Large scale fMRI parameter study on a production grid[1]

Remi S. Soleman[1], Tristan Glatard[3,4], Dick J. Veltman[2], Aart J. Nederveen[1], and Sílvia D. Olabarriaga[3,4]

[1] Radiology, [2] Psychiatry and [3] Clinical Epidemiology, Biostatistics and Bioinformatics
Academic Medical Center, University of Amsterdam, NL
[4] Institute of Informatics, University of Amsterdam, NL

Abstract

Functional magnetic resonance imaging (fMRI) analysis is usually carried out with standard software packages (e.g., FSL and SPM) implementing the General Linear Model (GLM) computation. Yet, the validity of an analysis may still largely depend on the parameterization of those tools, which has, however, received little attention from researchers. In this paper, we study the influence of three of those parameters, namely (i) the size of the spatial smoothing kernel, (ii) the hemodynamic response function delay and (iii) the degrees of freedom of the fMRI-to-anatomical scan registration. In addition, two different values of acquisition parameters (echo times) are compared. The study is performed on a data set of 11 subjects, sweeping a significant range of parameters. It involves almost one CPU year and produces 1.4 Terabytes of data. Thanks to a grid deployment of the FSL FEAT application, this compute and data intensive problem can be handled and the execution time is reduced to less than a week. Results suggest optimal parameter values for detecting amygdala activation, as well as the robustness of results obtained for the difference between two studied echo times.

1 Introduction

Functional magnetic resonance imaging (fMRI [12]) is a noninvasive method for detecting brain activation that is now applied extensively in neuroscience, neurosurgical planning and drug research. fMRI detects changes in oxyhaemoglobin/deoxyhaemoglobin ratio resulting from increased local perfusion in the brain following a rise in neural activity, the so-called blood oxygenation level dependent (BOLD) contrast. Functional MR data can be acquired rapidly and spatial resolution is high, so that large data sets are generated. fMRI data is usually analysed with software packages such as fMRIB Software Library (FSL) [13] and Statistical Parametric Mapping software (SPM) [8]. Although the GUI of these packages conceals much of the complexity of the image analysis process, the choice of parameters still plays an important role in fMRI analysis. Most researchers perform the analysis using standard (default) parameter settings without questioning the role of these parameters. Comparing results for different parameter setting is in most cases impractical due to the huge amount of computing resources required for analysis and data storage.

[1]This work was carried out in the context of the VL-e project, which is supported by a BSIK grant from the Dutch Ministry of Education, Culture and Science (OC&W) and is part of the ICT innovation program of the Ministry of Economic Affairs (EZ).

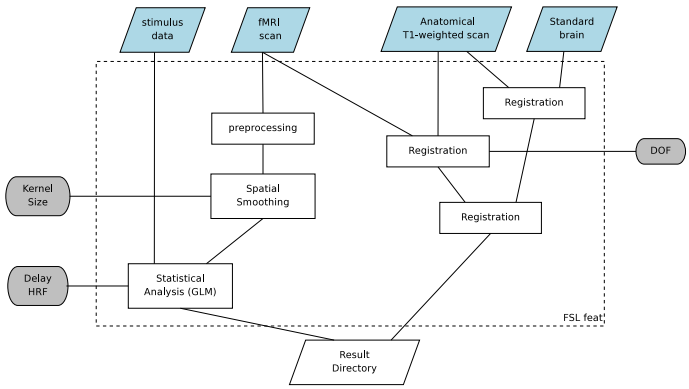

Figure 1: Simplified representation of FSL FEAT (fMRI Expert Analysis Tool) emphasising steps that are most relevant in this study. White boxes are processing steps, blue boxes are input files, and grey boxes are parameters investigated in this study.

In this paper we present a practical example of a parameter study in fMRI in which we varied three different parameters in a standard data preprocessing procedure and General Linear Model (GLM) analysis. We used for this example an emotional-provoking task which is known to activate the amygdala: a brain area mainly involved in proccessing of emotional reactions and memory. We also investigate the effect of one image acquisition parameter, namely echo time. The analysis is performed on a grid infrastructure, enabling this compute and data intensive problem to be tackled within a reasonable amount of time. This example illustrates the usefulness of such methodological experiments that employ massive computing resources to investigate the optimal parameters settings and the robustness of results obtained with fMRI.

