An Index of Scanner/Site Differences in fMRI Sensitivity: Method and Implications

L. Friedman[1,2], F. BIRN[3]

[1]The MIND Institute, Albuquerque, NM, United States, [2]Neurosciences, The University of New Mexico, Albuquerque, NM, United States, [3]National Center for Research Resources, Bethesda, MD, United States

Introduction:

The FIRST-BIRN (FBIRN) project is composed of a team of 11 universities studying brain dysfunctions related to the progression and treatment of schizophrenia. Although brain imaging techniques have generated remarkable progress in understanding how mental and neurological disorders develop, it has been nearly impossible for one laboratory to share and compare data with other labs. One goal of the FBIRN project is to develop procedures to enhance comparability of fMRI results across 11 sites, using different scanners at different field strengths. As a first step, the FBIRN group conducted a "Human Phantom Study", in which 5 subjects were scanned at 10 sites. A number of different paradigms were performed, including a sensorimotor paradigm, a resting state paradigm, breath-hold paradigm and a working memory paradign (Sternberg Item Recognition Paradigm, SIRP).

One interpretation of "comparability of fMRI results" is "comparability of activation pattern". We have developed an approach to enhancing "comparability of activation patterns" by adjusting the statistical threshold at each site. Although we are adjusting a statistical threshold, we are using the summary statistic (in this case Pearson r) simply as an index of site sensitivity. We contend that systematic site differences in this index of sensitivity can be partially explained by hardware and software factors at each site. Once these site factors are known, adjustments to data collection and analysis can be made to enhance comparability of fMRI results. We report here preliminary data from 4 sites (New Mexico – 1.5T, Iowa – 1.5T, Minnesota-3T and MGH-3T) which suggest that there are systematic site differences in this measure of sensitivity.

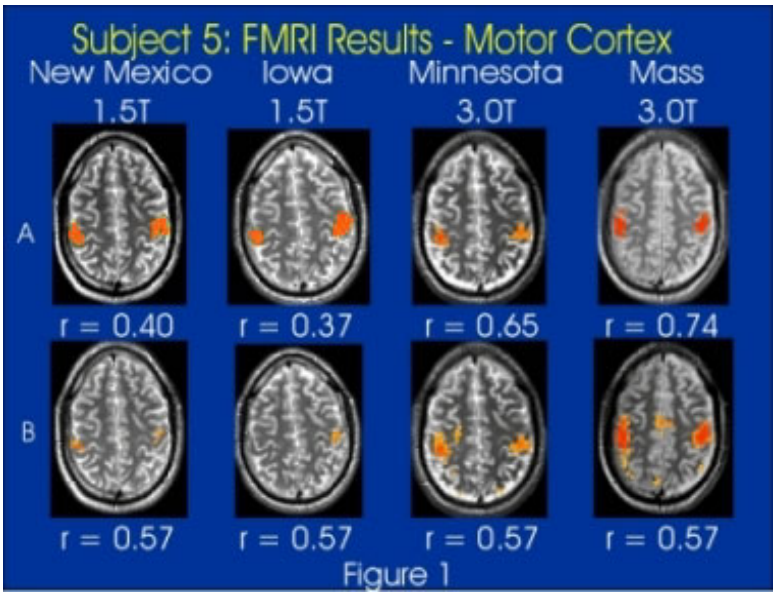


Figure 1

Methods:

Five "Human Phantoms" traveled to 10 sites and had identical fMRI studies performed twice (Visit 1 and Visit 2). The sensorimotor paradigm was designed to robustly activate primary motor cortex, primary auditory cortex and primary visual cortex. It was designed by Dr. Gary Glover, and consists of eight 30 second blocks. Each block consisted of an OFF period (fixation cross, silence, no motion, 15 sec) and an ON period. In the ON period, there was an alternating (3Hz) high contrast black and white checkerboard, a series of auditorily annoying tones which also changed at 3 Hz, and the subject was required to perform bilateral finger tapping (to a standardized response box) at 3 Hz.

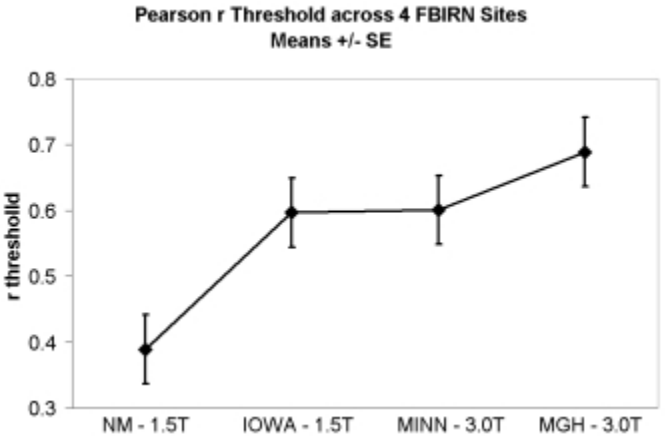
All the sites used a 3000 msec TR with 35 axial slices. The TE for 1.5T was 40 msec and for 3T was 30 msec. The voxel size was 3.44 X 3.44 X 4.00 mm. All scans were overlayed on T2-weighted images with the same slice thickness and slice positioning.

The fMRI analysis was performed with AFNI, and included slice-time correction, motion correction, detrending with a Fourier High-Pass filter, and smoothing with a Gaussian kernel (5 mm FWHM). The square wave of OFF/ON blocks was convolved with hemodynamic response functions (HRF) with several different properties and lags. AFNI chooses the function with the highest correlation, on a voxel by voxel basis, as the result.

We adjusted the correlation threshold (Pearson r) for each study from 4 sites to enhance comparability of activation patterns of the motor cortex ROI and the auditory cortex ROI.

Results:

Adjusting the r-threshold individually for each ROI resulted in much more comparable activation patterns across the four sites (Figure 1A – data from a single subject) than employing an average uniform threshold (Figure 1B). There were systematic differences across the four sites in this threshold (Figure 2). These differences were statistically significant ($F=18.5$, $p < 0.0001$). The NM site is significantly lower than all the other sites ($p<0.0001$). The Iowa site and the MINN site were not significantly different, but both were below the MGH site in sensitivity ($p<0.02$).



Discussion:

A key element explaining site to site differences in fMRI activation patterns is the relative sensitivity of different scanners, even within a field-strength. We are not using this threshold to determine the statistical significance of an activated voxel, but simply as a measure of the sensitivity of scanners and sequences to the BOLD effect. The fact that there are significant site differences in these thresholds indicates that this is a reasonable approach. Our next step is to explain these site sensitivity differences in terms of image properties (e.g., smoothness, T2* weighting), scanner differences (e.g., field strength, gradient characteristics, stability), and sequence differences (e.g., epi vs spiral). We hope to develop a sophisticated algorithm for multi-site calibration that will enhance the probability that different sites will produce similar activation patterns.