How to do a Functional Multi-Center Neuroimaging Study

Bryon Mueller, Ph.D.
Department of Psychiatry, University of Minnesota, Twin Cities

Introduction

As multi-center functional magnetic resonance imaging (fMRI) studies in both healthy and clinical populations become more common, questions about best practices are increasing. A multi-center fMRI study offers several advantaged over a single center study, most notably the potential to increase the number of subjects enrolled, to increase the demographic diversity of the subject population(s), and to include significant numbers of subjects from rare subgroups within clinical populations. While offering many potential advantages, multi-center fMRI studies also present multiple challenges. FMRI data is profoundly affected by experimental and methodological factors. Without careful planning and coordination across the participating centers there is a high likelihood of introducing undesirable inter-site variability into the data, reducing the benefit of the multi-center design. This talk presents results and recommendations from previous studies on how to design, implement and analyze a multi-center fMRI study in a way that improves data quality and reduce inter-site variability.

Some Previous Multi-center fMRI Studies

The first multi-center fMRI studies were published in 1998. In Casey et al.[1], 5-8 healthy subjects were scanned on 1.5T systems at 4 institutions while performed a working memory task. In the same year Ojemann et al. [2] published results from a two site study that looked at the reliability and comparability of fMRI and PET. Of note is that although different scanner vendors and fMRI pulse sequences were used within both studies, good reproducibility was reported. Recently, results from other multi-center fMRI studies have been published, including studies where the same healthy adults were scanned at multiple sites [3-7] as well as clinical studies [8,9] that pool data from multiple sites into a single analysis. The description and discussion found in these works provide an excellent background for anyone planning a multi-center study. In additional, resources and recommendations are available from the Functional Biomedical Information Research Network (FBIRN, http://www.nbirn.net/research/function/index.shtm), which provides best practices recommendations, tools developed to make multi-center fMRI studies easier, as well as actual data from several multi-center fMRI studies.

Potential Sources of Site Variation

There are many MR and non-MR related factors that can reduce over all data quality and/or produce site variability in an fMRI experiment. MR related factors include hardware differences (field strength, scanner manufacturer & model, head coil characteristics, field homogeneity), pulse sequence decisions (k-space trajectory, spin preparation), pulse sequence parameters (TR/TE/flip angle...), reconstruction algorithm and image smoothing, scanner signal-to-noise ratio (SNR), and scanner defects and instability. Potential non-MR factors include variability in stimulus (auditory, visual, tactile) and recording (button box, eye tracker, physiological monitoring) equipment, head restraint and positioning, acquisition positioning, subject training, and many other uncontrolled factors.

While multi-center studies utilizing identical hardware and software will have the lowest potential for site variability, real world factors may require the inclusion of scanners with different software versions, hardware, vendors, or even field strengths. Standardizing sequence parameters across site to the extent possible given site heterogeneity will help reduce the likelihood of site variability in the resulting fMRI data, although sequence standardization may force the study to use sequence parameters that have the "lowest common denominator" to allow the study to be performed at all sites.

Site Standardization and Certification

The most effective way to reduce site variability and insure image quality is through site standardization and certification. Site standardization includes selecting a standard fMRI protocol with sequence parameters that are matched to the extent possible, training local staff to insure uniform data acquisition occurs across site, and monitoring scanner and study data quality during collection to insure each data set achieves minimum data quality. Site certification involves verifying that each site is able to acquire data correctly with acceptable quality before study data collection begins. As a first step, a QA program should be implemented on each scanner. FMRI is stressful to the scanner equipment, and good scanner performance is necessary to obtain quality data. Poorly performing scanner will produce suboptimal data and will very likely introduce site variance. Several examples of scanner and image quality programs can be found in the literature [7,10-13]. After the scanner QA results have shown that the scanner is functioning well, the site should run a test subject through the full study protocol and those data should be evaluated for experimental problems. These test scans are a useful way to verify that the image acquisition (scan parameters, positioning) is correct, to look for image artifacts (ghosting, noise lines), to check the image SNR and smoothness, and to verify proper functioning of the ancillary equipment (e.g. does the software

register the button box presses, is the projector bright enough, etc.).). Minor issues (i.e. protocol parameter errors) can be resolved through feedback to the specific site while major issues (i.e. scanner wasn't able to run the complete protocol, poor SNR) may require that remedial action is taken (the protocol changed, scanner serviced, etc). Once the test data passes inspection, it can be helpful to send an experience MR scientist to every site to perform a final site evaluation as an experienced eye is often helpful in finding subtle site differences that can lead to variability in an fMRI study. Finally, when feasible, sending a group of test subjects to each site to perform the complete protocol will provide critical information on site variability from actual human data.