The paper is organised as follows. Section 2 presents details of the fMRI problem and the designed parameter sweep experiment. The implementation of this experiment on a grid infrastructure is presented in Section 3. fMRI and grid performance results are presented and discussed in Sections 4.1 and 4.2. Section 5 presents initial conclusions of this study.

2 fMRI Parameter Study

In this study the data was analysed with FSL, version 3.3 [13], using its software tool fMRI Expert Analysis Tool, version 5.63 (FEAT). Each fMRI data set is first individually submitted to first-level analysis, which calculates brain activation maps as a result of several image analysis steps. The most relevant ones for this study are illustrated in figure 1 and briefly presented below. After pre-processing (including motion correction, brain extraction and slice-timing), the fMRI scans are spatially smoothed using a kernel of size S. Then, the GLM analysis is performed taking into account the HRF delay H. In GLM, a model is created to fit the fMRI scan data with respect to the timing of the stimulation paradigm employed during the scanning session. It is assumed that a good fit between the model and the data means that changes in BOLD-signal are causally related to the stimulation paradigm [12]. The output of an fMRI analysis is a brain activation map containing standardised activation probabilities (Z-scores). Normalisation to the standard brain is performed in two steps, the first of which being the alignment of the fMRI data to the anatomical scan taking into account a number of degrees of freedom D. The parameters of interest are D, S and H, which were investigated with a parameter sweep experiment.

2.1 Parameters of Interest

The first part of the parameter sweep concerns the normalisation to a standard brain. To compare individual scans at a group level it is necessary to align the individual scans to a predefined template brain. In this study, it is the standard MNI152 brain distributed with FSL. As shown in figure 1, the normalisation to the standard brain is performed through 3 different registrations. First, the high-resolution T1-weighted scan is mapped to the standard brain and to the fMRI scan using two independent registration procedures. Then, those two results are combined to produce the fMRI-to-standard-brain mapping. The **degrees of freedom (*D*)** investigated in this study only concern the fMRI-to-T1 registration. The fMRI-to-standard-brain registration is always performed using 12 degrees of freedom. Depending on the amount of correction needed it is possible to restrict manually the transformation type by adjusting D to a lower level (in FSL one can choose between 12, 9, 7, 6 or 3) to prevent failures (e.g. complete flip of an image). A wrong choice of restriction will result in a poor registration and different brain areas will be compared with each other.

The second part of the parameter sweep concerns spatial smoothing. Reducing the spatial resolution of fMRI data increases the signal to noise ratio (SNR), which improves the sensitivity. On the other hand, when using large **smoothing kernels (*S*)** the spatial resolution will be lost without gaining any profit with regard to sensitivity. In many cases a smoothing kernel $S = 6$ mm (full width at half maximum) is used for a typical voxelsize of approximately 2-3 mm.

The third part of the parameter sweep concerns the **delay of hemodynamic response (*H*)**. Functional MRI provides an indirect measure of brain activity, since fMRI detects changes in blood oxygenation level resulting from increased local perfusion following a rise in neural activity. Therefore, fMRI analysis need to correct for this delayed hemodynamic response. This is performed by convolving the modeled activity pattern in the brain with a hemodynamic response function (HRF). In general a hemodynamic delay $H = 6$ s is used, although there is evidence that this value may vary within and between subjects [15, 1].

Apart from the effects of parameter manipulation, we are also interested in comparing MRI sequences. An important parameter in an MRI sequence is the **echo time (*T*)**, indicating the time window between the transmission of a radiofrequency pulse and the signal acquisition in fMRI. In general, the usage of a sequence with a shorter echo time will result in higher signal and smaller susceptibility artifacts in the images, whereas the contrast between high and low brain activity states will tend to decrease [6]. However, it is difficult to predict from theoretical considerations which acquisition scheme should be used to obtain optimal results. In addition we suspect that the difference between these two protocols might also depend on the parameter values used for fMRI analysis.

