A retrospective evaluation of clinical functional MRI quality and analysis methods

Victoria L Morgan[1,2], Megan Strother[2], and Reid C. Thompson[3]

[1]*Vanderbilt University Institute of Imaging Science, Vanderbilt University, Nashville, TN, United States,* [2]*Radiology and Radiological Sciences, Vanderbilt University, Nashville, TN, United States,* [3]*Neurosurgery, Vanderbilt University, Nashville, TN, United States*

Target Audience - This work would be of most interest to those who perform functional MRI for presurgical mapping of eloquent cortex.

Purpose – Due to the high variability in the anatomy and physiology of the patients who undergo functional MRI (fMRI) for clinical presurgical mapping of eloquent cortex, it is difficult to determine metrics for quality evaluation and to compare different analysis strategies. The purpose of this work is to 1) present a method for assessing quality of an individual dataset, and 2) to determine whether different analyses affect the amplitude and the relative spatial distribution of the activated regions.

Methods - We retrospectively investigated the fMRI results of patients who were referred for clinical presurgical functional mapping of finger tapping (n=84), face motor (n=82), reading (n=110), and word generation (n=131) tasks at our institution from 2006 to the present. Functional images were acquired using a 3T MRI scanner (Philips Healthcare, Inc., Best, Netherlands) using a T2* weighted, gradient-echo, echo planar fMRI scan (64 x 64, 3.75 mm x 3.75 mm, FOV = 240mm, 4.5 mm thick/0.5 mm gap, TE = 35 ms, TR = 2 sec, 100 volumes per series). Tasks were performed in blocks of 20 sec control/ 20 sec task repeated five times. Self-paced finger tapping was performed by touching fingers to thumb of either dominant (n=42), non-dominant (n=32) or both hands (n=10) contrasted with rest. Face motor involved the subject moving their cheeks, tongue and mouth contrasted with rest. The reading task involved reading descriptions of nouns (i.e. "a long yellow fruit – banana") contrasted with strings of symbols (i.e. "$%$!@#"). During the word generation task, the subject silently generated words that were related to the given category (i.e. "food") contrasted with rest. Preprocessing included slice timing correction, motion correction and 7mm spatial smoothing using SPM8 software [1].

For each dataset we examined 4 parameters: motion, temporal signal to noise ratio during the control periods (tSNR), amplitude of the activation map *t*-values from two analyses, and the relative spatial similarity of *t*-maps of the two analyses. The motion was quantified using framewise displacement (FD) [2]. The FD was the distance from one "frame" or time point to the next summed over the x, y, z translation and the x, y, z rotation after converting the angles to mm on the surface of a sphere of radius 50 mm [2]. We then averaged the FD across the 100 frames (FD Index, FDI).The tSNR of each dataset was the mean divided by the standard deviation of the first control block and the last 6 frames of each of the other four control blocks averaged across all voxels in the brain. We performed two analyses to determine the activated voxels for each dataset. Our original analysis uses the general linear model (GLM) of SPM8 including the block regressor and the 6 motion regression time series as confounds ("orig"). The second analysis was the same except with no motion confounds ("no_mot"). This represents some scanner-based and commercial clinical packages. From these we determined the average of the top 2% of *t*-values in the SPM *t*-map, and we performed a Pearson's correlation between the SPM *t*-maps calculated using the two GLM analyses (orig2no_mot) to quantify similarity between the maps.

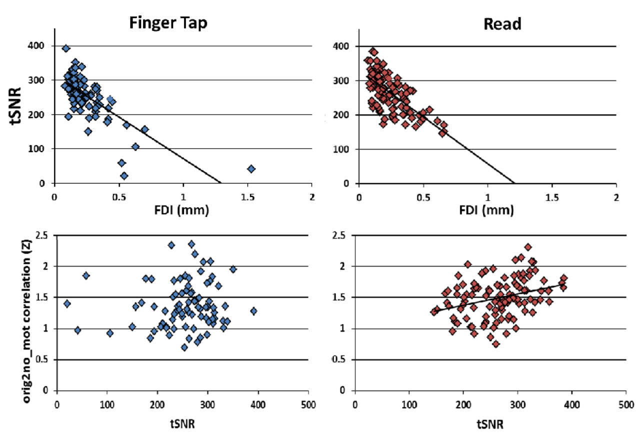

Figure 1. Relationship between signal (tSNR), motion (FDI) and spatial activation maps (orig2no_mot correlation) for one motor and one language task.

Results - In all four tasks, the FDI was significantly correlated with the tSNR (p<0.0001) with essentially the same linear regression line (Fig 1 – top). An FDI of 0.5 mm corresponded to a tSNR of approximately 200 in each task. Using a minimum tSNR of 200 (approximately 0.5 mm maximum FDI) as a threshold, 85% of both the finger tapping and reading datasets, 81% of the word generation datasets, and 70% of the face motor datasets can be categorized as having acceptable motion and tSNR.

For the finger tapping and face motor tasks, the *t*-values in the orig maps were not linearly related to tSNR and were not significantly different than in the no_mot maps. There was also no significant decrease in correlation between the orig and no_mot maps as tSNR decreased (Fig 1 – bottom, left). However, in both language tasks the *t*-values were reduced in the no_mot maps compared to the orig maps (paired *t*-tests). Also, in these language tasks, the correlation between the orig and no_mot maps linearly increases as the tSNR increases (read p<0.0001, word gen p=0.001)(Fig 1 – bottom, right).

Discussion and Conclusions – We showed that tSNR linearly decreases with increased motion in all four tasks we used, and that the face motor task has the highest potential for head motion and decreased tSNR. In both of the motor tasks, the *t*-values and the activation maps are similar using the GLM with and without motion regressors. But when performing the language tasks, the GLM without the motion regressors has decreased *t*-values and less similarity compared to the GLM with the motion regressors, especially at lower tSNR values. This suggests that when performing motor tasks the focal activation is robust across levels of tSNR and motion, and that either GLM will give the same results; but when performing language tasks, the activation levels and maps are not the same with the two analyses. Furthermore, the higher *t*-values using GLM with the motion regressors implies this analysis may be more robust.

References - [1] http://www.fil.ion.ucl.ac.uk/spm/software/spm8/ [2] Power JD, et al. NeuroImage 2012; 59:2142-2154.