Effect of slice orientation on reproducibility of fMRI motor activation at 3 Tesla

Sharon Gustard, Jalal M. Fadili, Emma J. Williams, Laurance Hall, T. Adrian Carpenter, Matthew Brett, Edward Bullmore

▶ **To cite this version:**

Sharon Gustard, Jalal M. Fadili, Emma J. Williams, Laurance Hall, T. Adrian Carpenter, et al.. Effect of slice orientation on reproducibility of fMRI motor activation at 3 Tesla. Magnetic Resonance Imaging, 2001, 19 (10), pp.1323-1331. hal-01121179

HAL Id: hal-01121179
https://hal.science/hal-01121179

Submitted on 27 Feb 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Effect of slice orientation on reproducibility of fMRI motor activation at 3 Tesla

Sharon Gustard[a,b], Jalal Fadili[c], Emma J. Williams[b], Laurance D. Hall[a], T. Adrian Carpenter[b,*], Matthew Brett[d], Edward T. Bullmore[e]

[a]*Herchel Smith Laboratory for Medicinal Chemistry, School of Clinical Medicine, University of Cambridge, Cambridge, UK*
[b]*Wolfson Brain Imaging Centre, School of Clinical Medicine, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK*
[c]*Centre for Speech and Language, Department of Experimental Psychology, University of Cambridge, Cambridge, UK*
[d]*MRC Cognition and Brain Sciences Unit, Cambridge, UK*
[e]*Department of Psychiatry, School of Clinical Medicine, University of Cambridge, Cambridge, UK*

Abstract

The effect of slice orientation on reproducibility and sensitivity of 3T fMRI activation using a motor task has been investigated in six normal volunteers. Four slice orientations were used; axial, oblique axial, coronal and sagittal. We applied analysis of variance (ANOVA) to suprathreshold voxel statistics to quantify variability in activation between orientations and between subjects. We also assessed signal detection accuracy in voxels across the whole brain by using a finite mixture model to fit receiver operating characteristic (ROC) curves to the data. Preliminary findings suggest that suprathreshold cluster characteristics demonstrate high motor reproducibility across subjects and orientations, although a significant difference between slice orientations in number of activated voxels was demonstrated in left motor cortex but not cerebellum. Subtle inter-orientation differences are highlighted in the ROC analyses, which are not obvious by ANOVA; the oblique axial slice orientation offers the highest signal detection accuracy, whereas coronal slices give the lowest. © 2001 Elsevier Science Inc. All rights reserved.

Keywords: Slice orientation; fMRI; Reproducibility; 3T; ROC

1. Introduction

FMRI activation pattern reproducibility has been reported at clinical [1–6] and high [7,8] magnetic field strengths, using well characterised visual, motor and cognitive paradigms. All studies report high within-subject and variable between-subject reproducibility, depending on the method of analysis. A multi-site study has also been reported [9] demonstrating highly consistent activation patterns in response to a working memory paradigm. Selection of the slice orientation depends on the brain area being studied, but axial or oblique axial slices are most commonly used in fMRI studies for comparison with PET images [10], and ease of acquisition because older gradient sets only acquire EPI data in the X-Y plane. Oblique axial slices are popular because they enable the frontal sinuses to be avoided, and lower numbers of oblique slices are needed for whole brain coverage, which allows the TR and therefore acquisition time to be reduced. The effect of slice order [11], slice thickness [11,12], and voxel size [13] on fMRI activation have been investigated but, to our knowledge, the effect of slice orientation has not yet been reported.

We present a whole brain study to assess reproducibility in 4 different slice orientations, using a standard motor paradigm. The data were analysed using two different methods; 1) statistical parametric maps [14] were used to identify suprathreshold "activated" voxels, and the characteristics of resulting activation clusters was assessed by two-way ANOVA (with orientation as a within-subject factor and subject as a between-subject factor), and 2) receiver operating characteristic (ROC) curves [2,15] were fitted (Methods section) to compare signal-detection accuracy in all voxels across the brain, for each of the 4 slice orientations.

