Using Reliability to Predict Validity in Clinical fMRI

Tynan Reid Stevens[1], Steven D Beyea[1], Ryan CN D'Arcy[1], and David B Clarke[2]
[1]IBD-Atlantic, NRC, Halifax, Nova Scotia, Canada, [2]QEII Health Science Centre, Halifax, Nova Scotia, Canada

Background: Functional MRI (fMRI) has been identified as a potentially useful tool for pre-surgical mapping. To establish the validity of fMRI for mapping critical functional zones, comparisons to cortical stimulation (CS) have been performed with a high degree of success [1,2]. However, setting thresholds for clinical fMRI remains a significant challenge, as the optimal threshold may vary between individuals, tasks, scanners, and sessions [1]. We hypothesize that setting thresholds based on test-retest reproducibility of activation patterns has the potential to identify critical functional zones reliably [3]. We aim to use test-retest ROC analysis to determine optimal thresholds using the ROC-r method [4,5], and compare the correspondence to CS achieved by traditional fixed thresholds with the automated ROC-r method.

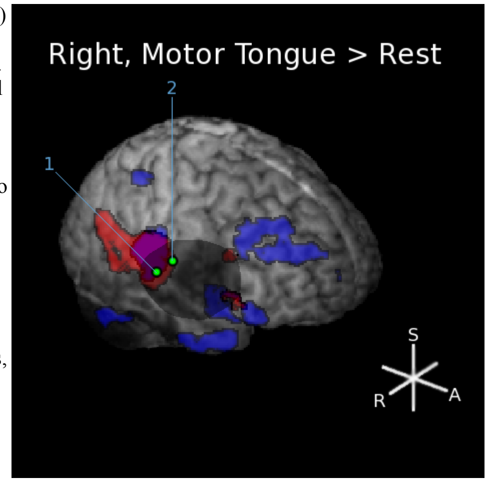

Figure 1: Example patient results. The test-retest fMRI images from a tongue movement paradigm are shown using ROC-r retest thresholds (image 1, red, t=1.7; image 2, blue, t=4.9), as well as the locations at which cortical stimulation caused speech arrest (1,2). Despite some discrepancy between sessions, the reliable activation in the pre-central gyrus (purple) accurately predicted the CS results.

Methods: Eleven patients with brain tumors (37 +/- 13 years, 6 female) were imaged by test-retest fMRI prior to receiving surgical treatment as part of their standard clinical care. Cortical stimulation was observed during the surgery, and recorded onto the MR images by means of an image guided surgery system. For ten patients, CS locations were recorded by means of the image guidance system, and corrected for brain shift. The remaining patient's CS results were recorded by digital photograph of tags placed on the cortical surface throughout CS. These tags were then digitized onto the corresponding locations in the MR image.

MR imaging was performed with a 4 T MRI. Structural (MP-FLASH; TI=500ms, TR=10ms, TE=5ms, α=11º, 256x256x170 matrix, 24x24x17cm FOV) and all functional (two-shot spiral, TR=2s, TE=15ms, α=60º, 64x64x22 matrix, 3.75x3.75x4mm voxels, 0.5mm gap) imaging were performed in a single session. Functional tasks included finger-tapping and tongue movement for motor mapping (20s blocks, 4 motor, 5 rest blocks), and comprehension (18s blocks, 6 sentence, 6 control, 7 rest blocks), and production tasks for language function (16s blocks, 6 naming, 6 control, 7 rest). The finger tapping task consisted of sequential thumb-to-digit tapping with either the left or right hand at 1Hz. The tongue movement task required the patient to move their tongue from one side of their mouth to the other at 1Hz. The language tasks used active control task blocks in addition to rest blocks to allow subtraction of non-essential task components. The comprehension task required patients to view short english sentences that were 50% correct (e.g. "He stroked her face with a feather"), and 50% incorrect (e.g. "John spread his ceiling on the floor."), and categorize the stimuli using one of two buttons. The control task required the same decision process and response, but presented single digit math statements that were correct (e.g. 2+2=4) or incorrect (e.g. 6-3=4). The language production task consisted of 3D renders of common objects, which the subjects were asked to name out-loud. The control task presented 3D nonsense shapes, and subjects were asked to passively view the objects without giving names to them.

