Quantifying differences in fMRI preprocessing pipelines via OGRE (One-step General Registration and Extraction)

Running title: OGRE fMRI preprocessing pipeline

Authors:

Mark P. McAvoy, PhD. mcavoy@wustl.edu. Affilitation: Program in Occupational Therapy, Washington University School of Medicine, St. Louis, MO, United States.

Lei Liu, PhD. lei.liu@wustl.edu. Affilitation: Department of Biostatistics, Washington University School of Medicine, St. Louis, MO, United States.

Ruiwen Zhou, PhD. zhour@wustl.edu. Affilitation: Department of Biostatistics, Washington University School of Medicine, St. Louis, MO, United States

Benjamin A. Philip, PhD.* bphilip@wustl.edu. Affilitation: Program in Occupational Therapy, Washington University School of Medicine, St. Louis, MO, United States.

*: Corresponding author.

Data availability statement: All software is openly available at https://github.com/PhilipLab/OGRE-pipeline. The data that support the findings of this study are available on request from the corresponding author, BAP, upon reasonable request.

Funding statement: This work was funded by NIH/NINDS R01 NS114046 to BAP.

Conflict of interest disclosure: Author BAP and Washington University in St. Louis have a licensing agreement with PlatformSTL to commercialize the drawing program used in this study.

Patient consent & ethics approval statement: All participants gave informed consent, and all procedures were approved by the Institutional Review Board at Washington University in St. Louis School of Medicine.

Clinical trial registration: The data used in this study were collected as part of clinical trial NCT05207878.

ABSTRACT

Volumetric preprocessing methods continue to enjoy great popularity in the analysis of functional MRI (fMRI) data. Among these methods, the software packages FSL (FMRIB, Oxford, UK) and FreeSurfer (LCN, Charlestown, MA) are omnipresent throughout the field. However, it remains unknown what advantages an integrated FSL+FreeSurfer preprocessing approach might provide over FSL alone. Here we developed the One-step General Registration and Extraction (OGRE) pipeline to combine FreeSurfer and FSL tools for brain extraction and registration, for FSL volumetric analysis of fMRI data. We compared preprocessing approaches in a dataset wherein adult human volunteers (N=26) performed a precision drawing task during fMRI scanning. OGRE's preprocessing, compared to traditional FSL preprocessing, led to lower inter-individual variability across the brain, more precise brain extraction, and greater detected activation in sensorimotor areas contralateral to movement. This demonstrates that the introduction of FreeSurfer tools via OGRE preprocessing can improve fMRI data analysis, in the context of FSL's volumetric analysis approach. The OGRE pipeline provides a turnkey method to integrate FreeSurfer-based brain extraction and registration with FSL analysis of task fMRI data.

KEYWORDS: image analysis pipeline; fMRI; brain; software; magnetic resonance imaging; humans; brain mapping

INTRODUCTION

Numerous methods and software packages exist for the analysis of fMRI data, and many of these approaches offer different advantages and tradeoffs. One commonly used analysis package is the fMRIB Software Library (FSL; Jenkinson et al., 2012), particularly its FEAT software for general linear model (GLM) volumetric model-based analysis of fMRI data (Woolrich et al., 2004; Woolrich et al., 2001). FSL is highly popular: as of September 2023, Jenkinson et al. (2012) has over 9300 citations, including over 2300 in the preceding eighteen months (Google Scholar). However, despite FSL's popularity, FSL preprocessing has some shortcomings, one of which is its reliance on the FSL Brain Extraction Tool (BET; Smith, 2002) for identifying brain voxels in the scan volume. BET is prone to under- and over-extraction of brain voxels (Boesen et al., 2004; Klein et al., 2010; Zhuang, Valentino, & Toga, 2006), and has limited options for adjusting its algorithmic criteria (Jenkinson, 2018). Together these shortcomings mean that BET either produces inaccurate brain extractions or requires time-consuming and error-prone manual adjustment (Mohapatra, Gosai, & Schlaug, 2023; Quilis-Sancho, Fernandez-Blazquez, & Gomez-Ramirez, 2020). Therefore, alternative FSL-compatible preprocessing approaches (including brain extraction) could improve the reliability and accuracy of fMRI data analysis.

One alternative preprocessing approach involves using FreeSurfer software for brain extraction and parcellation (e.g. Fischl & Dale, 2000; Fischl et al., 2002; Fischl et al., 2004). This has been further refined by the Human Connectome Project (HCP), which uses FSL tools for coarse alignment and registration, followed by FreeSurfer tools for fine alignment, brain extraction, and parcellation; critically, these two steps are applied simultaneously to functional data in a "one-step resampling" procedure (Glasser et al., 2013). This process removes spatial artifacts and distortions, creates cortical surface and myelin maps, performs precise within-individual cross-modal registrations, and performs volume-based registrations to standard spaces. The one-step resampling procedure is a small portion of the full HCP processing pipeline, which subsequently uses a surface-based approach with numerous further advantages (Coalson, Van Essen, & Glasser, 2018; Dickie et al., 2019; Glasser et al., 2018; Glasser et al., 2016). However, volumetric approaches continue to enjoy widespread popularity as lower resolution fMRI data remains common in the field (e.g. Holmes et al., 2015). Few studies have assessed the potential advantages gained by incorporating the one-step resampling procedure into FSL volumetric analysis, because direct, controlled tests (i.e. controlling for later processing steps) have not been possible because no turnkey solution previously existed to connect FreeSurfer with FSL FEAT.

Here, we created the OGRE (One-step General Registration and Extraction) pipeline to integrate the "one-step resampling" procedure (coarse registration with FSL followed by fine alignment and brain extraction with FreeSurfer) as a general-purpose fMRI preprocessing tool for FSL volumetric analysis. We performed a fully-controlled comparison of OGRE vs. FSL preprocessing (i.e. FSL+FreeSurfer vs. FSL-only) by evaluating the effects of the two preprocessing pipelines

on a fMRI motor task with the same subsequent volumetric FSL FEAT GLM analysis. The two preprocessing methods were compared for differences in task-based activation, brain extraction, and inter-individual variability.

MATERIALS & METHODS

Participants

Twenty-six right-handed adults (18 female; ages 52 ± 16, range 24-76) performed a precision drawing task (see *fMRI Task* below) in the fMRI scanner. 11/26 participants had peripheral nerve injuries to their right arm, but the differences between groups (categorical or performance-related) were not a focus of the current study. Two participants were unable to complete the drawing task with their right hand, so right hand (RH) drawing analyses were limited to 24 participants. All participants gave informed consent, and all procedures were approved by the Institutional Review Board at Washington University in St. Louis School of Medicine.

fMRI Acquisitions

Scans were performed on a Siemens PRISMA 3T MRI scanner. BOLD EPIs for fMRI were collected using a T2*-weighted gradient echo sequence, a standard 64-channel birdcage radio-frequency coil, and the following parameters: TR = 662 ms, TE = 30 ms, flip angle = 52°, 72 x 72 voxel matrix, FOV = 216 mm, 60 contiguous axial slices acquired in interleaved order, resolution: 3.0 x 3.0 x 3.0 mm, bandwidth = 2670 Hz/pixel, multi-band acceleration = 6x. Siemens auto-align was run at the start of each session.

