RESIDENT & FELLOW SECTION

Section Editor
Mitchell S.V. Elkind, MD, MS

Pearls & Oy-sters: Functional MRI

A primer for neurology residents

Peter Pressman, MD
Darren Gitelman, MD

Correspondence & reprint requests to Dr. Pressman: p-pressman@northwestern.edu

PEARL

- fMRI is becoming increasingly prevalent not just as a powerful research tool, but also in clinical settings such as presurgical localization of language function.

OY-STER

- Proper collection and interpretation of fMRI data is a multistep process with many potential opportunities for error. If fMRI data are used clinically, the clinician must question whether data collected in a population of carefully selected research subjects can be reasonably applied to a particular patient, or if unique patient characteristics could render such applications unreliable.

Since the first description of blood oxygen level–dependent (BOLD) contrast in animals and subsequently in humans,[1] fMRI has become an invaluable tool in neurologic research, and is also increasingly involved in presurgical assessment of brain function. For example, fMRI may complement other techniques in localizing language function, thereby permitting surgeons to operate in a nearby region while ensuring the patient's use of language remains intact.[2] Despite fMRI's growing utility, residents often lack sufficient comprehension of the technique to allow for critical appraisal of fMRI research. The goal of this article is to provide the neurology resident with a rudimentary understanding of the most common fMRI methods, along with potential sources of error in interpretation of results (table).

As the acronym BOLD suggests, fMRI does not measure neural activity directly, but relies on the phenomenon of neurovascular coupling to indicate changes in brain activity. Shortly following neuronal activation, oxyhemoglobin is converted to deoxyhemoglobin. Deoxyhemoglobin is paramagnetic and has a T2 shortening effect—that is, it lowers signal intensity on a T2-weighted scan.[3] About 1 to 5 seconds after the deoxyhemoglobin level rises, the brain's vasculature appears to overcompensate with a local increase in cerebral perfusion out of proportion to the deoxyhemoglobin level. As a result, levels of diamagnetic oxyhemoglobin actually become higher than at baseline. This "paradoxical" increase in the oxy- to deoxyhemoglobin ratio causes an increased MRI signal intensity. Studies have shown that these hemodynamic-related changes in magnetic resonance signal intensity are tightly coupled to the actual region of neuronal activation.[1,4]

The most common use of fMRI is to correlate brain activation (as implied by the hemodynamic response) with a specific task or stimulus. Benefits of fMRI include the ability to measure brain signals noninvasively, without radiation exposure. The technique has a relatively high spatial resolution (usual voxel sizes of $3 \times 3 \times 3$ mm^3) and is capable of recording signal throughout the brain. However, fMRI has lower temporal resolution when compared to EEG (several seconds vs 1 msec). Moreover, fMRI relies on a complex process, with several potential sources of error. These include loss of normal neurovascular coupling, difficulties in designing tasks that patients can perform, technical limitations of the MRI protocol, or inappropriate data analysis.[5]

DEPENDENCE ON NEUROVASCULAR COUPLING

fMRI depends on a consistently coupled relationship between neural activity and a hemodynamic response. However, neurovascular coupling is variable in certain patient populations. Neurovascular coupling may be altered in hypoxia, hypercapnia, hypertension, and even as part of normal aging.[6] Disease states such as vasculitis, angiopathies, tumors, and vascular malformations may also disrupt a "normal" neurovascular response.[2] While it may be impossible for a researcher to measure and confirm normal neurovascular coupling, the researcher's results may be called into question if there is reason to suspect altered coupling is present.

From the Department of Neurology (P.P.) and Cognitive Neurology and Alzheimer's Disease Center (D.G.), Northwestern University Medical School, Chicago, IL.

Disclosure: Author disclosures are provided at the end of the article.

