

A survey of patient motion in disorders of consciousness and optimization of its retrospective correction

Malte Hoffmann¹, T. Adrian Carpenter¹, Guy B. Williams¹, Stephen J. Sawiak^{1,2}

¹Wolfson Brain Imaging Centre, University of Cambridge, Cambridge, United Kingdom

²Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge, United Kingdom

Abstract

Functional magnetic resonance imaging (fMRI) can be seriously impaired by patient motion. The purpose of this study was to characterize the typical motion in a clinical population of patients in disorders of consciousness and compare the performance of retrospective correction with rigid-body realignment as implemented in widely used software packages. 63 subjects were scanned with an fMRI visual checkerboard paradigm using a 3-T scanner. Time series were corrected for motion, and the resulting transformations were used to calculate a motion score. SPM, FSL, AFNI and AIR were evaluated by comparing the motion obtained by re-running the tool on the corrected data. A publicly available sample fMRI dataset was modified with the motion detected in each patient with each tool. The performance of each tool was measured by comparing the number of supra-threshold voxels after standard fMRI analysis, both in the sample dataset and in simulated fMRI data. We assessed the effect of user-changeable parameters on motion correction in SPM. We found the equivalent motion in the patient population to be 1.4 mm on average. There was no significant difference in performance between SPM, FSL and AFNI. AIR was considerably worse, and took more time to run. We found that in SPM the quality factor and interpolation method have no effect on the cluster size, while higher separation and smoothing reduce it. We showed that the main packages SPM, FSL and AFNI are equally suitable for retrospective motion correction of fMRI time series. We show that typically only 80% of activated voxels are recovered by retrospective motion correction.

Keywords: retrospective motion correction, motion in disorders of consciousness, functional magnetic resonance imaging (fMRI), rigid-body registration.

1 Introduction

Functional magnetic resonance imaging (fMRI) is a powerful imaging modality in human brain research, providing spatial information on task-related changes in brain activity. Oxygen demand from active neurons is higher than that for inactive neurons, allowing the level of oxygenated blood to act as a marker for neural activity. The relationship between neural activity and blood oxygenation is known as the haemodynamic response. As haemoglobin has different magnetic susceptibilities in oxygenated and de-oxygenated blood [1], these neurons can be distinguished in blood oxygen level dependent (BOLD) contrast imaging.

Despite its success, fMRI analysis still poses a number of issues. One of the most problematic of these is gross patient motion, which can reduce the statistical power in fMRI studies by confounding voxel time courses. The presence of severe motion artefacts can significantly compromise image quality in research populations. Increased examination times cause considerable financial costs, and in some cases it is simply not possible to obtain images of acceptable quality. The problem of motion has been addressed with prospective and retrospective correction [1–8].

The most commonly used procedure for dealing with motion in an fMRI time series is retrospective correction: the images are realigned with a reference image, assuming that any head movement can be characterised by a translation in the three directions of space and a rotation about three axes. Most fMRI analysis pipelines use such an algorithm in preprocessing before statistical analysis. Some software packages include higher-degree-of-freedom (DOF) correction algorithms; compared to 6-DOF realignment, 12-DOF registration also allows for considering scale and shear [8].

Several software approaches are freely available for retrospective motion correction. Taking 100 randomly chosen fMRI studies published in 2013 from the abstract database Web of Science (Thomson-Reuters), 55% of studies used SPM [9], 15% used FSL [10] and 14% used AFNI [11]. Although the current SPM version is SPM8, 10% of the SPM users employed SPM2 and 25% used SPM5.

A comparison of motion correction packages was performed previously in healthy volunteers by Oakes et al. [12]. In addition to reporting the distribution of motion effects seen in a patient population, our work builds on that study by exploring newer editions of the same packages.

In the present study, we compared the motion levels detected from time series EPI data collected in fMRI experiments in patients with varying degrees of disorders of consciousness. We assessed the performance of the leading packages used for retrospective correction including SPM (version 8, April 2009), FSL (version 5, September 2012) and AFNI (version 2011-12-21) in three distinct ways. We compared these packages in addition to the oldest tool available to us, AIR (version 5.3, March 2011) [13]. Firstly, we compared the extent of motion correction transformations to an initial pass with each software package and a subsequent pass upon the corrected data with the same package.