High-resolution T1-weighted structural images were also acquired, using the 3D MP-RAGE pulse sequence: TR = 4500 ms, TE = 3.16 ms, TI = 1000 ms, flip angle = 8.0°, 256 x 256 voxel matrix, FOV = 256 mm, 176 contiguous axial slices, resolution: 1.0 x 1.0 x 1.0 mm. A T2-weighted image was also acquired at: TR = 3000 ms, TE = 409 ms, 256 x 256 voxel matrix, FOV = 256 mm, 176 contiguous axial slices, resolution: 1.0 x 1.0 x 1.0 mm. Spin echo field maps were also collected before the functional runs.

fMRI Task

Each participant completed 6 BOLD functional scans. During each scan, the participant used one hand (3 scans with right hand RH, 3 with left hand LH, interleaved order) to perform a precision drawing task based on the STEGA app (Philip & Frey, 2014; Philip et al., 2023). Participants saw hollow shapes, and were instructed to draw a line within the bounds of the shape, as fast as possible while prioritizing staying in-bounds over speed. Participants saw a video overlay of their (transparent) hand/pen over the drawn shape, using an MRI-compatible drawing tablet with bluescreen technology (Karimpoor et al., 2015), as shown in **Figure 1**. The task was presented in a block design with 15.2 seconds (23 images) of drawing, followed by 15.2 seconds (23 images) of rest (fixation cross). A scan comprised ten cycles of draw/rest, with an additional rest block at the start, and 3.3 seconds (5 images) of additional rest after the final rest block, leading to a total duration of 5:23 (230 images of drawing, 258 rest) per scan.

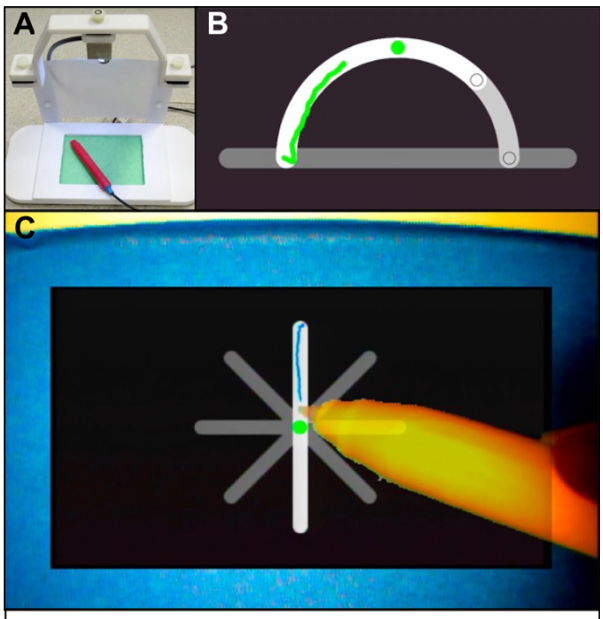

Figure 1. Motor task used for pipeline testing. A: MRI-compatible tablet. B: STEGA precision drawing task. C: Participant view of bluescreen integration.

Data Preprocessing: One-step General Registration and Extraction (OGRE) Pipeline

To integrate HCP preprocessing with FSL software via a Macintosh comparable turnkey analysis pipeline, OGRE (One-step General Registration and Extraction) was developed based on version 3.27 of the HCP pipeline (https://github.com/Washington-University/HCPpipelines/tree/v3.27.0).

OGRE uses the following preprocessing steps of the HCP pipeline: FreeSurfer brain extraction and parcellation, motion correction, field map distortion correction, and warping to the 2mm MNI atlas via the "one-step resampling" procedure which combines a coarse registration via FSL FLIRT and FNIRT (J. L. Andersson, M. Jenkinson, & S. Smith, 2007) with a fine registration via the FreeSurfer toolbox, as detailed in Glasser et al. (2013). New features added for OGRE included the option to specify native resolution as 0.7, 0.8, or 1.0 mm; the option to leave fMRI time series in native HCP space instead of registering to MNI atlas; extended compatibility to FreeSurfer versions 7.2.0, 7.3.2 and 7.4.0 (https://surfer.nmr.mgh.harvard.edu/); and the addition of three new steps to match FEAT: nonlinear spatial filtering (Smith & Brady, 1997), nonlinear temporal high pass filtering (Woolrich et al., 2001), and a switch from mean-based to median-based intensity normalization. OGRE also produces FEAT-formatted registration images and matrices for higher-level FEAT analysis. OGRE preprocessing outputs were entered as data for analysis in FEAT using its “Statistics” option (rather than “Full Analysis,” since OGRE has replaced FEAT preprocessing).

OGRE software and documentation are available online at: https://github.com/PhilipLab/OGRE-pipeline

Parallel analyses: OGRE vs. FSL-only. In the "OGRE analysis," the OGRE pipeline described above was used to preprocess data with the following options: FreeSurfer 7.4.0, native resolution of 1 mm, distortion correction using spin echo field maps, spatial smoothing kernel of 6 mm FWHM, intensity normalization via “grand median scaling” to a value of 10000, high pass temporal filtering with 60 sec cutoff, and output registration to 2 mm MNI atlas. Subsequently, GLM statistical analysis was performed with FEAT version 6.0.7 (Smith et al., 2004). Explanatory variables (EVs) were modeled, along with their temporal derivatives, according to the block design described in the previous section (*fMRI Task*). Volumes with excess head motion were addressed (Siegel et al., 2014) via additional confound EVs calculated within each time series as framewise displacement of 75th percentile + 1.5* interquartile range, using the FSL script FSLMotionOutliers. The hemodynamic response was accounted for by convolving the model with a double-gamma function. First-level contrasts of parameter estimates (COPEs) were calculated for Task vs. Rest. The first-level COPE (Task > Rest) served as input to higher-level analyses performed using a fixed-effects model. Z-statistic (Gaussianized T) images were thresholded using clusters determined by $Z \geq 3.1$ and a corrected cluster significance threshold of $p < 0.05$. The first-level COPEs were averaged across runs for each participant (second level). Region of interest (ROI) analyses were performed on second-level (i.e. participant-level) data, as detailed in *ROI Analyses* below.

In the "FSL-only analysis," preprocessing included the following steps: Non-brain structures were removed using BET. Head movement was reduced using MCFLIRT motion correction. Distortion correction was applied using spin echo field maps using FSL PRELUDE and FUGUE. Paralleling the OGRE implementation, spatial smoothing, intensity normalization, and the high-pass temporal filter were applied with the same settings as above. Functional data were registered with the high-resolution structural image using boundary-based registration (Greve & Fischl, 2009), and resampled to 2x2x2 mm resolution using FLIRT; the participant images were then registered to standard images (Montreal Neurological Institute MNI-152) using FNIRT nonlinear registration (J. Andersson, M. Jenkinson, & S. Smith, 2007). Subsequently, GLM statistical analysis was performed with FEAT as described for the OGRE analysis above.

Individual example participants were selected for illustration on an *a priori* basis via the following criteria: participants were ranked by the metric "magnitude difference between analysis methods, in the whole-cortex ROI" (averaged between LH condition and RH condition), from which the 25th percentile participant, median participant, and 75th percentile participant were selected. Single-scan example images were taken from the second scan of the Draw-RH condition.

ROI Analyses. To quantitatively compare the two analyses, a region-of-interest (ROI) approach was used, involving 42 ROI variables (21 ROIs, calculated separately for LH-draw and RH-draw conditions). These 21 ROIs comprised: 6 motor sub-regions (3 per hemisphere), 14 networks (7 per hemisphere), and 1 whole-cortex ROI. Specifically, the six primary motor cortex (M1) ROIs were chosen as 5 mm radius spheres based on normative activation maps (Smith & Frey, 2011): Hand-M1 ("Hand;" left hemisphere x = -38, y = -24, z = 54; right hemisphere x = 38, y = -24, z = 54), Lip-M1 ("Lip;" left hemisphere x = -54, y = -8, z = 42; right hemisphere x = 44, y = -12, z = 42) and Foot-M1 ("Foot;" left hemisphere x = -4, y = -28, z = 72; right hemisphere x = 8, y = -24, z = 74). The 14 network ROIs were the full networks from the NITRC (2022)

atlas of cortical networks in the Yeo et al. (2011) network parcellation: medial visual, sensorimotor, dorsal attention, ventral attention, frontoparietal, default mode (DMN), and subcortical; each in left and right hemispheres. Finally, the "whole-cortex" ROI was created as the union of all twelve cortical network-based ROIs. For each participant, the signal magnitude was calculated as the mean % BOLD signal change for the Task > Rest contrast in second-level (within participant) analyses.

