Reliability of fMRI for Studies of Language Recovery and Development

K. P. Eaton, PhD1, A. P. Ball, PhD2, M. Altaye, PhD1, S. K. Holland, PhD1, J. P. Szaflarski, MD, PhD2,3, B. M. Kissela, MD1, C. Banks2

1Imaging Research Center, Dept. of Radiology, CCHMC, Cincinnati, OH, 2Dept. of Neurology, Univ. of Cincinnati, Cincinnati, OH, 3Center for Imaging Research, Univ. of Cincinnati Medical Center, Cincinnati, OH

Introduction

- Quantifying changes in brain activation patterns associated with reorganization of language function over time (due to either normal development or post-insult recovery) requires a better understanding of inter-scan and inter-subject variability.
- Our goal was to quantify the reliability of fMRI as a tool for mapping neural reorganization and development by assessing the inter-scan and inter-subject variability of language activation patterns associated with verb-generation (VG) and semantic/tone decision (SDTD) tasks in healthy controls and aphasic patients with left middle cerebral artery (LMCA) infarcts.

Methods

- Variability measures are reported for fMRI activation patterns associated with VG and SDTD tasks in 4 healthy controls and 4 post-stroke aphasic subjects.
- A series of 10 fMRI scans was performed for each subject (40 total control scans for VG and SDTD each, 36 total aphasic scans for SDTD, 36 total aphasic scans for VG; one aphasic subject – 5 SDTD scans and 6 VG scans) using a 47 Varian scanner.
- T-score activation maps were thresholded at alpha = 1.96 (p = 0.05 or lower) to create maps of significant activation.
- Reliability was measured on a voxel-by-voxel basis using an intraclass correlation coefficient (ICC) [1,2] computed across all subjects and all trials for SDTD, VG, and combined SDTD/VG results.
- Regions of interest (ROIs) were computed for areas of reliable activation by spatially filtering the ICC map using a 2 mm Gaussian filter, thresholding at a value of 0.4, and taking all clusters greater than 10 contiguous voxels.
- The following benchmarks (similar to those of [3]) were used for categorizing reliability results: 0.4-0.59 (moderate reliability), 0.6-0.79 (substantial reliability), 0.8-0.99 (near-perfect reliability), and 1.0 (perfect reliability).

Results

- The patterns of reliable activation for the VG and SDTD tasks correspond well to those regions typically activated by these tasks in healthy and aphasic subjects (Fig. 1).
- ICCs for activation were consistently high (ICC ≈ 0.8) for individual tasks among both control and aphasic subjects (Fig. 2 and 3). These voxel-wise measures of reliability highlight regions of low inter-scan variability within language circuitry (i.e. regions that the language paradigm consistently activates or doesn’t activate over successive scans for a given subject).
- ICCs computed from the combined SDTD/VG data (insets of Fig. 2 and Fig. 3) were markedly reduced for both control and aphasic subjects as compared with the reliability coefficients for the individual tasks.
- The number of voxels displaying ICCs of 0.6 or greater stays relatively stable with variations in the threshold for significant activation (Fig. 1), suggesting that these methods for identifying reliable voxels are robust.
- More details and further discussion for this study can be found in [4].

Discussion

- The greater reliability measured in contralateral Broca’s homologue versus Broca’s area for aphasia supports the use of the current paradigms for studying language reorganization following left hemispheric insult, as subjects with early insult are more likely to demonstrate greater right hemispheric participation [6].
- Greater reliability in Broca’s homologue in control subjects may be explained by the more focal nature of activation in contralateral homologues shown in the repeatability data for control subjects as compared with a tendency for more diffuse activation patterns in the left hemisphere.
- The reliability of combined tasks decreases when compared to individual tasks, which may reflect the differences in task design or involvement of different structures underlying language production in the SDTD versus VG tasks.

Conclusions

One hurdle to longitudinal fMRI studies of neuroplasticity and brain development is the accurate distinction between variability arising from inter-scan and inter-subject sources and variability arising from the mechanisms of functional reorganization. This study shows excellent reliability of the fMRI language activation patterns in healthy and stroke subjects. These quantitative measures of inter-scan variability support the proposed use of these fMRI paradigms for longitudinal mapping of neural reorganization and development of language.

References


Acknowledgements

This study was supported by grants from The Neuroscience Institute in Cincinnati, OH, USA (JPS) and the NIH: R01-HD38578 (Holland) and K23-NR052468 (Szaflarski). Dr. Eaton is supported by a training grant from the NIH (T32-EB01656).

Figure 1. Group composite maps (upper figure) for control (top) and aphasic (bottom) subjects. Color scale: z = 1.96 (dark blue), z ≥ 6.0 (red). Talairach z-coordinates: +7 to +39 (left to right) for control SDTD; +3 to +35 (left to right) for all others. The graph at the bottom shows the reliability (color value) as a function of alpha for a subset of voxels with a given reliability occurring at each alpha threshold. The dashed line represents Ralpha = 0.4.

Figure 2. Control subject activation during SDTD (left) and VG (right) tasks. Top four rows (1-4) show voxel activation frequencies for each subject, while bottom row (R) shows ICCs for each task (subject 1 used for anatomical underlay). Inset at bottom shows reliability when SDTD and VG tasks are combined for each subject. Color range: dark blue (0.4), red (0.7-1.0). Talairach z-coordinates: +7 to +39 (left to right) for SDTD, +3 to +35 (left to right) for VG and SDTD/VG.

Figure 3. Aphasic subject activation during SDTD (left) and VG (right) tasks. Top four rows (1-4) show voxel activation frequencies measures for each subject, while bottom row (R) shows ICCs for each task (subject 1 used for anatomical underlay). Inset at bottom shows reliability when SDTD and VG tasks are combined for each subject. Color range: dark blue (0.4), red (0.7-1.0). Talairach z-coordinates: +7 to +39 (left to right) for SDTD and SDTD/VG, +3 to +37 (left to right) for VG.