2. Materials and methods

2.1. Subjects and task

All studies were performed with permission from Addenbrooke's Hospital Local Research Ethics Committee.

* Corresponding author. Tel.: +44-1223-762128; fax: +44-1223-331826.
E-mail address: tac12@wbic.cam.ac.uk (T.A. Carpenter).

Informed consent was obtained from six right-handed, healthy volunteers (3 male, 3 female, aged 21–35 years, mean age 26.7 ± 5 years) who were studied using a blocked periodic AB design (A = motor task, B = rest) repeated for 4 cycles. The duration of each epoch was 40 seconds. Subjects performed a self-paced (approximately 2 Hz) sequential finger opposition task with their right hand during the motor task epochs, and were informed of the beginning and end of each epoch by means of a tactile cue on their right ankle. During the rest periods subjects were asked to rest the hand used for tapping on their leg.

2.2. Data acquisition

Imaging was performed with a 3 Tesla Bruker Medspec Avance S300 system (Bruker Medical, Ettlingen, Germany), using a head gradient set insert and quadrature (diameter 28 cm) birdcage head coil for receive and transmit. Comparable eddy currents in all 3 gradient axes were ensured by adjustment of the preemphasis by the manufacturer upon installation of the magnet to give comparable values in each axis. Subjects were given earplugs and positioned on isocentre. Shimming was firstly performed manually in X, Y and Z and then adjusted to the second order using the automatic shimming software FASTMAP [16]. Whole brain, single shot T_2^* weighted EPI images were acquired in "true" axial, oblique axial, coronal and sagittal slice orientations (TR = 5000 ms, TE = 30 ms, $\alpha = 90°$, isotropic FOV = 256 mm, 40 contiguous slices, 4 mm slice thickness, 64 repetitions after removal of 8 dummy scans). The axial, coronal and sagittal slices were orthogonal with respect to each other and aligned to the gradient axes, and the oblique axial slices were positioned parallel to the base of the brain. The order of acquisition of the 4 orientations was randomised for each subject. To control the influence of partial volume effects across the orientations an isotropic resolution of 4 mm^3 was desired. This EPI sequence samples data during the gradient ramps as well as the gradient plateaus, and the interpolation of the ramp-sampled data points back to a rectilinear grid during reconstruction would result in a reduced effective resolution. To avoid this, data were acquired with a matrix of 128 × 64, and the ramp-sampled data removed prior to reconstruction, leaving a matrix of 64 × 64 and a verified reconstructed resolution of 4 mm^3.

2.2.1. Phantom experiment

To identify the contribution of hardware variations in different slice orientations, the experiment was repeated on a spherical phantom (100 mm diameter, containing copper sulphate solution) which was positioned exactly on isocentre. Only 25 slices were needed for complete coverage, but all other parameters were identical.

2.3. Data analysis

2.3.1. Statistical parametric mapping

Individual EPI timecourses were processed using Statistical Parametric Mapping (SPM99) (www.fil.ion.ucl.ac.uk/spm/) software. The processing steps involved slice timing correction, realignment, normalisation to the EPI template, smoothing to a FWHM of 8 × 8 × 8 mm, and fitting to a user-specified model on a voxel-by-voxel basis. The effect of the normalisation process on images acquired in different slice orientations was assessed by omitting the normalisation step and fitting the same model to realigned, smoothed data. Uncorrected statistical data ($p = 0.001$) were used in order to be comparable with ROC methods. Activation clusters in the left motor cortex (LMC) and right cerebellum (RC) were assessed by visual inspection and in terms of peak and regional mean t-statistic, and number of voxels, using two-way ANOVAS calculated with respect to orientation and subject. Regional mean t-statistics were obtained using matlab code which calculated the mean of the t-statistics across all voxels comprising LMC and RC clusters; LMC and RC clusters were defined as groups of voxels in those brain regions above an uncorrected threshold of $P < 0.001$. Fixed-effect analyses were performed using SPM99 [17] (www.mrc-cbu.cam.ac.uk/Imaging/) on the spatially normalised data, creating mean activation maps for each subject across orientations and each orientation across subjects. Statistical maps were assessed in the same way as for the individual data sets.