2.2 Experiment Design

Data Acquisition. To assess both research questions we used data of healthy volunteers acquired on a Philips 3.0 Tesla Intera scanner during an event-related task paradigm (viewing of emotional pictures, International Affective Picture System (IAPS) [9]). The task was presented twice during a session with different echo time: 28 ms and 35 ms. The order of presentation was counterbalanced between subjects. Whole brain imaging was performed by scanning axial slices with an in-plane resolution of 1.9 mm and a slice thickness of 3.0 mm. For $T = 28$ ms and $T = 35$ ms we used a repetition time (time between successive volumes) of 2.7 s and 3.1 s respectively. Except for the echo times and the repetition times, all scanning parameters were kept identical. In total 22 fMRI scans $F_{i,T}$ were acquired for 11 subjects $P_i, i \in [1,11]$ with echo times $T \in \{28,35\}$.

	CPU time	Input size (MB)	Results size (MB)
Indiv. analysis	49 min	104.5	150
Group analysis	8.6 min	1650	30.5
Group difference	22.8 min	3300	55.5

Table 1: Characteristics of the individual and group analyses. CPU times have been measured on a Dual-Core AMD Opteron 2613.427 MHz with 3.5 GB of RAM. Groups summarise results of 11 individuals.

Data Analysis. Each scan $F_{i,T}$ was analysed individually using FSL FEAT first-level analysis with varying parameters: HRF delay $H \in \{2.5, 3.5, 4.5, 5.5, 6.5, 7.5, 8.5, 9.5\}$, smoothing kernel size $S \in \{2,3,4,5,6,7,8,9,10,11,12\}$, and degrees of freedom for registration $D \in \{3,6,7,9,12\}$. Results of this phase consist of activation maps $\alpha_{i,T}(H,S,D)$ obtained with each combination of parameter values (H,S,D).

After all the individual activation maps were computed, an average activation map was calculated for groups of results using FSL FEAT high-level analysis. Results obtained for all P_i with identical (T,H,S,D) are averaged, generating group activation maps denoted as $\gamma_T(H,S,D)$. Finally, the results obtained with two echo times were compared using FSL FEAT group difference analysis in a similar way. The generated difference maps are denoted $\Delta(H,S,D)$.

Results Evaluation. The comparison of results obtained with different parameters was performed based on a region of interest with robust activation for the adopted stimulus. The IAPS is a picture set containing emotion-provoking pictures, both positive (e.g. erotic scenes) and negative (e.g., mutilations, attack scenes), which has often been used in emotion research. Previous studies have shown that viewing IAPS pictures vs. baseline activates regions within the brain that are involved in emotional processing such as the amygdalae [3, 10]. The amygdalae are almond shaped groups of nuclei within the left and right medial temporal lobe which are one of the key elements of emotion processing. Therefore, the mean Z-scores in both of the amygdalae were used as reference to compare activation maps obtained with different parameter settings. The amygdalae were located with a predefined anatomical atlas (AAL, [14, 7]) on the MNI152 template and the created mask was applied to the group activation maps $\gamma_{28}(H,S,D)$, $\gamma_{35}(H,S,D)$ and $\Delta(H,S,D)$, to extract the Z-scores within this region of interest and calculated their mean denoted $\mu_{28}(H,S,D)$, $\mu_{35}(H,S,D)$ and $\mu_D(H,S,D)$.

3 Grid implementation

The designed experiment involves large computation effort. The individual analysis is the most compute-intensive one due to pre-processing, registration and linear model computation. The group analysis is the most data intensive one, since it needs to access all the individual analysis results for averaging (See summary in table 1). Parameter sweeps are costly: in particular, when the parameter space has several dimensions (three in this study), widening the range of one of them swiftly increases the size of the computing problem. As a side effect, storing the produced data is often not manageable without a specific infrastructure. The total estimated resources for this study add up to almost one CPU-year and 1.4 Terabytes of data. Obviously, this experiment could not have been performed without a high performance infrastructure.