Functional MRI data were motion corrected, and registered to the high-resolution T1 weighted image using a 12 parameter affine transformation. The functional images were segmented into brain and non-brain tissue to provide a mask for fMRI analysis. The fMRI images were smoothed with a 6mm FWHM kernel prior to GLM analysis with AFNI.

Cortical stimulation was performed using an Ojemann OCS-1 stimulator, with a bipolar probe. 5 Hz stimulation was used for sensory-motor mapping, and 50-60Hz for language mapping. Pulses were delivered starting at a current of 4 mA peak-to-peak, and increased until an effect was elicited, or a maximum of 20 mA was reached. Involuntary movements or strange sensations were recorded for sensory-motor mapping. For language function, the patients were asked to count, or recite days of the week or months of the year. Any disruption of normal performance on these tasks was recorded if reproducible. Correlation between fMRI and CS was measured by the distance from positive CS locations to the nearest fMRI activation. The CS-to-fMRI distance was estimated for fixed thresholds, as well as for the ROC-r optimized thresholds.

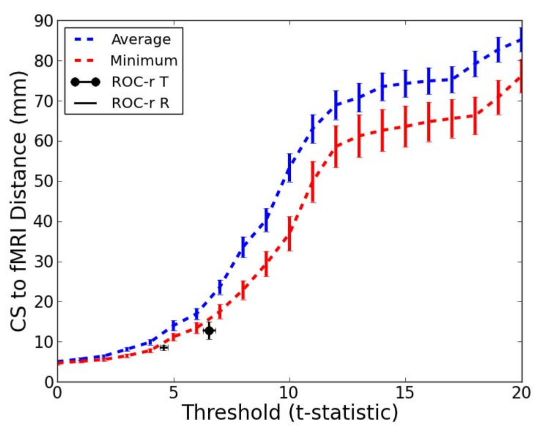

Figure 2: Distance versus threshold, +/- std. error. The blue curve represents the average CS-to-fMRI distance across all images of the test-retest pairs, at fixed threshold. The red curve shows the average distance when both images of the test-retest pair are overlaid simultaneously. The two points shown in black are the distances found when the ROC-r template (circle) and retest (no symbol) thresholds are used.

Results: Two of the patients had no response to cortical stimulation. For the remaining nine patients, CS-to-fMRI distance was calculated. An example patient dataset is shown in figure 1, and the average results are shown in figure 2. The minimum distance achieved between CS and fMRI was 5 +/- 3 mm, and increases with threshold. Using the minimum distance from the two test-retest images decreases the CS-to-fMRI distance at all thresholds. The ROC-r method returned template thresholds of 6.5 +/- 0.3, and retest thresholds of 4.6 +/- 0.2. The resulting CS-to-fMRI distance was 12 +/- 2 mm and 8.6 +/- 0.5 mm for the ROC-r template and retest thresholds respectively. At the same thresholds, the best distance achieved by the fixed confidence methods was 16 +/- 2 mm and 10.0 +/- 0.8 respectively.

Discussion & Conclusion: The minimum CS-to-fMRI distance of 5 mm observed suggests that some CS points are being localized outside the fMRI analysis volume. This is likely due to errors in the CS localization (e.g. brain shift), or caused by brain-masking in the fMRI analysis reducing the analysis volume. Test-retest imaging served to improve the correlation between fMRI and CS. Overlaying both images of the test-retest pair simultaneously reduced the CS-to-fMRI distance simply by expanding the activated volume compared to either replication alone. Optimizing the test-retest thresholds using the ROC-r method resulted in a further reduction of the CS-to-fMRI distance, supporting the hypothesis that reproducibility-optimized thresholding leads to reliable and valid activation patterns.

References: [1] Rutten GJM, et al, 2002. Ann Neurol, 51, 350-60. [2] Guissani C, et al, 2010. Neurosurgery, 66(1), 113-20. [3] Genovese CR, et al., 1997. Magn Reson Med, 38, 497-507. [4] Stevens T, et al., 2011. NeuroImage, submitted. [5] Stevens T, et al, 2010. ISMRM poster #1148.