Impact of noise-modelling methods on test-retest reliability of a covert verbal fluency fMRI task across 2 sites

J. O'Muircheartaigh[1], C. Vollmar[2], G. Barker[3], M. Symms[2], V. Kumari[4], P. Thompson[2], J. Duncan[2], M. Koepp[2], and M. Richardson[1]

[1]Department of Clinical Neuroscience, Institute of Psychiatry, Kings College London, London, United Kingdom, [2]The National Society for Epilepsy and the Department of Clinical and Experimental Epilepsy, Institute of Neurology, University College London, London, United Kingdom, [3]Department of Clinical Neuroscience, Centre for Neuroimaging Sciences, Institute of Psychiatry, Kings College London, London, United Kingdom, [4]Department of Psychology, Institute of Psychiatry, Kings College London, London, United Kingdom

Introduction

The use of fMRI to assess language lateralisation in epilepsy surgery candidates is increasingly popular due to it's non-invasive nature and it's good concordance with the intracarotid amobarbital test (Swanson et al, 2007). Covert languages tasks, where participants perform the task silently, minimising task-correlated movement, are the preferred option for this kind of assessment as they reduce possible false motion-correlated positives. However, this application of BOLD fMRI is not appropriate unless the results from a given paradigm are both reliable and sensitive on a single-subject level. Here intraclass correlations are used to assess reliability of a simple covert verbal fluency task assessed twice at each of two sites. This is done on both a group and a single subject level

Methods

MRI: Functional images were collected at two different sites, both equipped with a GE Signa Hdx 3.0T MR Scanner, resulting in four datasets per participant. The sequence was identical at both sites, and was collected using the following parameters:. A total of 120 volumes (TE=25ms, TR=2500ms) of 50 interleaved 2.5mm slices, with a 64 x 64 matrix and a 24x24cm field of view, were collected over 5 mins of scan time. *Participants:* Ten subjects (7M, 3F) took part on four separate occasions (twice at each site), with each session separated by at least 5 days. *Paradigm:* A simple block-design (30s on) covert verbal fluency task was employed in this study. Participants tried to generated words starting with a particular letter displayed on a screen every 3s for 30s. *Data Preprocessing:* Data was preprocessed using SPM5 (www.fil.ion.ucl.ac.uk): data was realigned to the mean image and then spatially normalised to the MNI template. *Statistical Modelling:* First level statistics were modelled using both standard parametric statistics and Bayesian modelling as implemented in SPM5. Both approaches were run with (a) no motion regressors (b) 6 motion regressors (c) 24 motion regressors (Lund et al, 2005). Robust weighted least square (Diedrichsen & Shadmehr, 2005) was also used to adjust for noise. *Reliability:* Reliability between sessions was assessed on both a group level and a single subject level using intraclass correlations (as implemented in the ICC toolbox, www.brainmap.co.uk). This was assessed in three left frontal regions of interest, the whole brain, a mask of all activated voxels (using a threshold of t>4) and a mask of all deactivated voxels (using a threshold of t<-3). These maps were derived from a group t-test from the first session at one site. Single subject ICCs were only calculated for the three regions of interest.

Results

Group level ICC:, Results show that standard parametric statistics with no regressors added to the GLM gave the highest group ICCs (see figure 1). The addition of volterra expanded movement regressors proved detrimental to reproducibility. Reliability in the 3 frontal regions of interest was quite high (ranging between 0.60 and 0.85, depending on processing strategy). *Single subject ICC values:* Of the ten subjects, only two showed ICC values in the three regions of interest of <0.50. Of these two subjects, one showed consistent movement throughout the task and one was right lateralised for language.

Discussion

This simple covert language task shows good reliability on both a single-subject and a group level. The detrimental impact of motion regressors on relaibility may reflect the small amount of motion recorded in the sample. This may also be due to motion being stimulus correlated, even in the context of a "silent" task such as covert verbal fluency. Patient groups should be looked at to examine the impact of motion in these groups, which are more liable to intra-session movement. In the context of a task commonly used for presurgical mapping, it is heartening to see that the only subject showing right-lateralised activation in this task, also showed the least reliability for left-sided activation,

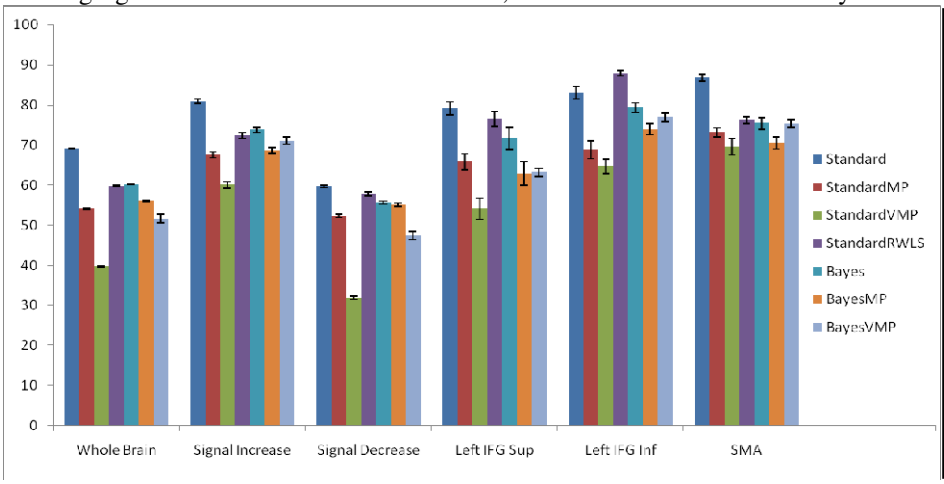

Figure 1: Group ICC values comparing 4 sessions using 7 different noise modelling approaches in the whole brain, all voxels with increased signal (t>4), all voxels with decreased signal (t<-3), and in three different ROIs – Superior and Inferior aspects of the inferior frontal gyrus, as well as the SMA. All regions are defined from group results (n=10) from one session at one site.