Secondly, we simulated an fMRI dataset which was then corrupted with the motion from each patient and subsequently corrected with each software package. The detected activation was then compared to the simulated activation pattern without motion across each package. Finally, we took some real fMRI data to repeat this experiment and compared the largest detected cluster size across packages.

We explored the effect of altering user-changeable parameters in the most widely used package, SPM and recommend optimal settings.

2 Methods

2.1 Patient population and image acquisition

Patients with various degrees of disorders of consciousness ($n = 63$) were scanned using a Siemens Trio system at 3 T with an fMRI visual checkerboard paradigm for 160 frames (EPI: TR/TE 2000/30 ms, matrix 64×64 , FOV 192×192 mm², 32 slices of 3-mm thickness with 0.75 mm gap). A total of 121 series were produced for analysis.

All patients gave informed consent and the study was approved by the Local Research Ethics committee (LREC 99/391).

2.2 Survey of motion

Time series were corrected for motion with each package and the resulting transformations were used to calculate a motion score. Following Ref. [3], this score was defined by taking the RMS translation of temporal frames with rotational components included as the equivalent translation of points on a 64-mm sphere.

2.3 Comparison of packages

For a consistent comparison between patients, a publicly available sample fMRI dataset (SPM auditory data, [14]) was artificially modified with the motion detected in each patient with each tool. Each package was evaluated by comparing the motion score obtained by re-running the tool on the corrected data.

The performance of each tool was measured by comparing the number of supra-threshold voxels in the largest significant cluster after standard fMRI analysis with SPM8 ($p_{\text{FWE}} < 0.05$ in the dataset without motion). The fMRI analysis was done following Ref. [14]. All packages were run with default parameters (see appendix). However, we reduced the default number of 12 DOF to 6 in FSL for consistency with the other packages. We assessed the effect of user-changeable parameters on motion correction performance in SPM (separation, smoothing FWHM, quality factor and interpolation method) on the motion trajectory from each time series.

To evaluate the effect of motion on a known area of activation, we simulated a task-related BOLD signal according to Ref. [15] in a cube with a side length of 6 voxels. Each package was evaluated by comparing the number of activated voxels ($p_{\text{FWE}} < 0.05$) inside the cube after fMRI analysis.

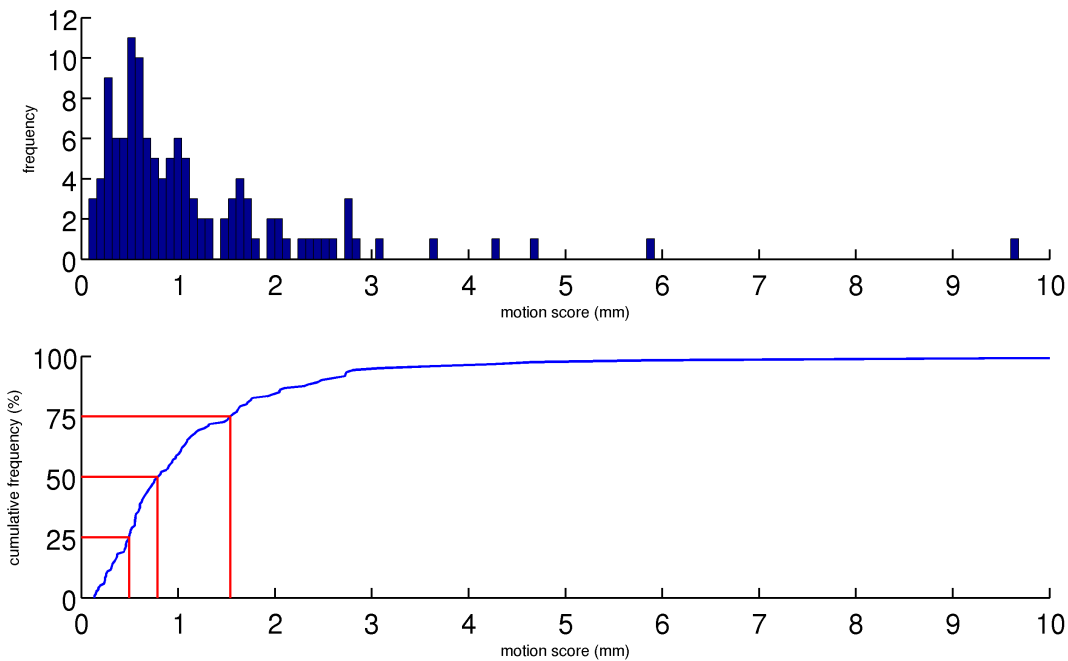


Figure 1. Distribution of motion estimates and its cumulative density function (%) for 121 fMRI time series obtained using SPM. Motion scores (mm) are plotted on the x-axis. One time series with an extreme motion score of 22.39 mm was omitted from the figure.