Data Monitoring

Once the study starts, the data should be processed, analyzed, and reviewed by experienced personal as soon as possible to evaluate the data for potential problems with the scanner, subject (i.e. motion) or data collection. Errors in protocol parameters and scanner induced artifacts may be discovered using automated processing, while more subtle problems such as a decrease in signal to noise ratio or lack of observed activation within expected brain regions might indicate significant problems with the data acquisition, subject issues (motion, lack of attention), or scanner issues. Visual inspection of the results from each run remains an important final QA step. Behavioral data results should also be inspected to insure proper functioning of the equipment. And a simple way to avoid potential left-right orientation confusion in the analysis is by placing an MRI-visible marker near the right side of the head for every subject. The technologist should verify that the marker is visible on the correct side of the image (beware image wrap)!

Data Analysis

The analysis approach taken for multi-center imaging data will depend on the size of the task activation and the amount of residual site effect observed in the data. If the between-site reliability is known to be high (say 0.75 or higher), it may be acceptable ignore site in the analysis, simply merging the data into a single pool and performing the analysis. If residual or unknown site variation remains after site standardization, there are several strategies for performing the group analysis that calibrate site variance out of the data. In a multi-center study that used multiple vendors, field strengths and k-space methods, FBIRN showed that a dramatic reduction in site variance component could be achieved by covarying for scanner differences in signal-to-fluctuation-noise-ratio and smoothness [14,15]. Analysis can also be explicitly modeled as a site and site-by-group effect by incorporating site as an additive effect in an analysis of variance. An analysis of covariance can also be used to calibrate out site effects, for example using a breath-hold task and/or perfusion measurement as a covariate to adjust for scanner, site, or uninteresting subject differences in the analysis of a cognitive task. Finally, analyses can be performed that treat each site separately, and then integrate these results into a meta-analysis. This approach makes no assumption about the nature of site effects.

Conclusions

Multi-center fMRI studies have become a method of choice for investigating cognitive function in both healthy and clinical populations, allowing for unprecedented populations sizes. Best practice recommendations are still evolving because the technical methods are improving, changes are occurring in the hardware and software, and basic understanding of the underlying cognitive neuroscience is increasing. Methods used to perform multi-center fMRI studies are in flux because neuroimaging is a developing science, but current research has produced good recommendations and resources that are now available to help those designing and executing such studies.

References

1. Casey BJ et al. Reproducibility of fMRI results across four institutions using a spatial working memory task. Neuroimage. 1998 Oct;8(3):249-61.
2. Ojemann JG et al. Functional MRI studies of word-stem completion: reliability across laboratories and comparison to blood flow imaging with PET. Hum Brain Mapp. 1998;6(4):203-15.
3. Zou KH et al. Reproducibility of functional MR imaging: preliminary results of prospective multi-institutional study performed by Biomedical Informatics Research Network. Radiology. 2005 Dec;237(3):781-9.
4. Costafreda SG et al. Multisite fMRI reproducibility of a motor task using identical MR systems. J Magn Reson Imaging. 2007 Oct;26(4):1122-6.
5. Friedman L et al. Test-retest and between-site reliability in a multicenter fMRI study. Hum Brain Mapp. 2008 Aug;29(8):958-72.
6. Suckling J et al. Components of variance in a multicentre functional MRI study and implications for calculation of statistical power. Hum Brain Mapp. 2008 Oct;29(10):1111-22.
7. Sutton BP et al. Investigation and validation of intersite fMRI studies using the same imaging hardware. J Magn Reson Imaging. 2008 Jul;28(1):21-8.
8. Schneider F et al. Neural correlates of working memory dysfunction in first-episode schizophrenia patients: an fMRI multi-center study. Schizophr Res. 2007 Jan;89(1-3):198-210.

9. Bosnell R et al. Reproducibility of fMRI in the clinical setting: implications for trial designs. Neuroimage. 2008 Aug 15;42(2):603-10.
10. Stocker T et al. Automated quality assurance routines for fMRI data applied to a multicenter study. Hum Brain Mapp. 2005 Jun;25(2):237-46.
11. Friedman L et al. Report on a multicenter fMRI quality assurance protocol. J Magn Reson Imaging. 2006 Jun;23(6):827-39.
12. Olsrud J et al. A two-compartment gel phantom for optimization and quality assurance in clinical BOLD fMRI. Magn Reson Imaging. 2008 Feb;26(2):279-86.
13. Gedamu EL et al. Automated quality control of brain MR images. 1: J Magn Reson Imaging. 2008 Aug;28(2):308-19.
14. Friedman L et al. Reducing inter-scanner variability of activation in a multicenter fMRI study: role of smoothness equalization. Neuroimage. 2006 Oct 1;32(4):1656-68.
15. Friedman L et al. Reducing interscanner variability of activation in a multicenter fMRI study: controlling for signal-to-fluctuation-noise-ratio (SFNR) differences. Neuroimage. 2006 Nov 1;33(2):471-81.