Production grids have been designed to support such computing and data needs, but they are still seldom used by medical imaging researchers for such studies. Roadblocks need to be identified and carefully targeted by

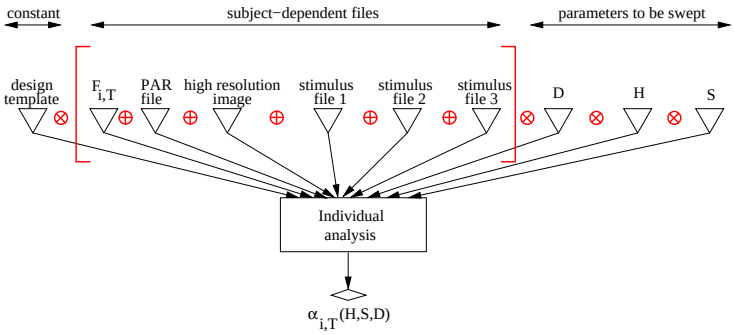

Figure 2: Individual analysis workflow described in Scufl. Iteration strategies are drawn in red ($\otimes$ is the cross-product and $\oplus$ is the dot product). Scufl iteration strategies allows for easily distinguishing sweeping parameters from data-related inputs.

the grid community with new developments. To do that, the Dutch Virtual Laboratory for e-Science (VL-e[1]) set up a tight interaction between domain scientists and grid developers and grid service providers. The project gives access to a Proof-of-Concept grid infrastructure that is part of EGEE. It also offers a fertile ground for exchange of expertise to develop and deploy grid-enabled applications in answer to real demands of domain scientists as it is the case in this fMRI study.

3.1 Grid Application Design

As part of the VL-e software distribution, the FSL package is pre-installed in all computing nodes, therefore scripts can be submitted directly as grid jobs to perform FSL FEAT or other image analysis tasks. The gLite command line utilities could have been used to submit and monitor jobs on the grid, however this approach was found to be too unfriendly. Instead, we adopted a workflow approach by wrapping image analysis tasks as web-services, composing workflows to describe the parameter sweep experiment, and enacting the workflows on the production grid. More details are presented below.

The Scufl workflow language from the Taverna workbench [11] was used to describe workflows. Even if this use-case does not exhibit strong workflow requirements (like complex data dependencies between components), using such a language is useful to describe the parameter sweep in an elegant way. In short, Scufl supports lists to be indicated as inputs, and the workflow is iterated for each list element. Operators specify how list elements should be combined as workflow inputs. The "cross product" $A \otimes B$ indicates that the workflow is iterated for all combinations of elements $a_i \in A$ and $b_j \in B$, and the "dot product" $A \oplus B$ indicates that the workflow is iterated for pairs (a_i, b_i). Using such a generic workflow framework instead of ad-hoc scripts is a step towards an autonomous usage of the grid by end-users, which is an important aim of our research. Two workflow descriptions were used to implement the parameter sweep experiment, one for the individual analyses and one for the group analyses.

The **individual analysis workflow** takes 10 inputs and produces a single output (see figure 2). The performed task consists of FSL FEAT first-level analysis and some data logistics (download/upload data). Six of the inputs correspond to the fMRI scan $F_{i,T}$ (100 MB) and other subject-related data: MRI acquisition parameters (1.5 MB), the high resolution anatomical (T1-weighted) scan (1.5 MB) and 3 files with stimulus timing information (500 bytes each). Three other inputs correspond to the parameters (H,S,D) to

[1] `http://www.vl-e.nl`

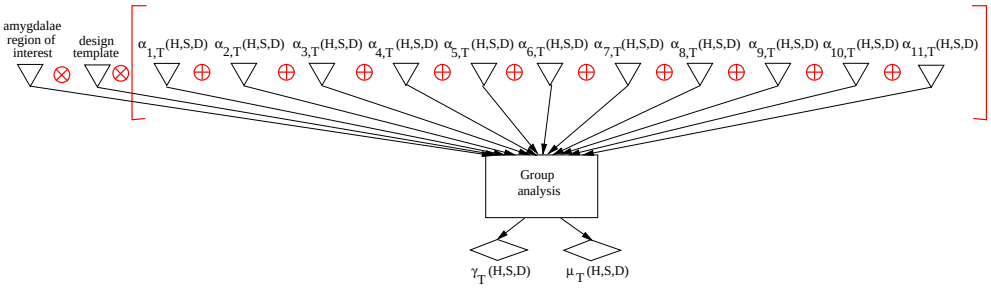

Figure 3: Group analysis workflow described in Scufl.

sweep upon. The last input (design file) contains all the parameters for FSL FEAT that remain constant in all analyses. Each individual analysis produces a single directory with all the output files (150 MB when compressed), one of them being the brain activation map $\alpha_{i,T}(H,S,D)$ used in group analyses. In this experiment, the inputs assume lists of values associated with different scans, subjects, and parameters. The iteration strategies of Scufl allow to express the sweep on individual analysis tasks in a simple way: the 6 first data-related inputs are combined with the dot product operator, and the other parameters are swept with the cross product.