2.3.2. ROC methods

The receiver operating characteristic (ROC) curve is commonly used to quantify signal detection accuracy. It is a plot of the sensitivity $1 - \beta$ versus α for different thresholds of the rating scale, where α and β respectively correspond to the probability of type I error (false positives fraction, FPF) and the probability of type II error (false negatives fraction, FNF). Several useful measures of detection accuracy can be derived, such as area under the ROC curve [15,18], where a greater area corresponds to better detection accuracy.

ROC curves can be fitted to data using parametric and non-parametric approaches. Non-parametric methods [2,19] specify a number of assumptions, which include repeating the experiment a minimum of 3 times, and that the observations (voxels) and individual trials are independent. Parametric methods [20,21] also assume voxel independence and that the true positives and true negatives form a binormal distribution.

We propose a mixture-based model to fit ROC curves to the observed data [22–26]. This model does not make the assumption of binormally distributed data or of a minimum number of replications, but it does specify independent voxels. The model was therefore fitted to the t-statistic maps which were processed without the smoothing step.

The model can be formulated in the following way:

$$f(x) = \lambda f_{H_1}(x) + (1 - \lambda) f_{H_0}(x) \quad (1)$$

where λ is the mixing proportion (the true proportion of activated voxels), $f(x)$ is the probability of observing an activated voxel, $f_{H_0}(x)$ and $f_{H_1}(x)$ are the probability density functions (pdf) under the null and the alternative hypotheses [27]. The model is fitted to the observed data using a maximum likelihood estimator (MLE) (see Appendix).

Reliability criteria can be determined from ROC curves. While classical approaches for statistical map thresholding focus on the FPF, we propose another approach which involves maximising a criterion to determine the best trade-off between the FPF and True Positive Fraction (TPF). The reliability criterion (RC) maximises the TPF and the true negatives fraction (TNF) in the thresholded map, using,

$$RC(\alpha, \beta) = \hat{\lambda}(1 - \beta) + (1 - \hat{\lambda})(1 - \alpha) \quad (2)$$

where $\hat{\lambda}$ is the MLE estimate of λ and $(1 - \alpha)$ is the TNF. Differentiating RC gives the threshold corresponding to the point on the ROC whose slope is $\frac{1-\hat{\lambda}}{\hat{\lambda}}$. Another approach could consist of choosing the point on the ROC curve with a gradient of 1. If any prior knowledge is available on the expected true positive or false positive fractions, other weighting coefficients can be used instead of $\hat{\lambda}$ and $1 - \hat{\lambda}$.

2.3.3. Phantom analysis

EPI images were reconstructed and analysed using in-house software. Six slices evenly spaced through the phantom were cut out from a total of 25 slices (slices 2, 6, 10, 14, 18, and 22) in each orientation, to observe signal-to-noise across the whole phantom. A 200 pixel circular region of interest, on isocentre (to minimise effects of B_0 drift) and in the centre of all slices and orientations, was used to determine the mean and standard deviation of the signal in that region for all 64 time points. The mean and standard deviation of each timecourse was calculated in all 6 slice numbers and all 4 orientations, and used to plot line graphs. The timecourse means and standard deviations were used to estimate a global mean and standard deviation for all 4 slice orientations. One-way ANOVAs were performed on the individual slice data with respect to slice orientation.

