-veterans). (b) ROI Type (functional, anatomical) $\times$ Group within PFC. (c) ROI (functional, anatomical) $\times$ Group (collegiate, military-veterans) within MC. (d) Data Type (surface, volume) $\times$ Group (collegiate, military-veterans) within MC. (e) Data Type (surface, volume) $\times$ Health (healthy, Gulf War Illness) within MC. For simple contrasts: *** p .001, ** p .01, * p .05. For interactions: ††† p .001, †† p .01.

Figure 5.

Effects of data processing on extent of activation. (a) Brain Region (PFC, MC) $\times$ ROI (functional, anatomical) for volume analyses (voxels). (b) Brain Region (PFC, MC) $\times$ ROI (functional, anatomical) for surface analyses (nodes). (c) Brain Region (PFC, MC) $\times$ ROI (functional, anatomical) for volume analyses (voxels) in native space. (d) Brain Region (PFC, MC) $\times$ ROI (functional, anatomical) for volume analyses (voxels) in standard space. (e) Brain Region (PFC, MC) $\times$ ROI (functional, anatomical) for surface analyses (nodes) in native space. (f) Brain Region (PFC, MC) $\times$ ROI (functional, anatomical) for surface analyses (nodes) in standard space. For simple contrasts: *** p .001, ** p .01, * p .05. For interactions: ††† p .001, †† p .01.

Figure 6.

Effects of data processing on group differences in extent of activation. (a) Metric (voxels, nodes) × Group (collegiate, military-veterans) within PFC. (b) Brain Space (native, standard) × Group (collegiate, military-veteran) within PFC. (c) ROI (functional, anatomical) × Group (collegiate, military-veteran) within PFC. (d) Brain Space (native, standard) × Group (collegiate, military-veteran) within MC. (e) ROI (functional, anatomical) × Group (collegiate, military-veteran) within MC. (f) Brain Space (native, standard) × Group (collegiate, military-veteran) within MC. (g) ROI (functional, anatomical) × Group (collegiate, military-veteran) within MC. For simple contrasts: *** p .001, ** p .01, * p .05. For interactions: ††† p .001, †† p .01.