The **group analysis workflow** used to calculate average activation has 13 inputs and two outputs – see figure 3. Besides running the FSL FEAT high-level analysis, this workflow also calculates the mean value in the amygdalae using with the FSL AVWMATHS utility. Eleven inputs indicate directories containing pre-computed results of individual analyses for 11 fMRI scans $F_{i,T}$. These inputs are provided by lists in which the elements correspond to directories with results computed with the same parameter settings (H,S,D). The dot product is used to guarantee that the activation maps averaged during the workflow iteration are $\alpha_{i,T}(H,S,D), i \in [1,11]$. The remaining inputs are constant in this experiment, namely the design file with parameters for FSL FEAT, and the mask indicating the voxels in the amygdalae. The group analysis produces 2 outputs: a directory with results (30 MB when compressed), including the group activation map $\gamma_T(H,S,D)$, and a file containing the mean Z-score $\mu_T(H,S,D)$ in the amygdalae region.

Similarly to the above, a workflow with 23 inputs and 2 outputs is used for **group difference analysis** to compare results obtained with different echo times T. Twenty two inputs indicate directories with pre-computed individual analyses corresponding to two groups fMRI data $F_{i,28}, F_{i,35}, i \in [1..11]$ acquired with different echo times. The outputs include the difference activation maps for these two groups ($\Delta(H,S,D)$) and the mean values in the amygdalae ($\mu_D(H,S,D)$).

3.2 Application Deployment

The grid infrastructure of the Dutch VL-e project was used to run the experiment. At the time the experiment was performed the `vlemed` Virtual Organisation (VO) had access to 687 CPUs spread over 4 sites through 8 batch queues federated by the gLite middleware. As a production grid, it is shared among several VOs, therefore there is no way to control the number of available CPUs at a given instant. Four Storage Elements (SE) accessible through an SRM interface are also available and files can be organised in a single Logical File Catalog (LFC) independently from their physical location. In addition, a gridFTP and a Storage Resource Broker servers are also available.

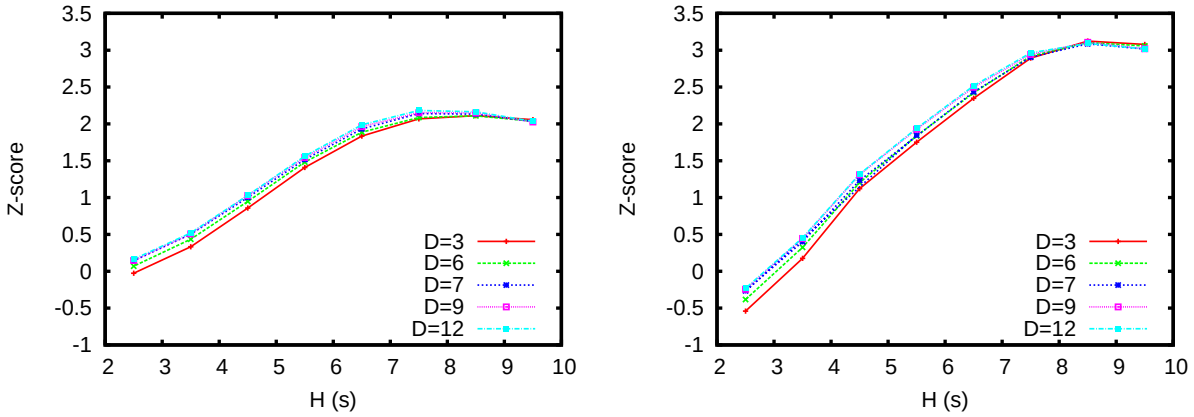

Figure 4: Influence of degrees of freedom D in the mean activation in the amygdalae as a function of delay for $T = 28$ ms using smoothing kernels $S = 5$ mm (left) and $S = 11$ mm (right).