OGRE analysis was compared with FSL-only analysis via two approaches. The first approach focused on inter-individual variability. The primary variable was the standard deviation (across the signal magnitudes of the 26 participants, within each ROI). Direct statistical comparisons within each ROI used a pairwise Pitman-Morgan test (Morgan, 1939) to compare the inter-individual variability of OGRE signal magnitudes vs. the inter-individual variability of FSL-only signal magnitudes. A big-picture overview was developed by comparing the 14 network ROI standard deviations for OGRE analysis, vs. the 14 standard deviations for FSL-only analysis, via a paired-sample t-test. Finally, to explore the spatial patterns of where the two analyses differed, a 2 (Method: OGRE vs. FSL-only) * 9 (Area) * 2 (Hemisphere: left, right) * 2 (Condition: LH-draw, RH-draw) ANOVA was performed on the standard deviations. The "area" factor had 9 levels instead of 10 because the Foot region was omitted as an outlier (high difference between OGRE and FSL-only in Foot, Right Hemisphere).

The second approach for comparing OGRE vs FSL-only analysis focused on mean signal magnitude. The same process was followed as above, except: the primary variable was the mean signal magnitude (across the 26 participants). Direct statistical comparisons used paired-sample t-tests. In the signal magnitude analysis, no outliers were omitted from the ANOVA, making it 2 * 10 * 2 * 2 instead of 2 * 9 * 2 *2.

Individual example ROIs were selected for illustration on an *a priori* basis via a similar method to the selection of example participants: ROIs were ranked by the metric "magnitude difference between analysis methods", from which the 25th percentile ROI (ventral attention, right hemisphere, RH-draw condition), median ROI (ventral attention, right hemisphere, LH-Draw condition), and 75th percentile ROI (default mode network, right hemisphere, left hand) were selected.

Brain extraction comparisons. To quantify the difference in brain extraction and registration results between analysis methods, we compared the size of post-registration (i.e. standard space) brain masks from each task fMRI scan, following Zhang et al. (2019).

RESULTS

Lower inter-individual variability with OGRE compared to FSL-only analysis.

In all forty-two ROIs, OGRE analysis led to numerically lower inter-individual variability (standard deviation) than FSL-only analysis, as shown in **Figure 2** and detailed in **Table 1**. This difference between analysis methods was statistically significant in 41/42 ROIs (98%).

To provide a big-picture overview of how our analysis methods affected data variability across the brain, we compared standard deviations (across participants) across our network-based ROIs. We found significantly lower standard deviations for OGRE than FSL (paired-samples t-test: T = -16.7, p = 9.94×10^{-16}); in these network-based ROIs, OGRE standard deviations were 0.283 ± 0.0548, while FSL standard deviations were 0.504 ± 0.101, an effect size of 1.60. Moreover, OGRE also had lower variability in the whole-cortex ROI under either task condition (p ≤ 0.005; **Table 1** bottom two rows), demonstrating broadly lower variability for OGRE analysis than FSL-only analysis.

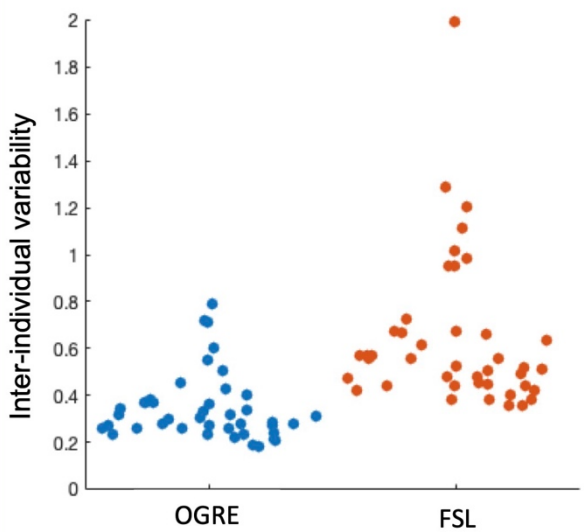

Figure 2. Visualization of lower inter-individual variability with OGRE preprocessing than FSL-only preprocessing. Each point represents the SD across participants for one ROI.

To explore the spatial pattern of where the two analyses differed, we performed a 4-way ANOVA: 2 (Method: OGRE vs. FSL-only) * 9 (Area: see **Table 1** Column 1) * 2 (Hemisphere: left, right) * 2 (Condition: LH-draw, RH-draw) on standard deviations. We found significant main effects of Method ($F(1,33) = 106.81$, $p < 0.001$), Area ($F(8,33) = 15.45$, $p < 0.001$), and Condition ($F(1,33) = 16.22$, $p = 0.0003$), but not Hemisphere ($p = 0.288$). We found a significant interaction for Area * Hemisphere ($F(8,33) = 3.23$, $p = 0.008$), and a trend toward a Hemisphere * Condition interaction ($F(1,33) = 3.69$, $p = 0.0633$), but no other interactions ($p > 0.6$). Post-hoc tests revealed that the main effect of Method was driven by lower variability for OGRE (0.338 ± 0.137) than FSL-only (0.569 ± 0.188), the main effect of Area was driven by significantly higher variability for motor ROIs than for network ROIs, likely due to the motor ROIs' smaller size; and the main effect of Condition was driven by higher variability for LH-draw (0.498 ± 0.206) than RH-draw (0.408 ± 0.187). The Area * Hemisphere interaction was driven by high variability for bilateral Hand and left hemisphere Lip; the Hemisphere * Condition trend was driven by higher variability for LH-draw (either hemisphere) than "right hemisphere, RH-draw." In summary, OGRE showed improved reliability (lower inter-individual variability) throughout the brain; this effect was more pronounced in smaller ROIs, but otherwise not specific to task-related areas or networks (e.g. sensorimotor network).

ROI			Standard deviation (across participants)			
Area	**Hemisphere**	**Condition**	**OGRE**	**FSL-only**	**Improvement (FSL-OGRE)**	***p***
Hand	Left	LH	0.456	0.671	0.216	7.43×10^{-6}
Hand	Right	LH	0.789	1.113	0.324	1.97×10^{-5}
Hand	Left	RH	0.600	0.985	0.386	1.29×10^{-4}
Hand	Right	RH	0.402	0.519	0.118	0.011
Lip	Left	LH	0.712	1.013	0.301	3.96×10^{-4}
Lip	Right	LH	0.427	0.659	0.232	8.24×10^{-5}
Lip	Left	RH	0.549	0.950	0.401	5.36×10^{-7}
Lip	Right	RH	0.320	0.452	0.132	0.009
Foot	Left	LH	0.506	1.204	0.698	1.14×10^{-7}
Foot	Right	LH	0.721	1.996	1.275	6.97×10^{-13}
Foot	Left	RH	0.256	0.951	0.695	6.24×10^{-10}
Foot	Right	RH	0.344	1.289	0.946	1.13×10^{-9}
Med Visual	Left	LH	0.385	0.557	0.172	5.92×10^{-5}
Med Visual	Right	LH	0.370	0.571	0.201	2.69×10^{-6}
Med Visual	Left	RH	0.277	0.509	0.232	1.37×10^{-9}
Med Visual	Right	RH	0.275	0.526	0.251	5.49×10^{-7}
Sensorimotor	Left	LH	0.334	0.554	0.220	4.42×10^{-6}
Sensorimotor	Right	LH	0.364	0.673	0.309	1.96×10^{-7}
Sensorimotor	Left	RH	0.258	0.441	0.183	6.26×10^{-6}
Sensorimotor	Right	RH	0.260	0.422	0.162	1.02×10^{-6}
Drsl Attention	Left	LH	0.369	0.668	0.299	4.70×10^{-6}
Drsl Attention	Right	LH	0.372	0.727	0.355	3.57×10^{-7}
Drsl Attention	Left	RH	0.311	0.634	0.323	2.19×10^{-8}
Drsl Attention	Right	RH	0.298	0.612	0.314	3.29×10^{-10}
Vtrl Attention	Left	LH	0.306	0.482	0.176	1.30×10^{-4}
Vtrl Attention	Right	LH	0.315	0.569	0.255	4.52×10^{-8}
Vtrl Attention	Left	RH	0.230	0.379	0.150	2.08×10^{-6}
Vtrl Attention	Right	RH	0.285	0.402	0.117	0.002
Frontoparietal	Left	LH	0.269	0.568	0.299	5.26×10^{-9}