3 Results

Motion scores and the cumulative density function for motion in all patients are shown in [Figure 1](#). There are time series with extreme motion. However, 85% of the times series have mean translation amplitudes under 1 mm, and 88% of them have mean rotation angles under 1° . The mean score was 1.4 mm during the time course, though 84% of patients showed motion less than 2 mm. Mean translation amplitudes, rotations and motion scores are given in [Table 1](#). There was good agreement between the scores as measured by each package (see [Figure 2](#)). The time taken to correct one time series is shown in [Figure 3](#).

Residual motion scores on the second pass of each tool are shown in [Figure 4](#). Numbers represent the mean ratio of second to first-pass motion scores. On the first pass, SPM, FSL, AFNI and AIR detect 1.33, 1.25, 1.16 and 2.06 mm of average equivalent distance in data with motion estimated by SPM. SPM estimates the least motion in the second pass, followed by FSL, AFNI and AIR. The residual motion detected by AIR is comparable in magnitude to the amount of motion induced initially.

[Figure 5](#) compares the mean cluster size obtained following motion correction by each package of data with motion detected by SPM. SPM, FSL and AFNI match the data without motion equally well, followed by AIR. For comparison, the interpolation error was measured by displacing the sample dataset with no

Table 1. Mean translation amplitude, rotation and motion score in the clinical population estimated for rigid-body realignment by the different software packages.

	T_x (mm)	T_y (mm)	T_z (mm)	R_x ($^\circ$)	R_y ($^\circ$)	R_z ($^\circ$)	Score (mm)
SPM	0.21	0.39	0.47	0.36	0.27	0.30	1.36
FSL	0.20	0.36	0.44	0.33	0.25	0.28	1.25
AFNI	0.34	0.46	0.49	0.33	0.23	0.30	1.39
AIR	0.18	0.26	0.41	0.29	0.31	0.30	1.18

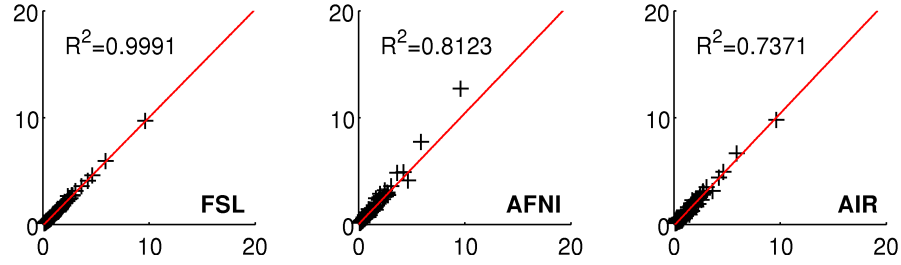


Figure 2. Comparison of motion parameters from different packages, with SPM plotted on the x-axis in each case. Units are equivalent distance (mm).

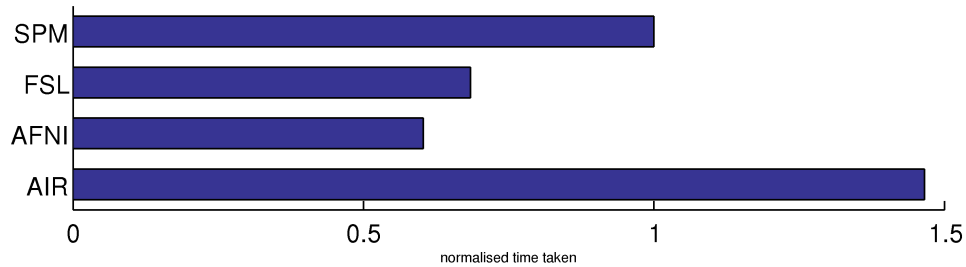


Figure 3. Time taken to correct the motion of one dataset (relative to time taken by SPM, 56 sec).