Table Questions in reading an fMRI publication

Potential source of error	Questions to consider
Neurovascular coupling	Is there a reason to suspect neurovascular coupling is compromised, e.g., hypoxemia, hypercapnia, or uncontrolled hypertension?
Task design	Does the task design allow the examination of interactions between different factors or is an aspect of the task varied to allow parametric analyses?
	Are tasks of appropriate difficulty?
	If patient and control groups are being studied, have they been trained similarly and can they perform the task at an appropriate accuracy?
Technical considerations	Does the study focus on a region likely to be affected by susceptibility artifact, e.g., the amygdala, basal temporal region, or orbitofrontal cortex? If so, has this been examined and what adjustments have been made?
Data analysis	Is a general linear model used or if there is a small number of subjects (<10) have nonparametric statistics been used?
	How are the statistics calculated, and are they corrected for multiple comparisons?
	How are activations selected?
	Are the investigators focusing on signal intensity, number of pixels activated, or a different measure?
	If results are meant to apply to a general population, has a random (or mixed) effects model been used for data analysis?
Clinical application	How easily could these data be applied to an individual patient?

TASK DESIGN Unlike classic lesion-based localization, which identifies brain areas essential for performing a task, fMRI reveals all regions demonstrating task-related changes in brain activity. This includes areas that are essential for the task, and also regions with correlated or supplementary activity that may be nonessential for that task. For example, when subjects are asked to generate a list of words, areas related to both language and attention are activated, and this could lead to erroneous conclusions about the brain regions necessary for language.[2]

Early functional imaging studies were designed in light of the concept of "pure insertion," which held that adding components to a task resulted in incremental recruitment of brain regions.[7] Proper interpretation of functional imaging studies was thought to depend on the researcher's ability to subtract a resting or control condition from images acquired on task. It is now understood that the correspondence between tasks and brain activity is more complex, and designing tasks based on pure insertion can lead to erroneous conclusions about the association between brain areas and cognitive functions—for example, that Wernicke area is not involved in language processing.[8] More sophisticated parametric designs (which vary task complexity) and factorial designs (which allow an examination of interactions) can now be used to investigate the correspondence between brain activity and task performance.[9]

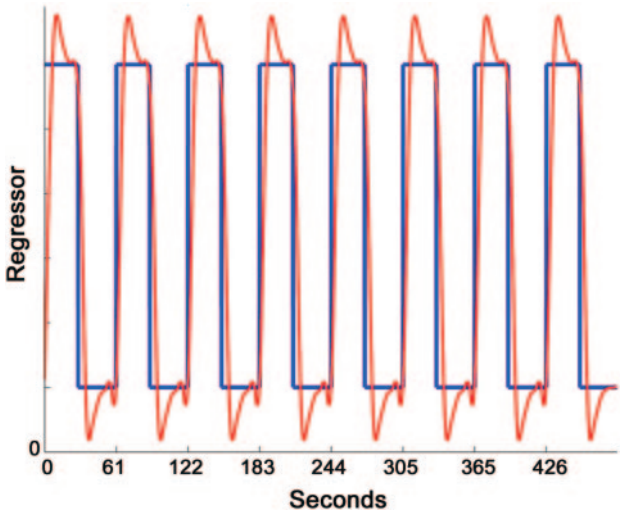
Another important aspect of task design is ensuring the task is tailored to a patient's abilities. If a task is too easy, the patient may need to exert little effort and the corresponding brain activation may be minimal. If a task is too difficult, patients may be unable to perform it, or additional brain regions may be recruited for task performance, making interpretation of the activation patterns more challenging.[6] Furthermore, while a patient may need to practice a task to perform it properly, overfamiliarity with the task may change activation patterns as the patient develops different performance strategies. Therefore, when interpreting studies it is important to consider whether subject groups were trained equally and performed appropriately on the task.[5]

TECHNICAL CONSIDERATIONS Most fMRI studies are done at magnetic field strengths of 1.5 Tesla, and increasingly 3 Tesla. Higher field strengths can improve BOLD signal detection and spatial resolution. This helps to decrease type II errors (false-negative results) and increases statistical power. However, increasing magnetic field strength also increases the possibility of susceptibility artifacts, particularly in the orbitofrontal cortex and anterior inferior temporal lobes.[10] Several techniques may reduce these artifacts: for example, the frequency matrix may be increased, voxel size reduced, or slice thickness reduced. Each of these techniques has its benefits and drawbacks, which are beyond the scope of this article. Other artifacts include artifacts of resolution limitations and data filtering, field inhomogeneities, and movement.[5]