Porting the image analysis application to the grid was easily done using the Generic Application Service Wrapper (GASW) [4], which integrates executables in Scufl workflows via a command-line descriptor and the Web Service Description Language (WDSL). The FEAT script was wrapped as shipped by the Fhttp://www.thefreedictionary.com/timeSL distribution. Workflows are executed on this infrastructure using the MOTEUR workflow engine [4] and adopting the software architecture detailed in [5]. The various storage resources are accessed via the homogeneous Virtual File System interface provided by the VL-e software Toolkit[2].

4 Results and Discussion

4.1 fMRI results

Figure 4 illustrates the mean activation within the amygdalae using different degrees of freedom D for the fMRI-to-anatomical scan registration. The plot shows that no significant activation differences were observed, a pattern that is also observed for other smoothing kernel sizes. In particular, it should be noted that a registration with $D = 3$ (translation only) produced similar results as obtained with $D = 9$ and $D = 12$. Given that both scans belong to the same subject and have been acquired during a single scanning session, only minor differences are expected, which could explain why only translations seem sufficient to capture them. The main contribution to the normalisation process is likely to be the anatomical-to-standard-brain registration, the second step in the registration process, which was fixed to 12 degrees of freedom in this study. A further analysis of the complete registration process will be part of our future work.

Figure 5-left shows the mean activation within the amygdalae obtained with different spatial smoothing kernel sizes and HRF delays. Note that a maximum for the Z-score is obtained for a delay of $H = 8.5$ s, which differs from the value adopted by default in FSL feat ($H = 6$ s). The figure also shows an activation increase when using larger spatial smoothing kernels. A visual inspection of the results (figure 6) suggests that with $S = 12$ mm spatial resolution is still acceptable and that even larger smoothing kernels might further increase sensitivity.

[2] `http://www.science.uva.nl/~ptdeboer/vlet`

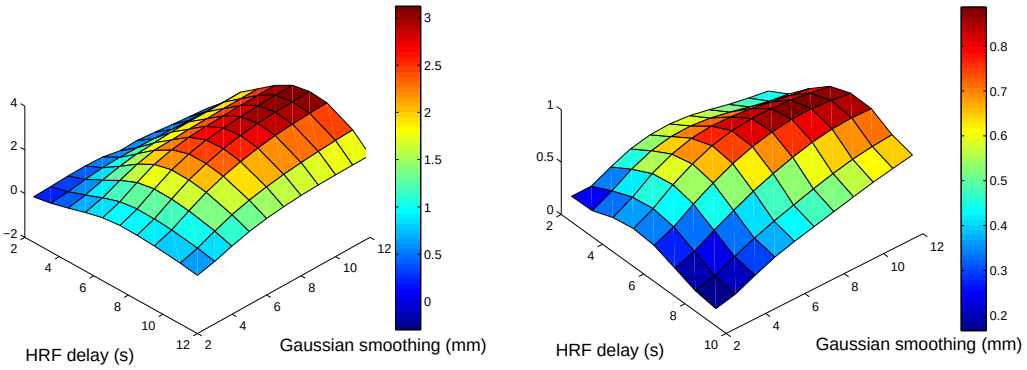

Figure 5: Results obtained for group analyses for varying smoothing kernel sizes S and HRF delays H. Left: Mean activation (Z-scores) in the amygdalae for $T = 28$ ms, $D = 12$. Right: Mean differences between amygdalae activation with $T = 28$ ms and $T = 35$ ms.

Our second research question concerns the effects of manipulating echo time (T). Pairwise T-tests were used to compare results obtained with both echo times. No significant differences in activation between the scan sequences could be detected, as shown in figure 5-right: none of the parameter combinations cross the threshold of significance. We used a Z-value of 2.33 (which equals a probability of 0.01) as threshold to distinguish noise from an actual signal. ([16]).