Frontoparietal	Right	LH	0.276	0.557	0.282	6.19×10^{-11}
Frontoparietal	Left	RH	0.218	0.448	0.229	9.49×10^{-10}
Frontoparietal	Right	RH	0.234	0.474	0.240	2.95×10^{-12}
DMN	Left	LH	0.261	0.479	0.218	1.20×10^{-7}
DMN	Right	LH	0.269	0.492	0.223	1.32×10^{-9}
DMN	Left	RH	0.187	0.442	0.256	7.45×10^{-12}
DMN	Right	RH	0.231	0.442	0.211	7.36×10^{-8}
Subcortical	Left	LH	0.332	0.385	0.052	0.350
Subcortical	Right	LH	0.240	0.356	0.116	0.005
Subcortical	Left	RH	0.181	0.355	0.174	6.67×10^{-5}
Subcortical	Right	RH	0.206	0.385	0.179	6.93×10^{-5}
Whole cortex	n/a	LH	0.276	0.508	0.233	5.51×10^{-8}
Whole cortex	n/a	RH	0.211	0.419	0.208	2.57×10^{-10}

Table 1. Inter-individual variability (standard deviation of BOLD % signal change during Task > Rest) was consistenyly lower for OGRE analysis than FSL-only analysis. Condition = data collected during drawing with left hand (LH) or right hand (RH). P-values calculated via Pitman-Morgan test. Med = medial, Drsl = dorsal, Vtrl = ventral.

Inconsistent increases in BOLD signal magnitude for OGRE compared to FSL-only analysis

In terms of signal magnitude (participant mean % BOLD signal change for Task > Rest), OGRE analysis led to numerically greater activity in 28/42 ROIs (67%), as shown in **Table 2**. Of the 28 ROIs with numerically higher activity for OGRE, 15 (54%) were statistically significant. Of the 14/42 ROIs with numerically lower activity for OGRE, 4 (29%) were statistically significant.

To provide a big-picture overview of how our analysis methods affected data variability across the brain, we compared mean signal magnitudes (across participants) across our network-based ROIs. We found significantly higher magnitudes for OGRE than FSL ($T = 2.25$, $p = 0.033$); in these network-based ROIs, OGRE means were 0.059 ± 0.195 (mean ± SD across participants), while FSL means were -0.003 ± 0.336, an effect size of 0.23. However, we found no differences between analysis methods in the whole-cortex ROI for either condition ($p > 0.25$; **Table 2**, bottom two rows), demonstrating that OGRE analysis increased BOLD signal magnitudes in some networks but not overall.

To explore the spatial pattern of where the two analyses differed, we performed a 4-way ANOVA: 2 (Method: OGRE vs. FSL-only) & 10 (Area: see **Table 2** Column 1) * 2 (Hemisphere: left, right) * 2 (Condition: LH-draw, RH-draw) on mean signal magnitudes. We found a significant main effect of Area ($F(9,37)=11.28$, $p < 0.001$), and a significant Hemisphere * Condition interaction ($F(1,37) = 6.36$, $p = 0.016$), but no other main or interaction effects ($p > 0.3$ except the following example), including no significant main effect of Method ($F(1,37) = 2.74$, $p = 0.106$). Post-hoc analysis indicated that the main effect arose from lower means for Foot, Medial Visual, and Default Mode than other ROIs; and the Hemisphere * Condition interaction arose from higher means for "left hemisphere, RH-draw" than "right hemisphere, RH-draw." Therefore, OGRE did not lead to systematic changes in signal magnitude. However, the meaning of "changes in signal magnitude" is not necessarily clear since this is real data with no known ground truth.

ROI			**Mean (across participants)**			***p***
Area	**Hemisphere**	**Condition**	**OGRE**	**FSL-only**	**Increase (OGRE-FSL)**	
Hand	Left	LH	0.067	0.012	0.055	0.356
Hand	Right	LH	0.905	1.129	-0.225	0.022
Hand	Left	RH	1.138	1.455	-0.317	0.012

Hand	Right	RH	-0.121	-0.303	0.181	0.001
Lip	Left	LH	0.472	0.524	-0.052	0.602
Lip	Right	LH	0.142	0.084	0.058	0.406
Lip	Left	RH	0.420	0.557	-0.137	0.185
Lip	Right	RH	-0.028	-0.183	0.155	0.009
Foot	Left	LH	-0.115	-0.622	0.507	0.011
Foot	Right	LH	-0.121	-0.946	0.826	0.008
Foot	Left	RH	-0.068	-0.678	0.609	0.001
Foot	Right	RH	-0.165	-0.872	0.706	0.005
Med Visual	Left	LH	-0.265	-0.554	0.289	0.000
Med Visual	Right	LH	-0.237	-0.499	0.262	0.000
Med Visual	Left	RH	-0.202	-0.395	0.192	0.001
Med Visual	Right	RH	-0.255	-0.529	0.274	0.000
Sensorimotor	Left	LH	0.197	0.162	0.034	0.572
Sensorimotor	Right	LH	0.262	0.274	-0.012	0.874
Sensorimotor	Left	RH	0.319	0.397	-0.079	0.123
Sensorimotor	Right	RH	0.100	0.036	0.063	0.135
Drsl Attention	Left	LH	0.308	0.393	-0.084	0.320
Drsl Attention	Right	LH	0.348	0.466	-0.118	0.193
Drsl Attention	Left	RH	0.349	0.552	-0.203	0.015
Drsl Attention	Right	RH	0.392	0.636	-0.244	0.002
Vtrl Attention	Left	LH	0.130	0.123	0.007	0.898
Vtrl Attention	Right	LH	0.174	0.124	0.051	0.426
Vtrl Attention	Left	RH	0.141	0.143	-0.002	0.961
Vtrl Attention	Right	RH	0.146	0.171	-0.025	0.544
Frontoparietal	Left	LH	0.001	-0.128	0.128	0.086
Frontoparietal	Right	LH	0.051	-0.027	0.079	0.226
Frontoparietal	Left	RH	0.008	-0.110	0.118	0.035
Frontoparietal	Right	RH	0.085	0.052	0.033	0.533
DMN	Left	LH	-0.162	-0.425	0.263	0.000
DMN	Right	LH	-0.104	-0.316	0.212	0.000
DMN	Left	RH	-0.162	-0.447	0.286	0.000
DMN	Right	RH	-0.113	-0.328	0.215	0.000
Subcortical	Left	LH	-0.025	-0.001	-0.024	0.726
Subcortical	Right	LH	0.068	0.032	0.036	0.456
Subcortical	Left	RH	0.024	0.009	0.015	0.781
Subcortical	Right	RH	0.076	0.102	-0.026	0.636
Whole cortex	n/a	LH	0.075	0.010	0.064	0.273
Whole cortex	n/a	RH	0.086	0.056	0.030	0.523

Table 2. Task-related signal magnitude (BOLD % signal change during Task > Rest) was sometimes, but not always, higher for OGRE analysis than FSL-only analysis. For details see Table 1.

Increased signal in whole-brain analyses, including in task-specific areas, when using OGRE

To assess the effects of OGRE analysis on the whole brain, we conducted a whole-brain repeated measures analysis, treating OGRE analysis and FSL analysis as the two measures for each participant. For both LH-draw and RH-draw

conditions, the OGRE > FSL-only contrast showed significant effects in a small but distributed set of cortical networks, as shown in **Figure 3**. No areas were significant for FSL-only > OGRE.