DATA PROCESSING AND ANALYSIS Analysis usually begins with preprocessing the images. This involves realigning the time series of images to minimize the effects of motion, correcting the data for differences in the time at which each of the image slices are acquired (slice timing correction), normalizing the images to a standard template, if multiple subjects are to be combined, and smoothing the images to further minimize noise and improve signal detection.[11,12]

The next step is the "first-level" or "single-subject" analysis. One of the most common approaches is to use the general linear model to design a multiple linear regression. To do this the investigator constructs a model of the predicted BOLD signal by designating which time periods (blocks) or trials (events) in a task will be associated with changes in brain activity. This prediction includes a time lag for the hemodynamic response, which underlies the BOLD signal measurement. A regression analysis then calculates the parameter estimate or "goodness of fit" between the predicted model and the fMRI BOLD signal at each of the voxels in the brain (figure 1). By comparing this measure to the SD of the data,

Figure 1 A graph of a "boxcar" function



The boxcar represents a neuronal response to a stimulus. This response is assumed to start and stop in a manner exactly timed with the stimulus. Overlapping the boxcar is a convolved predicted hemodynamic response function. This function is the model to which the measured blood oxygenation level–dependent signal will be compared for statistical significance.

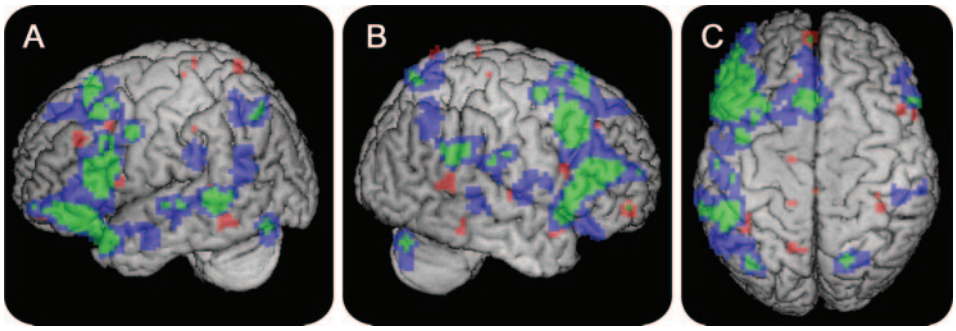
a *t* score can be calculated at an individual subject level.[11] More commonly, however, the parameter estimates from individual subjects are combined in a "second level" or random-effects analysis. The use of a random-effects model, which accounts for intersubject variance, ensures the results are likely to be applicable to the general population and not just that particular group of subjects.[12]

The investigator must next determine which voxels show statistically significant activations. Because of the large number of voxels in an average fMRI scan (upwards of 50,000), conventional criteria for significance ($p < 0.05$) at each individual voxel could show a high number of false-positives (e.g., 2,500). This is known as the multiple comparisons problem. This problem may be addressed by adjusting the statistical thresholds to ensure no more than 5% false-positives across the entire brain based on the significance level at either individual voxels or clusters of voxels. A voxel-level analysis is better at identifying the precise location of focal brain activation, but may have less sensitivity for detecting an activation because the statistical threshold is quite stringent with so many voxels in the brain. In contrast, cluster-level analyses are less localizing, but more powerful for spatially extended signals. Which method is used will depend on the goals of the experiment, the strength and location of the activation, and the number of subjects in the study.