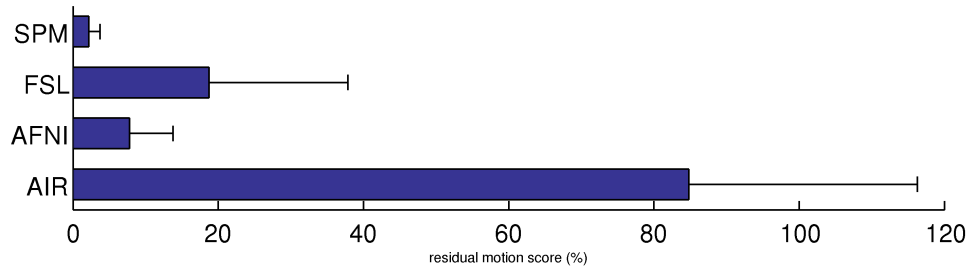


Figure 4. Residual motion (%) detected in a second pass of each tool on the corrected images. Numbers represent the mean ratio of second to first-pass motion in 121 time series. Error bars show standard deviations.

motion by half a voxel in each dimension and this resulted in a cluster size of 93.4% of its original size.

Cluster sizes are represented in [Figure 6](#) for different values of user-changeable parameters in SPM. The sizes are relative to the result obtained with default parameters. The quality factor and interpolation method have no effect on the cluster size, while higher separation and smoothing reduce it.

[Figure 7](#) compares the mean number of activated voxels inside a cube of simulated fMRI data after motion correction with each tool. Uniformly, motion resulted in a reduction in mean cluster size independent of which package was used. The variability was such that that the three main packages SPM, FSL and AFNI do not show a significant difference. The mean activation outside the cube was similar to the result with no motion for all tools (2 voxels).

4 Discussion

The majority of patients showed motion comparable to the voxel size of acquisition, with some severe outliers. We showed that 84% of patients moved less than 2 mm on average. This compares with the maximum motion detected by Ref. [7] in a group of 1600 paediatric patients: 79% had motion under 3 mm (though 41% of patients in this study were under general anaesthesia).

For all packages assessed, the fundamental algorithm is the same: fitting a matrix based on a function of

image similarity. The difference is the optimisation strategy and the starting parameters. SPM [9] and AFNI use least-squares minimisation. For FSL and AIR, the default cost functions are normalised correlation and standard deviation of ratio images, respectively. SPM performs a local search for minima [9], whereas FSL searches globally [10].

There was no significant difference in performance between SPM, FSL and AFNI. AIR was considerably worse, and took more time to run. When testing 12-DOF registration on data with motion detected by SPM, there was no significant change in performance for FSL. 12-DOF correction with AIR even resulted in a decrease of cluster size from 47.6% to 37.8% compared to 6 DOF. However, 12-DOF algorithms took considerably more time to run than their 6-DOF counterparts (FSL: +86.5%, AIR: +35.0%).

We did not find any combination of realignment parameters that increased the size of the fMRI activation cluster in SPM. We therefore recommend choosing parameters that require the shortest time without affecting statistical power: realignment takes 12.9 sec with 8 mm separation, 5 mm FWHM (default), bicubic interpolation and a quality factor of 0.001. This is more than 4 times faster than with the default values, and the cluster size remains unchanged (98.6%).

There was no significant activation outside a cube of simulated fMRI data after motion correction. Therefore none of the tools used for motion correction have introduced artificial activation. However, AIR performed considerably worse than the other packages.

Our overall findings are in agreement with those presented in Ref. [12] in 2005, with the exception of the performance obtained by AIR: the newer versions of the main packages AFNI, SPM and FSL still perform equally well, with AFNI taking the least time to run in both studies. While we obtained a significantly lower number of activated voxels with AIR, that study found it to produce similar results as the other tools. They did not find any combination of motion correction parameters that would improve the performance noticeably either, i.e. using so-called "accurate" parameters did not lead to more activation.

In contrast to Ref. [12], which concerned motion in healthy volunteers, we assessed it in a clinical patient population. For this reason, we had to use a different methodology (a standard fMRI dataset for each comparison) to prevent the confounding effects of possibly strikingly different activation patterns that, moreover, are likely to correlate with the motion detected.