4.2 Grid performance

Table 2 presents a summary of performance results of the individual and the group analyses grid application as measured during the execution of this parameter study on the VL-e grid. A total of 9680 individual analyses, 880 group analyses and 440 group differences was computed. CPU time was benchmarked on a node of the infrastructure with representative performance.

Individual analyses were submitted in 4 batches of about 2500 jobs each to avoid infrastructure flooding. The global speed-up of 115 times enabled us to compute the individual analyses in less than 3 days. Note that this is a compute-intensive experiment, therefore the speed-up value provides an indication of the number of CPUs that were used concurrently. Note also the small job failure rate of 0.5% measured for this experiment.

Although pleasantly parallel, group analyses faced quite poor speed-ups (global 2.9). One of the reasons is they are dominated by data transfers, since all individual analyses are downloaded to the worker node before executing FSL FEAT. Even if each analysis only transfers about 1 GB, concurrently initiating hundreds of

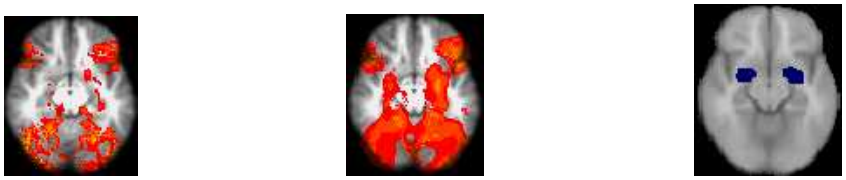

Figure 6: Impact of the smoothing kernel size for $H = 2.5$ s and $D = 12$. Left and Centre: slice of the brain activation map respectively for $S = 5$ mm and $S = 12$ mm smoothing. Right: slice of the amygdalae region of interest (in blue) overlaid on the template brain.

	$\#P_i$	#T	#S	#D	#H	# Analyses	CPU (days)	Data (TB)	Elapsed (hours)	Speed -up	# Submit Jobs	Failure (%)
Individual Analyses												
batch 1	11	1	5	5	8	2200	74.9	0.31	14.9	120.5	2200	0.00
batch 2	11	1	6	5	8	2640	89.8	0.38	11.6	186.6	2642	0.08
batch 3	11	1	6	5	8	2640	89.8	0.38	32	67.38	2687	1.75
batch 4	11	1	5	5	8	2200	74.9	0.31	10.2	176.8	2203	0.14
total	11	2	11	5	8	9680	329.4	1.38	68.7	115	9732	0.53
Group Analyses								(MB)				
batch 1		1	6	5	8	240	1.4	7.1	8.0	4.3	401	40.15
batch 2		1	6	5	8	240	1.4	7.1	9.5	3.6	240	0.00
batch 3		1	5	5	8	200	1.2	6	14.9	1.9	200	0.00
batch 4		1	5	5	8	200	1.2	6	11.3	2.5	600	66.67
total		2	11	5	8	880	5.2	26.2	43.7	2.9	1441	38.93
Group Difference Analyses								(MB)				
batch 1			11	5	8	440	7	23.8	44.3	3.8	2650	83.40

Table 2: Performance results for individual, group and difference analyses for each job batch: number of subjects, echo times, smoothing sizes *S*, degrees-of-freedom *D*, HRF delays *H*, and successful analyses; total CPU time, amount of produced data, total elapsed time from first submitted to last completed job, speed-up of the experiment (CPU/elapsed), and job failure rate ((#submitted - #successful) / #submitted).

such data transfers rapidly reduces the data servers throughput. Moreover, one of the Storage Elements limited the number of concurrent connections, causing jobs to fail and partially explaining the high 38.93% global failure ratio. This problem was even larger in the case of group difference analyses, where each of the 440 jobs concurrently attempted to download more than 3 GB. The global failure rate here was 83%. In spite of these problems, the experiment could be carried out in a reasonable time in the end, which could not have been possible without adopting a grid implementation. All in all, those failure ratios look good compared to other medical experiments conducted on EGEE (see *e.g.* [2]). Two main reasons can explain such a difference. First, our experiment is set up in a quite controlled environment, with pre-installed software and strong support team on every grid site. Second, it has been carefully tuned in order to avoid some of the problems encountered in [2]: for instance, the total workload has been split into batches of reasonable size to avoid proxy expirations.