3. Results

3.1. SPM analysis

Fig. 1 shows group activation maps for a) each orientation averaged over subjects and b) each subject averaged over orientations, both maps being thresholded at a P value of 0.001. Characteristic left motor cortex (LMC), right cerebellum (RC) and supplementary motor area (SMA) activations are observed in all subjects and orientations, and RMC activation is present in axial, coronal and sagittal orientations and subjects 2, 4, 5 and 6. There is a difference in extent of LMC activation pattern between slice orientations. By visual inspection, the order of decreasing size of activation pattern was axial, sagittal, coronal and oblique axial. In the cerebellum the corresponding order was oblique axial, coronal, sagittal and axial. Fig. 2 shows histograms of the peak and regional mean t-statistic, and number of activated voxels in the LMC and RC. Table 1 shows results of a two-way analysis of variance (slice orientation and subject being the within- and between-subject factors respectively) with each of these three parameters treated separately as the dependant variable. Figs. 2a and 2b suggest that regional mean t-statistic is a more robust measure than peak t-statistic. Both inter-subject variability and inter-orientation variability in the LMC are reduced using mean t values compared to peak t values (Table 1). Fig. 2b suggests that signal change in cerebellar activation is as reproducible as primary cortical activation, although there are large variations in extent (Fig. 2c). Fig. 2 and Table 1 show that there are no significant differences between slice orientations except in the number of activated voxels in the LMC. It should also be noted (Figs. 1b and 2, Table 1) that inter-subject variability is of a similar order of magnitude to inter-orientation differences.

3.2. Phantom experiment

Fig. 3 shows the mean and standard deviation of signal intensity over time (64 repetitions) in a 200 pixel ROI. Results for six slices through a spherical phantom in 4 slice orientations are shown. All 4 orientations gave results of the same order of magnitude and there are no significant differences between the orientations (Table 1).

3.3. ROC analysis

Estimated ROC curves averaged over 6 subjects for the 4 slice orientations are shown in Fig. 4. These curves are calculated over a relatively conservative range of false positive thresholds (0–0.025) in terms of fMRI analyses. The curves do not cross each other and visual inspection of the area under each one suggests a decreasing relative order of signal detection accuracy to be oblique axial, axial, sagittal and coronal. The ellipses, indicative of the range of suitable thresholds over which signal detection accuracy is maximised, show that the axial orientation provides the widest TPF range and smallest FPF range. The oblique axial orientation has a similar TPF range to axials but the FPF range is twice as large. The largest FPF range is seen in the sagittals, and while coronals offer the smallest TPF range, they also have the second smallest FPF range. The solid black dots (Fig. 4) mark the point at which the TPF/TNF trade-off is maximised, and suggest that different false positive thresholds might be appropriate for different slice orientations. The lowest false positive threshold would be achieved using an oblique axial slice orientation. The area