The cluster list in **Table 3** illustrates how the OGRE > FSL-only activity includes deep structures that may represent noise (e.g. ventricles and white matter), but also includes areas with expected task-specific actiity (sensorimotor, parietal, and premotor areas contralateral to movement). Therefore, while OGRE may increase spurious signals in areas with high noise susceptibility, it also enhances signal in task-specific areas that are likely of primary interest.

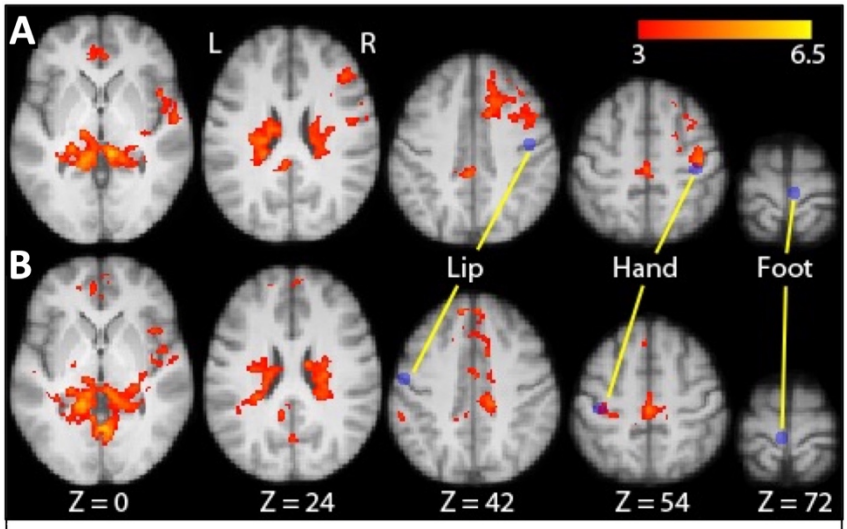

Figure 3. Significant differences in BOLD signal for OGRE > FSL-only contrast. (No significant differences for FSL-only > OGRE.) Purple circles: M1 regions of interest. **A:** LH drawing condition. **B:** RH drawing condition.

Condition	Cluster area	Peak coordinates			Peak Z-score	# voxels
		x	y	z		
LH-draw	Cingulate gyrus, hippocampal subiculum, thalamus, lateral ventricles, basal ganglia, internal capsule, midbrain	12	-38	-8	6.42	7872
	Left cingulate gyrus	-4	-34	38	5.72	616
	Left subcallosal cortex (inferior prefrontal)	6	30	-12	4.97	174
	Right primary sensorimotor	42	-18	56	4.79	148
	Right superior frontal gyrus (premotor cortex)	24	4	60	3.9	77
RH-draw	Lingual gyrus, cingulate gyrus, thalamus, basal ganglia, internal capsule, lateral ventricles, midbrain, pons	8	-68	-4	5.73	8495
	Midline paracingulate gyrus	-2	56	8	4.61	1132
	Right frontal operculum	38	14	10	4.43	232
	Left precuneus & cuneal cortex	-12	-70	20	4.61	125
	Left inferior parietal lobule	-42	-68	18	4.34	117
	Left primary sensorimotor	-36	-24	52	4.57	96
	Left supramarginal gyrus (inferior parietal lobule)	-56	-38	40	3.77	83

Table 3. Significant clusters in OGRE > FSL-only contrast. No significant clusters for FSL-only > OGRE.

OGRE identified smaller brain volumes, primarily through removal of non-brain voxels

To quantify the difference in brain extraction results between analysis methods, we compared the size of post-registration (i.e. standard space) brain masks of task fMRI runs. OGRE masks were 235,222 ± 3,868 voxels (range 226,807-242,917), whereas FSL-only masks were 311,474 ± 8,072 voxels (range 297,666-329,132); OGRE masks were smaller for all individuals, and this difference was statistically significant at the group level (paired-samples t-test: $p = 1.65 \times 10^{-35}$, $t = -114$). These smaller OGRE masks largely reflected the successful removal of non-brain volume: for each participant, we evaluated the OGRE-specific change (i.e., the area outside the OGRE mask but within the FSL mask) by using the MNI-152 standard brain mask to identify brain/non-brain voxels, and found that OGRE removed over 15 times

more non-brain voxels (72,249 ± 7,206, range 61,236-90,945) than brain voxels (4,746 ± 1,878, range 2,840-10,063). This ratio of non-brain/brain removal, while already high, reflects an over-estimation of removed brain voxels because the MNI brain mask includes non-brain areas such as the medial longitudinal fissure and inferior areas around the hypothalamus, as illustrated by our three example participant scans in **Figure 4**. Overall, these numbers show that the one-step FSL+FreeSurfer process (via its implementation in OGRE) did not produce perfect automated brain extraction and registration, but it generally provided superior performance compared to FSL-only brain extraction and registration.

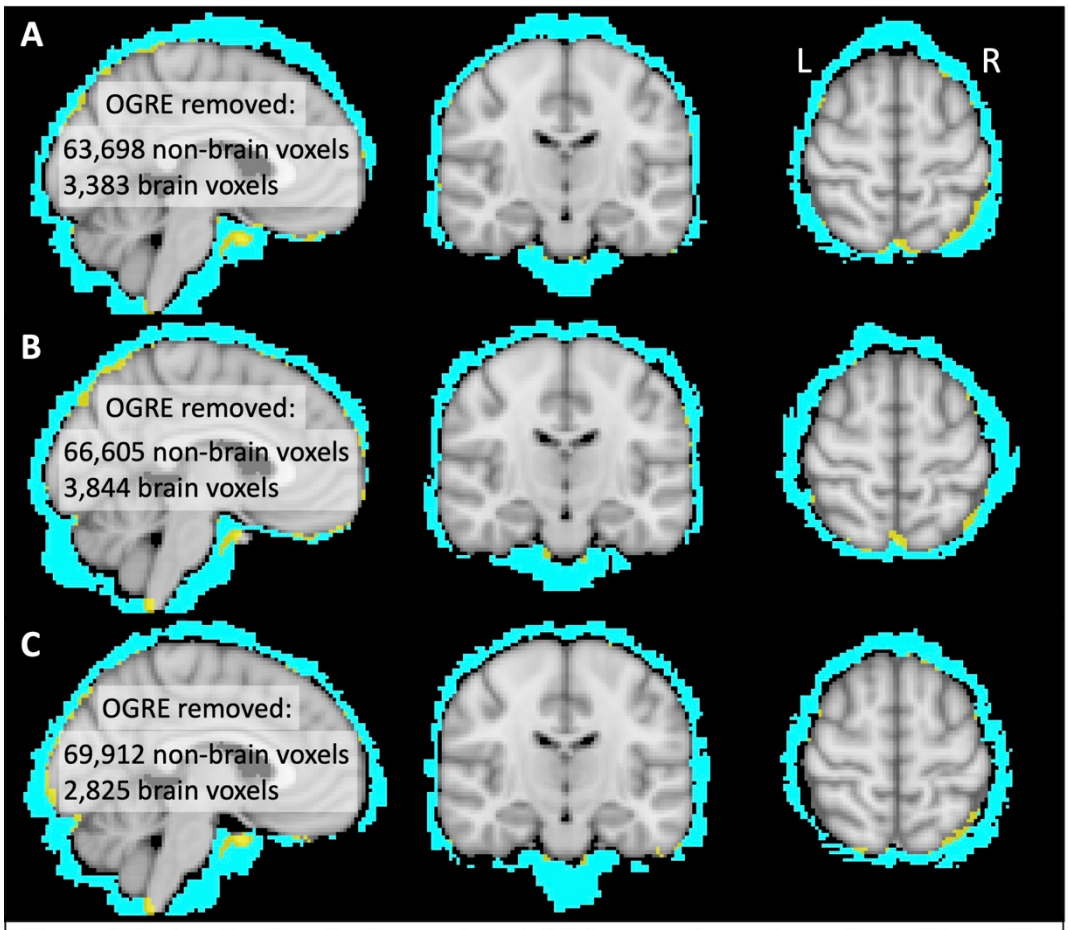

Figure 4. Brain extractions for 3 example task fMRI scans, shown at x = -4, y = -15, z = 60. Text and colored area: voxels removed by OGRE but not FSL. Brain voxels (yellow) and non-brain (blue) defined by MNI-152 standard brain mask. Participants chosen by change in signal magnitude across cortex, OGRE vs. FSL-only. **A:** Participant with modest (25th percentile) change in signal magnitude. **B:** Participant with median change in signal magnitude. **C:** Participant with large (75th percentile) change in signal magnitude.