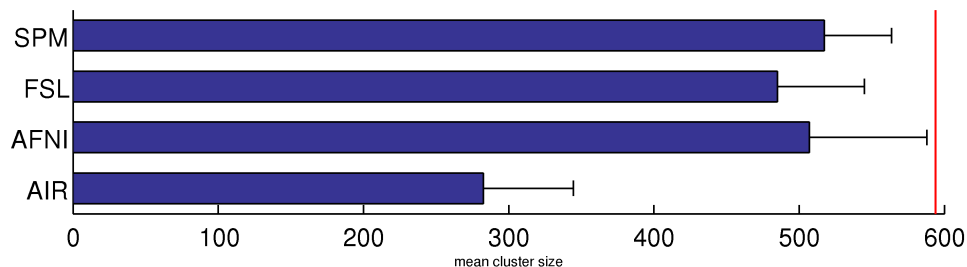


Figure 5. Mean cluster sizes obtained for each software package used for motion correction. A cluster size of 594 is obtained without motion (corresponding to $p_{FWE} < 0.05$). Error bars show standard deviations.

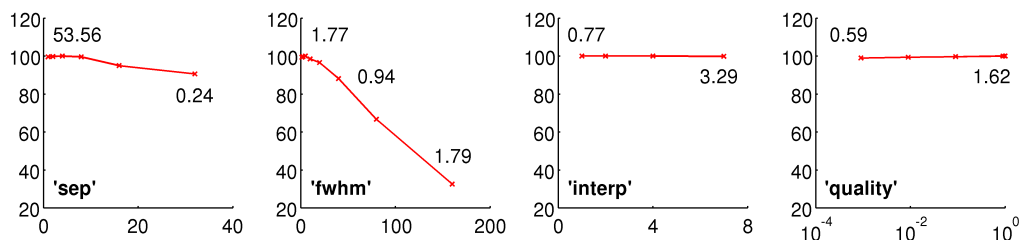


Figure 6. Cluster sizes (relative to result with default parameters, %) for different realignment parameters in SPM. The time taken (relative to time with default parameters) is shown for the first and last values.

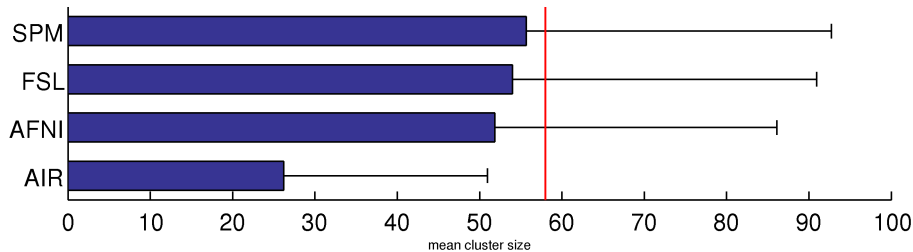


Figure 7. Cluster sizes (corresponding to $p_{FWE} < 0.05$) obtained for each software package in a cubic area of simulated activation of 216 voxels. A cluster size of 57 is obtained without motion. Error bars show standard deviations.

5 Conclusion

We have presented the statistical distribution of motion parameters in a patient population drawn from disorders of consciousness. Our findings suggest that SPM, FSL and AFNI are equally suitable for retrospective motion correction of fMRI time series. Substantial time improvements can be gained by decreasing the quality factor setting in SPM without compromising the final results.

We note that regardless of the package used, even after retrospective motion correction motion reduces sensitivity with only $\sim 80\%$ of activated voxels being detected. It is likely that the actual loss in practice is worse than this, as we only simulated a pure motion effect here by displacing uncorrupted data. Motion during acquisition will result in further artefacts causing fewer activated voxels to be found.

Acknowledgements

The authors would like to thank the Cusanuswerk funding body (Bonn, Germany) for funding, the James S. McDonnell Foundation Collaborative Activity Grant (Recovery of Consciousness After Severe Brain Injury) for funding the image acquisition, and the Cambridge Research into Impaired Consciousness Group for sharing data.

Appendix

The commands used for the different software packages are as follows, where angle brackets need to be replaced with the actual file names.