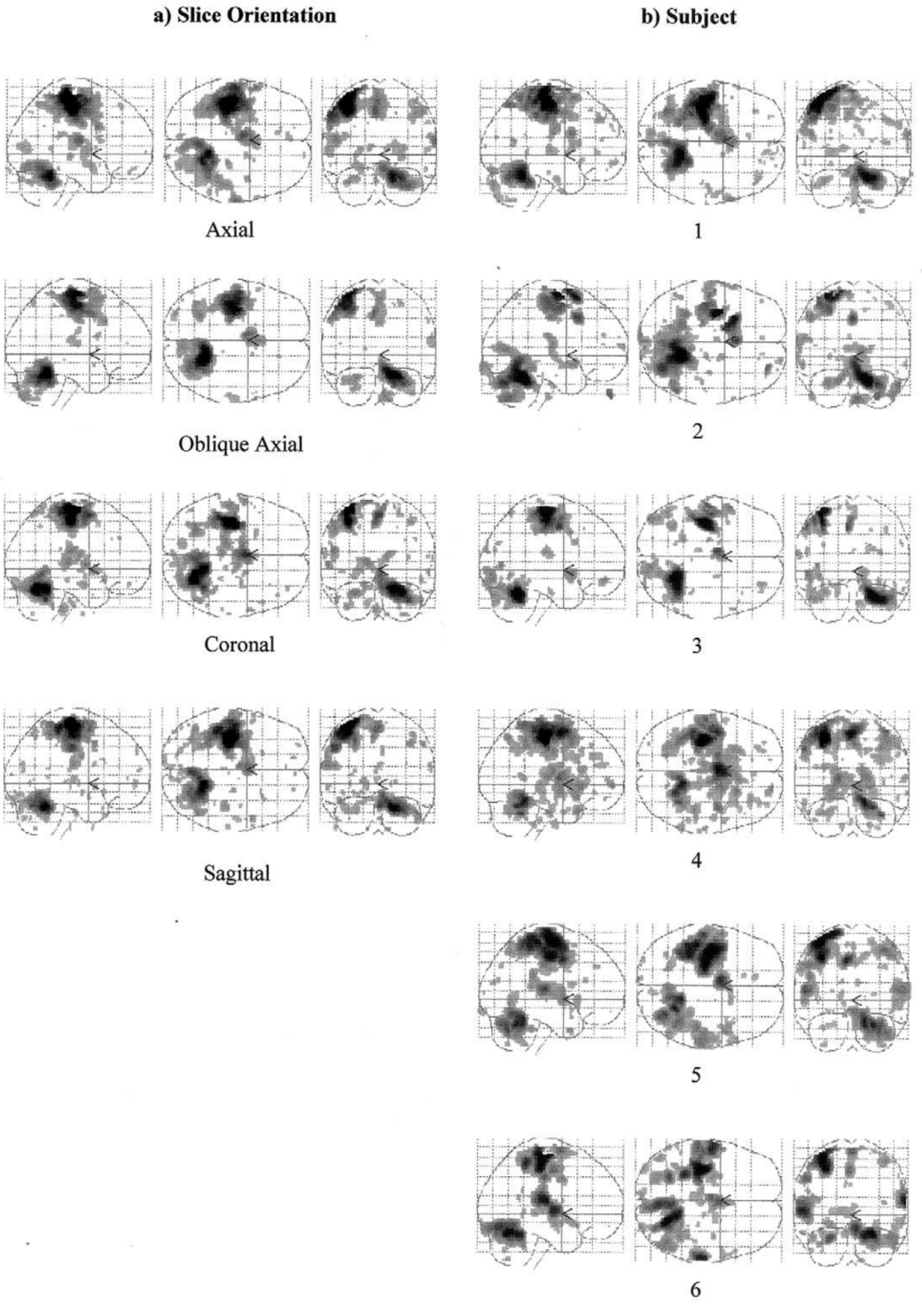

Fig. 1. Group activation maps for a) each orientation averaged over subjects and b) each subject averaged over orientations, thresholded at an uncorrected p value of 0.001.

under the ROC curves over a more stringent range of false positive thresholds typical of fMRI analyses (0–0.005) (Fig. 5) illustrates an incrementally decreasing signal detection accuracy of order oblique axial, axial, sagittal and coronal.

4. Discussion

The effect of slice orientation on motor task activation has been investigated using parametric methods to look at specific regions of activation, and ROC methods which