5 Conclusion

In this paper, the benefit of performing parameter sweeps for methodological fMRI studies has been established by assessing amygdala activation for different analysis parameter choices. Detecting activation in this brain area is in general cumbersome, emphasising the need of a good choice of parameters to obtain reliable results. Our initial results indicate that the optimal values deviate from the default parameters that are typically used in this type of analysis in FSL FEAT. In the future we plan to perform this analysis for different fMRI paradigms and different brain areas and expect to find different optimal parameter values. Additionally, we could study the robustness of the difference between two different MRI scan protocols over different parameter sets for analysis. Results indicate that the observed difference between the considered

echo times is not significant and does not depend on the parameters used for analysis.

Although the grid implementation enabled the experiment to be completed within a reasonable time, the set-up is still far from ideal. Data transfers, rather than computation, seem to be the limiting factor now and data distribution or even replication among several Storage Elements might be a way to improve their reliability. Then, a proper strategy would have to be determined in order to avoid blind replication of terabytes of data. Moreover, failure management is still a challenging issue in grid applications deployed in production grids, with the consequence that much (expert) manual intervention is still needed to complete the experiment. This not only causes a significant overhead in the management of this experiment, but also makes the dream of having end-users performing medical image analysis on grids seem further away than we initially hoped for. We continue to work on improvements in the grid application to support the many experiments to come motivated by the promising results obtained in the study reported here.

References

[1] G. Aguirre et al. The variability of human bold hemodynamic responses. *NeuroImage*, 8(4):360–369, 1998.

[2] G. Aparicio et al. A Highly Optimized Grid Deployment: the Metagenomic Analysis Example. In *Global Healthgrid: e-Science Meets Biomedical INformatics (Healthgrid'08)*, pages 105 –115, Chicago, USA, May 2008. Healthgrid, IOS Press.

[3] H. Breiter et al. Response and Habituation of the Human Amygdala during Visual Processing of Facial Expression. *Neuron*, 17(5):875–887, 1996.

[4] T. Glatard et al. A Service-Oriented Architecture enabling dynamic services grouping for optimizing distributed workflows execution. *Future Generation Computer Systems*, 24(7):720–730, 2008.

[5] T. Glatard et al. Workflow integration in VL-e medical. In *21st IEEE International Symposium on Computer-Based Medical Systems (CBMS'08)*, pages 144–146, Jyväskylä, Finland, June 2008.

[6] M. Gorno-Tempini et al. Echo Time Dependence of BOLD Contrast and Susceptibility Artifacts. *NeuroImage*, 15(1):136–142, Jan. 2002.

[7] C. Holmes et al. Enhancement of MR images using registration for signal averaging. *Journal of Computer Assisted Tomography*, 22(2):324–333, Apr. 1998.

[8] F. K.J. Statistical parametric mapping and other analysis of functional imaging data. In *Brain Mapping: The Methods*, pages 363–385. Academic Press, 1996.

[9] P. Lang et al. International affective picture system (IAPS): Technical manual and affective ratings. Technical report, University of Florida, Center for Research in Psychophysiology, Gainesville, 1999.

[10] I. Liberzon et al. Extended amygdala and emotional salience: a PET activation study of positive and negative affect. *Neuropsychopharmacology*, 28(4):726–33, Apr. 2003.

[11] T. Oinn et al. A tool for the composition and enactment of bioinformatics workflows. *Bioinformatics journal*, 17(20):3045–3054, 2004.

[12] S. Smith. Overview of fMRI analysis. *The British Journal of Radiology*, 77:S167–S175, 2004.

[13] S. Smith et al. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23(1):S208–S219, 2004.

[14] N. Tzourio-Mazoyer et al. Automated anatomical labelling of activations in spm using a macroscopic anatomical parcellation of the MNI MRI single subject brain. *NeuroImage*, 15(1):273–289, Feb. 2002.

[15] M. Woolrich et al. The variability of human bold hemodynamic responses. *NeuroImage*, 21:1748–1761, 2004.

[16] K. Worsley et al. A three-dimensional statistical analysis for CBF activation studies in human brain. *Journal of Cerebral Blood Flow and Metabolism*, 12:900–918, 1992.