Differences between OGRE and FSL-only analyses are broadly consistent across participants

Our participant group included at least one major unmodeled source of variance: the distinction between typical adults and patients with peripheral nerve injury. This provides an advantage for practical testing, because it highlights OGRE's robustness in the face of complex real-world data. However, we must rule out the possibility that this sample choice distorted our results. To address this, we first assessed whether participant group (patient vs. control) affected our measure of task-related BOLD signal (% BOLD signal change during Task > Rest). We performed a between-groups t-test to assess group differences in BOLD signal, and repeated this analysis separately for each ROI and analysis method (i.e. 82 comparisons). We found a between-group difference for only 1 comparison (Foot, OGRE), and it was weak enough that would disappear under any amount of multiple comparison correction (t = 2.10, p = 0.0475). Therefore, peripheral nerve injury seemed to have no systematic effect on the measures tested here.

To further address the underlying concern about between-methods effects potentially being unevenly distributed across individuals, we examined the correlation between OGRE signal intensities and FSL-only signal intensities (% BOLD signal change, Task > Rest) for each ROI. Correlations were high (Spearman's ρ 0.843 ± 0.090, range 0.533-0.954), with $\rho \geq 0.8$ in 33/42 (79%) of ROIs, indicating that the OGRE:FSL-only relationship was broadly linear. Three sample ROIs are shown in **Figure 5**, illustrating the absence of within-ROI clusters of participants with different between-analysis relationships. Overall, the advantages of OGRE analysis over FSL-only analysis did not seem to depend on subgroups or systematic differences between participants.

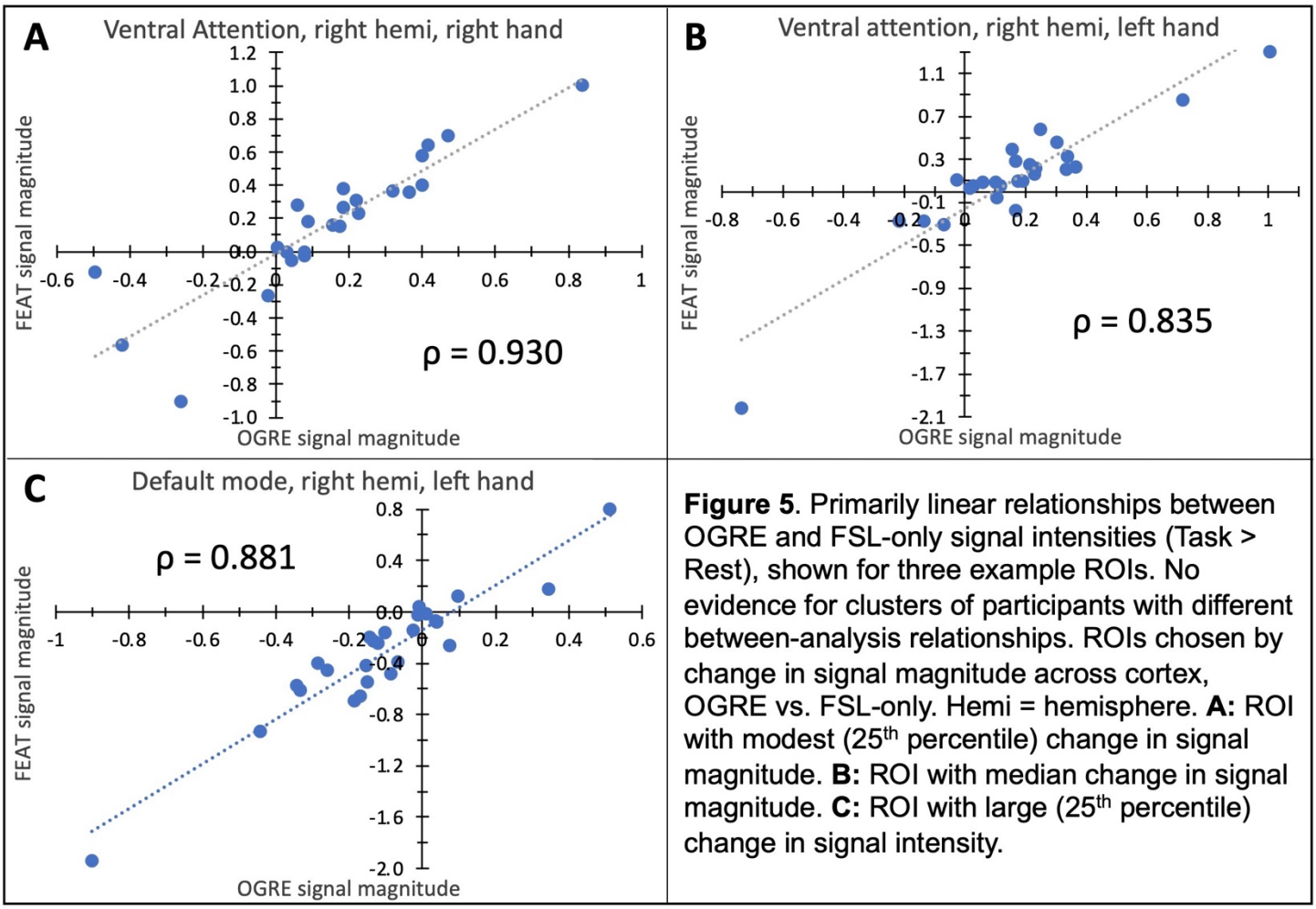

Figure 5. Primarily linear relationships between OGRE and FSL-only signal intensities (Task > Rest), shown for three example ROIs. No evidence for clusters of participants with different between-analysis relationships. ROIs chosen by change in signal magnitude across cortex, OGRE vs. FSL-only. Hemi = hemisphere. **A:** ROI with modest (25th percentile) change in signal magnitude. **B:** ROI with median change in signal magnitude. **C:** ROI with large (25th percentile) change in signal intensity.

DISCUSSION

OGRE preprocessing demonstrated the advantages of introducting FreeSurfer brain extraction and registration into FSL via the one-step resampling process. The one-step resampling process (FSL+FreeSurfer, as implemented by OGRE), compared to FSL-only preprocessing, led to fewer voxels incorrectly registered (to brain vs. non-brain), and changes in FSL FEAT higher-level results consistent with increased signal in task-relevant areas. Critically, OGRE preprocessing led to decreased inter-individual variability in BOLD signal magnitude across the brain during a motor task. OGRE software will allow any FSL user to take advantage of these preprocessing benefits.

Reduced inter-individual variability and increased task-related signal in functional analysis

We directly compared OGRE preprocessing with FSL preprocessing by analyzing the same data with the same subsequent volumetric GLM analysis (in FSL FEAT). Our ultimate finding was that OGRE led to significantly increased detection of BOLD signal in task-relevant areas of the brain (sensorimotor, parietal, and premotor areas contralateral to movement). This likely arises from reduced inter-individual variability in BOLD signal, which in turn may arise from increased precision in the process of registering individual data to standard space, as discussed further below.

OGRE also led to a statistically significant increase in BOLD signal in deep structures that likely represent noise (e.g. white matter, ventricles). This may reflect amplification of real artifacts present in the data (e.g. movement-related). While amplification of artifacts is unhelpful, it is easily distinguishable from the expected cortical areas and networks involved in our task (precision drawing), and thus the heightened signal in deep structures should not present an obstacle to interpretation of task fMRI results.