SPM

```
spm_realign(<4D_IMAGE_FILE>)
```

AFNI

```
3dvolreg -base 0 -prefix NULL -1Dmatrix_save <OUTPUT_FILE> <4D_IMAGE_FILE>
```

FSL with 6 DOF

```
mcflirt -in <INPUT_VOLUME> -out <OUTPUT_FILE> -refvol 0 -mats -dof 6
```

FSL with 12 DOF

```
mcflirt -in <INPUT_VOLUME> -out <OUTPUT_FILE> -refvol 0 -mats -dof 12
```

AIR with 6 DOF

```
alignlinear <REF_VOLUME> <INPUT_VOLUME> <OUTPUT_FILE> -m 6 -t1 10 -t2 10
```

AIR with 12 DOF

```
alignlinear <REF_VOLUME> <INPUT_VOLUME> <OUTPUT_FILE> -m 12 -t1 10 -t2 10
```

References

- [1] Ogawa S, Lee TM, Kay AR, Tank DW. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci USA*, 1990;87(24):9868–72.
- [2] Welch EB, Manduca A, Grimm RC, Ward HA, Jack Jr CR, Clifford R. Spherical navigator echoes for full 3D rigid body motion measurement in MRI. *Magn Reson Med*, 2002;47(1):32–41.
- [3] Tisdall MD, and Hess AT, Reuter M, Meintjes EM, Fischl B, van der Kouwe AJW. Volumetric navigators for prospective motion correction and selective reacquisition in neuroanatomical MRI. *Magn Reson Med*, 2012;68(2):389–99.
- [4] Zaitsev M, Dold C, Sakas G, Hennig J, Speck O. Magnetic resonance imaging of freely moving objects: prospective real-time motion correction using an external optical motion tracking system. *NeuroImage*, 2006;31(3):1038–50.
- [5] Muraskin J, Ooi MB, Goldman RI, Krueger S, Thomas WJ, Sajda P, Brown TR. Prospective active marker motion correction improves statistical power in BOLD fMRI. *NeuroImage*, 2013;68:154–61.
- [6] Sengupta S, Tadanki S, Gore JC, Welch EB. Prospective real-time head motion correction using inductively coupled wireless NMR probes. *Magn Reson Med*, 2014;72(4):971–85.
- [7] Holdsworth SJ, Aksoy M, Newbould RD, Yeom K, Van AT, Ooi MB, Barnes PD, Bammer R, Skare S. Diffusion tensor imaging (DTI) with retrospective motion correction for large-scale pediatric imaging. *J Magn Reson Imaging*, 2012;36(4):961–71.
- [8] Maclaren J, Herbst M, Speck O, Zaitsev M. Prospective motion correction in brain imaging: a review. *Magn Reson Med*, 2013;69(3):621–36.
- [9] Friston KJ, Ashburner J, Frith CD, Poline J-B, Heather JD, Frackowiak RSJ. Spatial registration and normalization of images. *Hum Brain Mapp*. 1995;3(3):165–89.
- [10] Jenkinson M, Smith S. A global optimisation method for robust affine registration of brain images. *Med Image Anal*, 2001;5(2):143–56.
- [11] Cox RW, Jesmanowicz A. Real-time 3D image registration for functional MRI. *Magn Reson Med*, 1999;42(6):1014–8.
- [12] Oakes TR, Johnstone T, Ores Walsh KS, Greischar LL, Alexander AL, Fox AS, Davidson RJ. Comparison of fMRI motion correction software tools. *NeuroImage*, 2005;28(3):529–43.
- [13] Jiang AP, Kennedy DN, Baker JR, Weisskoff R, Tootell RBH, Woods RP, Benson RR, Kwong KK, Brady TJ, Rosen BR, Belliveau JW. Motion detection and correction in functional MR imaging. *Human Brain Mapping*, 1995;3(3):224–35.
- [14] Single subject epoch (block) auditory fMRI activation data [Internet]. The FIL Methods group, Wellcome Trust Centre for Neuroimaging; c1991 [cited 2014 April 10]. Available from: <http://www.fil.ion.ucl.ac.uk/spm/data/auditory>.
- [15] Kwong KK, Belliveau JW, Chesler DA, Goldberg IE, Weisskoff RM, Poncelet BP, Kennedy DN, Hoppel BE, Cohen MS, Turner R. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc Natl Acad Sci USA*, 1992;89(12):5675–9.