Multiple comparisons may be corrected for using a variety of methods, including a Bonferroni correction (usually considered overconservative for functional imaging data since the voxels are not all independent), familywise error rate (FWER), which ensures no more than 5% false-positives in every analysis, and false discovery rate (FDR), which ensures no more than 5% false-positives on average. Each method has strengths and weaknesses—for example, use of FWER at voxel levels may be more specific, whereas FDR may be more sensitive (figure 2).[12]

CONCLUSION fMRI has produced robust and replicable data across various studies, but the critical reader must remember that these compelling images are in fact statistically analyzed representations of cerebral blood flow, and represent only an indirect correlate of neural activity and brain function. Brain processes are usually more complex and nonlocalized than these colored blobs suggest. The statistical methods used may present false-positives or false-

Figure 2 Functional images obtained from a research volunteer asked to perform a verbal recognition task



Views are from left (A), right (B), and superior (C) aspects. Different methods of multivariate correction are used, with the results superimposed. In red, activations are set at an uncorrected *p* level of 0.001. Blue demonstrates the overlap of activations with an false discovery rate (FDR) threshold of 3, with a cluster level of 47 (representing a corrected FDR level). In green are activations at a familywise error of 0.05, with a cluster of 0. Note how sensitivity and specificity differ between these methods.

negatives. With these basic precautions, the neurologist may appreciate these studies with a discerning eye, and better understand the underlying methodology and clinical applications.

AUTHOR CONTRIBUTIONS

Dr. Pressman: concept and design, drafting/revising the manuscript. Dr. Gitelman: drafting/revising the manuscript, figure design.

DISCLOSURE

Dr. Pressman is a member of the *Neurology*® Resident & Fellow Section editorial team. Dr. Gitelman serves on editorial advisory boards for *Neuroimage* and *Frontiers in Integrative Neuroscience*; serves as a consultant for Abiant, Inc.; receives research support from the NIH (NIDDK, NIAMS, NICHD), Yale University, the Michael J Fox Foundation, and the Paul Ruby Foundation; and has served as an expert in medico-legal cases.

REFERENCES

1. Kwong K, Belliveau J, Chesler D, et al. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. Proc Natl Acad Sci USA 1992;89:5675–5679.
2. Bookheimer S. Pre-surgical language mapping with functional magnetic resonance imaging. Neuropsychol Rev 2007;17:145–155.
3. Thulborn KR, Waterton JC, Matthews PM, Radda GK. Oxygenation dependence of the transverse relaxation time of water protons in whole blood at high field. Biochim Biophys Acta 1982;714:265–270.
4. Logothetis N, Pauls J, Augath M, Trinath T, Oeltermann A. Neurophysiological investigation of the basis of the fMRI signal. Nature 2001;412:128–130.
5. Faro SH, Mohamed FB, Haughton V. Functional MRI: Basic Principles and Clinical Applications. New York: Oxford University Press; 2006.
6. Krings T, Reinges M, Willmes K, et al. Factors related to the magnitude of T2* MR signal changes during functional imaging. Neuroradiology 2002;44:459–466.
7. Friston K, Price C, Fletcher P, Moore C, Frackowiak R, Dolan R. The trouble with cognitive subtraction. NeuroImage 1996;4:97–104.
8. Petersen S, Fox P, Posner M, Mintun M, Raichle M. Positron emission tomographic studies of the cortical anatomy of single-word processing. . Nature 1988;331:585–589.
9. Amaro E, Barker G. Study design in fMRI: basic principles. Brain Cogn 2005;60:220–232.
10. Devlin JT, Russell RP, Davis MH, et al. Susceptibility-induced loss of signal: comparing PET and fMRI on a semantic task. NeuroImage 2000;11:589–600.
11. Smith SM. Overview of fMRI analysis. Br J Radiol 2004; special issue:S167–S175.
12. Carter CS, Heckers S, Nichols T, Pine DS, Strother S. Optimizing the design and analysis of clinical functional magnetic resonance imaging research studies. Biol Psychiatry 2008;64:842–849.

Pearls & Oy-sters: Functional MRI: A primer for neurology residents

Peter Pressman and Darren Gitelman

Neurology 2012;78;e68-e71
DOI 10.1212/WNL.0b013e318248e57a