Because we used real data rather than simulated data, it is impossible to determine whether increased signal in task-relevant areas reflects a more-precise measurement of a strong signal, or a less-precise measurement of a weak signal.

However, the latter interpretation – artifactual inflation of BOLD signal strength – is incompatible with our finding that OGRE's primary impact did not primarily arise from a change in individual magnitudes (**Table 2**), but instead from decreased inter-individual variability (**Table 1**).

Overall, the decreased inter-individual variability with OGRE supports the value of introducing FreeSurfer brain extraction and registration via one-step (FSL+FreeSurfer) preprocessing, for task fMRI analysis. Inter-individual variability is a longstanding challenge in fMRI research (Dubois & Adolphs, 2016; Van Horn, Grafton, & Miller, 2008), because group averaging creates a tradeoff between statistical power and relevance to individual participants. Numerous methods exist to tackle inter-individual variability in fMRI data, but the best methods require specific data collection tools or lengthy data collection sessions (Gordon et al., 2017; Michon et al., 2022; Napadow et al., 2008), so analysis-based approaches will remain useful in many situations.

Increased precision in registration to standard space

OGRE performs its registration and brain extraction by combining the established algorithms of FSL FLIRT and FNIRT (J. L. Andersson, M. Jenkinson, & S. Smith, 2007) with FreeSurfer (e.g. Fischl & Dale, 2000; Fischl et al., 2002; Fischl et al., 2004). OGRE, unlike standard FreeSurfer, allows a direct comparison to FSL's brain extraction. The OGRE pipeline led to substantially smaller brain volume estimates for task fMRI scans, which were overwhelmingly caused by increased precision: for identification of non-brain voxels missed by FSL's BET tool, OGRE had a hit:false-positive ratio above 15:1. While this result is not "perfect," FreeSurfer is known to be more robust than FSL for identifying and quantifying gray matter (Heinen et al., 2016; Klauschen et al., 2009). Improved precision in the brain extraction process may underpin most or all of OGRE's improved fMRI results. Further improvement may be possible by adjusting the FreeSurfer settings and manual editing of the FreeSurfer output.

Use case of OGRE: improved preprocessing for researchers who prefer volumetric analysis

OGRE uses code from developed by Human Connectome Project (HCP), but should not be mistaken for an implementation of the HCP pipeline, which incorporates many other features including a surface-based approach to fMRI data analysis (Dickie et al., 2019; Glasser et al., 2018; Glasser et al., 2016). Surface-based analysis has many demonstrated benefits (Coalson, Van Essen, & Glasser, 2018), but not all analysis approaches are suitable for all research questions, groups, and protocols. Here we have quantified the advantages of the one-step resampling procedure (FSL+Freesurfer) during preprocessing, and demonstrated that those advantageous preprocessing steps can be leveraged to improve analyses with the popular, accessible FSL toolbox.

Limitations

Here we compared two preprocessing approaches, OGRE and FSL, using a motor task fMRI dataset and two traditional volumetric magnitude-based analysis methods (ROI, whole-brain). Many future directions present themselves, including using similarity-based fMRI analyses (Nili et al., 2014); and evaluating other data sets such as other tasks, task-free MRI, and simulated data with known ground truth. Given the reduced inter-individual variability provided by OGRE preprocessing, we expect it to demonstrate benefits across a broad range of analyses and datasets, and we welcome further testing of OGRE's robustness.

OGRE version 1.0.0 uses code derived from HCP 3.27.0, and as such may not reflect updates or bug fixes in future versions of the HCP code. Future updates to OGRE will link it to the main HCP repository to avoid this problem.

CONCLUSION

Task fMRI analysis can be improved with preprocessing approaches that integrate tools from FSL and FreeSurfer. In our fMRI precision drawing motor task, we demonstrated that the “one-step resampling” preprocessing approach (i.e. introducing FreeSurfer tools to augment FSL brain extraction and registration) can reduce inter-individual variability throughout the brain, and increase signal in task-related regions of the brain. We developed the OGRE (One-step General Registration and Extraction) software to provide FSL users with a turnkey general-purpose solution to

implement and integrate this preprocessing stream for improved results in FSL FEAT analysis of task fMRI data. OGRE is available online at: https://github.com/PhilipLab/OGRE-pipeline.

ENDING STATEMENTS

Information sharing statement: All software is openly available at https://github.com/PhilipLab/OGRE-pipeline. The data that support the findings of this study are available on request from the corresponding author, BAP, upon reasonable request.

Data and code availability statement: All software is openly available at https://github.com/PhilipLab/OGRE-pipeline. The data that support the findings of this study are available on request from the corresponding author, BAP, upon reasonable request.

Author contribution statement: MPM was involved in software development (primary) and manuscript writing/editing. BAP was involved in conceptualization, software development, data analysis, funding acquisition, supervision, and manuscript writing/editing. LL and RZ were involved in data analysis and manuscript editing.

Funding statement: This work was funded by NIH/NINDS R01 NS114046 to BAP.

Declaration of competing interests: Author BAP and Washington University in St. Louis have a licensing agreement with PlatformSTL to commercialize the drawing program used in this study.

Ethics statement: All participants gave informed consent, and all procedures were approved by the Institutional Review Board at Washington University in St. Louis School of Medicine.

References

Andersson, J., Jenkinson, M., & Smith, S. (2007). Non-linear registration, aka Spatial normalisation FMRIB technical report TR07JA2. *FMRIB Analysis Group of the University of Oxford*. http://fsl.fmrib.ox.ac.uk/analysis/techrep/tr07ja2/tr07ja2.pdf

Andersson, J. L., Jenkinson, M., & Smith, S. (2007). Non-linear registration, aka Spatial normalisation FMRIB technical report TR07JA2. *FMRIB Analysis Group of the University of Oxford*, *2*(1), e21.

Boesen, K., Rehm, K., Schaper, K., Stoltzner, S., Woods, R., Luders, E., & Rottenberg, D. (2004). Quantitative comparison of four brain extraction algorithms. *NeuroImage*, *22*(3), 1255-1261. https://doi.org/10.1016/j.neuroimage.2004.03.010

Coalson, T. S., Van Essen, D. C., & Glasser, M. F. (2018). The impact of traditional neuroimaging methods on the spatial localization of cortical areas. *Proc Natl Acad Sci U S A*, *115*(27), E6356-E6365. https://doi.org/10.1073/pnas.1801582115

Dickie, E. W., Anticevic, A., Smith, D. E., Coalson, T. S., Manogaran, M., Calarco, N., Viviano, J. D., Glasser, M. F., Van Essen, D. C., & Voineskos, A. N. (2019). Ciftify: A framework for surface-based analysis of legacy MR acquisitions. *NeuroImage*, *197*, 818-826.

Dubois, J., & Adolphs, R. (2016). Building a Science of Individual Differences from fMRI. *Trends Cogn Sci*, *20*(6), 425-443. https://doi.org/10.1016/j.tics.2016.03.014

Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences*, *97*(20), 11050-11055.

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., Van Der Kouwe, A., Killiany, R., Kennedy, D., & Klaveness, S. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, *33*(3), 341-355.

Fischl, B., Salat, D. H., van der Kouwe, A. J., Makris, N., Segonne, F., Quinn, B. T., & Dale, A. M. (2004). Sequence-independent segmentation of magnetic resonance images. *NeuroImage*, *23 Suppl 1*, S69-84. https://doi.org/10.1016/j.neuroimage.2004.07.016

Glasser, M. F., Coalson, T. S., Bijsterbosch, J. D., Harrison, S. J., Harms, M. P., Anticevic, A., Van Essen, D. C., & Smith, S. M. (2018). Using temporal ICA to selectively remove global noise while preserving global signal in functional MRI data. *NeuroImage*, *181*, 692-717.

