A method for generating voxelwise between-groups statistical inferences from fMRI data using Independent Component Analysis

V. J. Schmithorst[1], S. K. Holland[1]

[1]Imaging Research Center, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

Introduction

Data-driven methods such as Independent Component Analysis (ICA) [1] are being increasingly used for the analysis of functional MRI (fMRI) data. ICA operates by linearly unmixing the data into spatially independent component maps. One advantage of ICA is that the hemodynamic response function (HRF) need not be accurately specified beforehand. In General Linear Model (GLM)-based approaches, greater flexibility in modeling the HRF comes at the cost of sensitivity. A method has also been developed to generate within-groups voxelwise statistical inferences using ICA [2]. Here we describe a method to extend ICA for the generation of voxelwise between-groups statistical inferences, or correlations with covariates of interest. The scaling ambiguity inherent with ICA is resolved by using the effect sizes, rather than the IC signal intensities, as the dependent variables.

Theory

Data for each of N subjects is grouped into an n-X-m array, assuming that there are m voxels in the brain and n time frames. Similar to the method of Calhoun *et al.* [2], the data for each subject is reduced via Principal Component Analysis (PCA) into a p-X-m array, assuming p components are present in each dataset. The data are then concatenated across subjects into a pN-X-m array, and a second PCA reduction is performed. ICA is then performed on the doubly-reduced dataset. By keeping track of the matrices involved in the PCA reductions for each subject, individual IC maps may be generated for each subject. These maps are problematic to use in a between-groups analysis since there is a scaling ambiguity in the IC maps (since the ICA model is only unique up to a scaling factor). The scaling ambiguity may be resolved, however, by considering the unmixing matrix for each subject as the pseudoinverse of a GLM design matrix, and by using the effect sizes (t-scores) of the regression parameters as the dependent variables for a second-level analysis.

Materials and Methods

Eighty-four subjects (three females and three males for each year of age from 5 through 18), were scanned on a Bruker 3T scanner performing a passive story-listening task using a thirty-second on-off block design paradigm. FMRI-EPI scan parameters were: TR/TE = 3000/38 ms, slice thickness = 5 mm, matrix = 64 X 64, FOV = 25.6 X 25.6 cm, BW = 125 kHz. The data was pre-processed by spatially filtering the data with a 4 mm Gaussian filter, and voxelwise de-meaning and normalization to a percent change from the mean. The data was reduced to 40 components per subject, and to 50 components after the second data reduction using an Expectation Maximization (EM) version of PCA (due to the large matrix size) [3]. The Fast ICA algorithm [4] was used to find the independent components. Three task-related components (using the criterion of $R > 0.65$ between the average associated time course with the on-off task reference function) were selected for display. For comparison, a GLM analysis was also performed, using the on-off task reference function (shifted by 3s to account for the hemodynamic delay) as the task regressor. A within-group analysis was performed on the t-scores for the selected components and the GLM (Figure 1, left). For voxels with p < 1e-5 (uncorrected) for cortical activation the t-scores were correlated with subject age in years (Figure 1, right) using a threshold of $p < 0.05$ and a spatial extent threshold of 7 voxels.

Results

One task-related component yields results similar to the GLM-based analysis for cortical activation and correlation with subject age, with the superior temporal gyrus bilaterally and Broca's area (left inferior frontal gyrus) exhibiting activation increases with age. Another task-related component exhibits activation in the most anterior portion of the superior temporal gyrus bilaterally, with no age-related changes detected. A third task-related component, with activation mainly in the angular gyrus, exhibits increases with age only for regions likely related to attention, in the posterior cingulate and left prefrontal cortex. A possible explanation for these findings is that syntactic processing, which recruits Wernicke's area, is a skill that develops with age during childhood [5]. Audio-visual association, which recruits the angular gyrus, may be a skill which is already developed in early childhood, prior to the age range encompassed by the study, as is melodic/contour processing, which recruits the anterior portion of the superior temporal gyrus bilaterally [6].

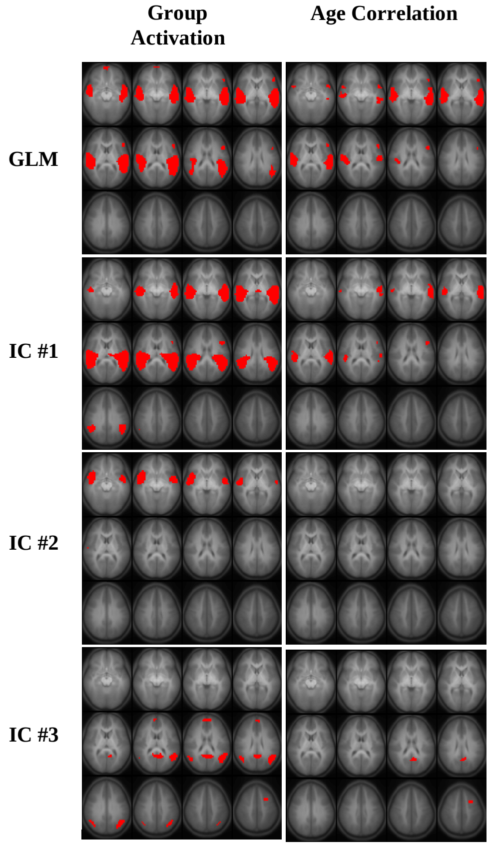

Figure 1. Composite results from GLM analysis (top) and task-related ICA components (bottom 3 rows) from 84 subjects performing a passive story listening task. For composite activation, results are significant at p < 1e-5 (uncorrected); for age correlation at p < 0.05 (uncorrected) with a spatial extent threshold of 7 voxels. All images in radiological orientation.

Discussion

As shown previously [7], the group ICA procedure results in the detection of similar, but not identical, cortical regions as compared to a standard GLM, due to the flexibility of ICA in varying the time courses across subjects. For between-group comparisons or correlations with regressors of interest, the ICA results signify changes in signal intensity across subjects, rather than reflecting how well the voxel time courses fit a previously specified model. Thus, the group ICA procedure is able to distinguish between regions with true BOLD signal intensity changes, and regions where the impulse HRF varies across groups. To aid interpretation of the results, a useful complementary test to the voxelwise analysis will be to investigate the individual subject time courses for "task-relatedness" according to some pre-specified criterion (e.g. correlation with the task reference function or power-spectrum analysis [8]) rather than relying on the subject-averaged time course.

Conclusion

A method for generating voxelwise between-groups statistical inferences using ICA is proposed. The method provides the capability of relating ICA results to demographic and environmental influences on cortical activation patterns. The method compares favorably with the GLM approach and may offer more flexibility and power since it is not dependent on the accuracy of an *a priori* model.

References

[1] McKeown M.J., *Neuroimage*, 11, 24, 2000. [2] Calhoun V.D., Adali T., Pearlson G.D., et al., *Hum Brain Mapp*, 14, 140, 2001. [3] Roweis S., *Neural Information Processing Systems*, 10, 626, 1997. [4] Hyvarinen A., *IEEE Trans Neural Networks*, 10, 626, 1999. [5] Huttenlocher J., Vasilyeva M., Cymerman E., et al., *Cognit Psychol*, 45, 337, 2002. [6] Schmithorst V.J., Holland S.K., *Neurosci Lett*, 348, 65, 2003. [7] Schmithorst V.J., Brown R.D., *Neuroimage*, 22, 1414, 2004. [8] Moritz C.H., Rogers B.P., Meyerand M.E., *Hum Brain Mapp*, 18, 111, 2003.