Glasser, M. F., Smith, S. M., Marcus, D. S., Andersson, J. L., Auerbach, E. J., Behrens, T. E., Coalson, T. S., Harms, M. P., Jenkinson, M., & Moeller, S. (2016). The human connectome project's neuroimaging approach. *Nature neuroscience*, *19*(9), 1175-1187.

Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J. R., Van Essen, D. C., Jenkinson, M., & Consortium, W. U.-M. H. (2013). The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage*, *80*, 105-124. https://doi.org/10.1016/j.neuroimage.2013.04.127

Gordon, E. M., Laumann, T. O., Gilmore, A. W., Newbold, D. J., Greene, D. J., Berg, J. J., Ortega, M., Hoyt-Drazen, C., Gratton, C., Sun, H., Hampton, J. M., Coalson, R. S., Nguyen, A. L., McDermott, K. B., Shimony, J. S., Snyder, A. Z., Schlaggar, B. L., Petersen, S. E., Nelson, S. M., & Dosenbach, N. U. F. (2017). Precision Functional Mapping of Individual Human Brains. *Neuron*, *95*(4), 791-807 e797. https://doi.org/10.1016/j.neuron.2017.07.011

Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. [Evaluation Study]. *NeuroImage*, *48*(1), 63-72. https://doi.org/10.1016/j.neuroimage.2009.06.060

Heinen, R., Bouvy, W. H., Mendrik, A. M., Viergever, M. A., Biessels, G. J., & de Bresser, J. (2016). Robustness of Automated Methods for Brain Volume Measurements across Different MRI Field Strengths. *PLoS ONE*, *11*(10), e0165719. https://doi.org/10.1371/journal.pone.0165719

Holmes, A. J., Hollinshead, M. O., O'Keefe, T. M., Petrov, V. I., Fariello, G. R., Wald, L. L., Fischl, B., Rosen, B. R., Mair, R. W., Roffman, J. L., Smoller, J. W., & Buckner, R. L. (2015). Brain Genomics Superstruct Project initial data release with structural, functional, and behavioral measures. *Sci Data*, *2*, 150031. https://doi.org/10.1038/sdata.2015.31

Jenkinson, M. (2018). *BET User Guide*. FMRIB ANalysis Group. Retrieved August 2 from https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/BET/UserGuide

Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, *62*(2), 782-790.

Karimpoor, M., Tam, F., Strother, S. C., Fischer, C. E., Schweizer, T. A., & Graham, S. J. (2015). A computerized tablet with visual feedback of hand position for functional magnetic resonance imaging. *Frontiers in Human Neuroscience*, *9*, 150.

Klauschen, F., Goldman, A., Barra, V., Meyer-Lindenberg, A., & Lundervold, A. (2009). Evaluation of automated brain MR image segmentation and volumetry methods. *Hum Brain Mapp*, *30*(4), 1310-1327. https://doi.org/10.1002/hbm.20599

Klein, A., Ghosh, S. S., Avants, B., Yeo, B. T., Fischl, B., Ardekani, B., Gee, J. C., Mann, J. J., & Parsey, R. V. (2010). Evaluation of volume-based and surface-based brain image registration methods. *NeuroImage*, *51*(1), 214-220. https://doi.org/10.1016/j.neuroimage.2010.01.091

Michon, K. J., Khammash, D., Simmonite, M., Hamlin, A. M., & Polk, T. A. (2022). Person-specific and precision neuroimaging: Current methods and future directions. *NeuroImage*, *263*, 119589. https://doi.org/10.1016/j.neuroimage.2022.119589

Mohapatra, S., Gosai, A., & Schlaug, G. (2023). Brain extraction comparing segment anything model (sam) and fsl brain extraction tool. *arXiv preprint arXiv:2304.04738*.

Morgan, W. (1939). A test for the significance of the difference between the two variances in a sample from a normal bivariate population. *Biometrika*, *31*(1/2), 13-19.

Napadow, V., Dhond, R., Conti, G., Makris, N., Brown, E. N., & Barbieri, R. (2008). Brain correlates of autonomic modulation: combining heart rate variability with fMRI. *NeuroImage*, *42*(1), 169-177.

Nili, H., Wingfield, C., Walther, A., Su, L., Marslen-Wilson, W., & Kriegeskorte, N. (2014). A toolbox for representational similarity analysis. *PLoS Computational Biology*, *10*(4), e1003553.

NITRC. (2022). *Resting State Networks cortical and subcortical atlas (rsnatlascxsubcx: Atlas RSN v1)*. NeuroImaging Tools & Resources Collaboratory. Retrieved August 2 from https://www.nitrc.org/projects/rsnatlascxsubcx/

Philip, B. A., & Frey, S. H. (2014). Compensatory Changes Accompanying Chronic Forced Use of the Nondominant Hand by Unilateral Amputees. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience*, *34*(10), 3622-3631. https://doi.org/10.1523/JNEUROSCI.3770-13.2014

Philip, B. A., Li, F., Hawkins-Chernof, E., Swamidass, V., & Zwir, I. (2023). STEGA: Initial Validation of an iPad App to Measure Precision Drawing and Handwriting. *American Journal of Occupational Therapy*, *77*(3).

Quilis-Sancho, J., Fernandez-Blazquez, M. A., & Gomez-Ramirez, J. (2020). A comparative analysis of automated MRI brain segmentation in a large longitudinal dataset: Freesurfer vs. FSL. *bioRxiv*. https://doi.org/10.1101/2020.08.13.249474

Scholar, G. (2023). *Results of "cited by" for Jenkinson et al. 2012*. Alphabet. Retrieved September 26 from scholar.google.com

Siegel, J. S., Power, J. D., Dubis, J. W., Vogel, A. C., Church, J. A., Schlaggar, B. L., & Petersen, S. E. (2014). Statistical improvements in functional magnetic resonance imaging analyses produced by censoring high-motion data points. *Hum Brain Mapp*, *35*(5), 1981-1996. https://doi.org/10.1002/hbm.22307

Smith, J. C., & Frey, S. H. (2011, Feb 09). *Use of independent component analysis to define regions of interest for fMRI studies.* International Society for Magnetic Resonance in Medicine (ISMRM), Montreal, Canada. http://cds.ismrm.org/protected/11MProceedings/files/1618.pdf

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, *17*(3), 143-155.

Smith, S. M., & Brady, J. M. (1997). SUSAN—a new approach to low level image processing. *International journal of computer vision*, *23*(1), 45-78.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. [Review]. *NeuroImage*, *23 Suppl 1*, S208-219. https://doi.org/10.1016/j.neuroimage.2004.07.051

Van Horn, J. D., Grafton, S. T., & Miller, M. B. (2008). Individual Variability in Brain Activity: A Nuisance or an Opportunity? *Brain Imaging Behav*, *2*(4), 327-334. https://doi.org/10.1007/s11682-008-9049-9

Woolrich, M. W., Behrens, T. E., Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2004). Multilevel linear modelling for FMRI group analysis using Bayesian inference. *NeuroImage*, *21*(4), 1732-1747. https://doi.org/10.1016/j.neuroimage.2003.12.023

Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of FMRI data. *NeuroImage*, *14*(6), 1370-1386.

Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zollei, L., Polimeni, J. R., Fischl, B., Liu, H., & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*(3), 1125-1165. https://doi.org/10.1152/jn.00338.2011

Zhang, X., Feng, Y., Chen, W., Li, X., Faria, A. V., Feng, Q., & Mori, S. (2019). Linear Registration of Brain MRI Using Knowledge-Based Multiple Intermediator Libraries. *Front Neurosci*, *13*, 909. https://doi.org/10.3389/fnins.2019.00909

Zhuang, A. H., Valentino, D. J., & Toga, A. W. (2006). Skull-stripping magnetic resonance brain images using a model-based level set. *NeuroImage*, *32*(1), 79-92. https://doi.org/10.1016/j.neuroimage.2006.03.019