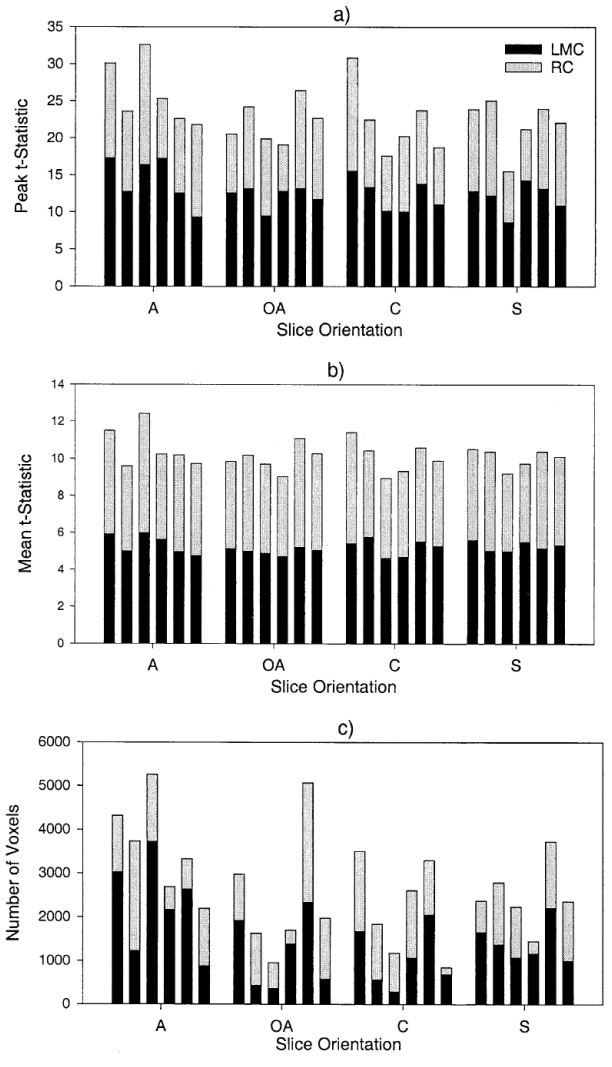

Fig. 2. Uncorrected ($p < 0.001$) a) peak t-statistic, b) mean t-statistic, and c) number of voxels in left motor cortex (LMC) and right cerebellum (RC) for axial (A), oblique axial (OA), coronal (C) and sagittal (S) slice orientations. Subjects 1 to 6 are displayed from left to right in each of the 4 orientations.

evaluate signal detection accuracy across the whole brain.

Parametric analyses which reflect cluster-level characteristics (regional mean t-statistic) demonstrate improved reproducibility across orientations and subjects compared to voxel-level characteristics (peak t-statistic) (Table 1 and Fig. 2). Cluster-level cerebellar activation appears as robust as cortical activation [10,28,29], rendering it equally suitable for reproducibility studies. Inter-subject variability was generally observed to be of the same order of magnitude as inter-orientation differences for both cluster- and voxel-level characteristics, suggesting that a larger number of subjects may need to be studied. This is frequently a limiting factor in quantitative fMRI studies. A variety of statistical measures were used as dependent variables in the study, and analysis of variance results show that certain suprathreshold cluster characteristics reveal differences between orientations, whereas others do not. Peak t-statistic and number of suprathreshold voxels are conventional statistical measures, however the former is derived from a single voxel and the latter reported in the literature to be susceptible to high inter-subject variability [4,30,31]. The regional mean t-statistic provided an estimate of the average t-statistic across all voxels in the cluster, which although unconventional, proved to be highly reliable across subjects and orientations. Therefore careful selection of the measurement parameters is important in methodological fMRI studies. The only significant difference between orientations was in terms of the number of voxels in the LMC, for which there are a number of possible causes.

One possible source of the observed differences may lie in the instrumentation. Phantom experiments were performed to assess the contribution of hardware, specifically gradient performance, to these differences. Methods of monitoring scanner stability have already been reported [32] and their implementation could provide a quantitative measure of the contribution of hardware variations to the overall inter-subject variability which is observed in fMRI. Our results suggest that there are no significant differences ($p = 0.279$) in performance of the different gradients.

It was also hypothesised that images acquired in different orientations may suffer from different amounts of geometric distortion, exacerbated at 3 Tesla [33], resulting in variations in the quality of normalisation to a relatively undistorted 2 Tesla template. However, the normalisation parameters (not shown) and suprathreshold voxel analyses indicate comparable quality across slice orientations. The

Table 1

	Slice Orientation					Subject				
Dependent Variable	SS	df	MS	F	P	SS	df	MS	F	P
LMC - Peak t-statistic	20.140	3	6.714	1.548	0.243	42.750	5	8.551	1.971	0.142
LMC - Mean t-statistic	0.439	3	0.146	0.890	0.496	0.463	5	0.093	0.563	0.727
LMC - No. Voxels	5.48e + 0.6	3	1.83e + 06	5.157	0.012	7.48e + 06	5	1.50e + 06	4.223	0.014
RC - Peak t-statistic	14.500	3	4.835	0.698	0.568	36.130	5	7.277	1.043	0.429
RC - Mean t-statistic	0.690	3	0.233	0.756	0.536	2.121	5	0.424	1.376	0.288
RC - No. Voxels	1.81e + 05	3	6.04e + 04	0.140	0.935	2.46e + 06	5	4.91e + 05	1.136	0.384
Phantom	7.09e + 15	3	2.36e + 15	1.375	0.279					

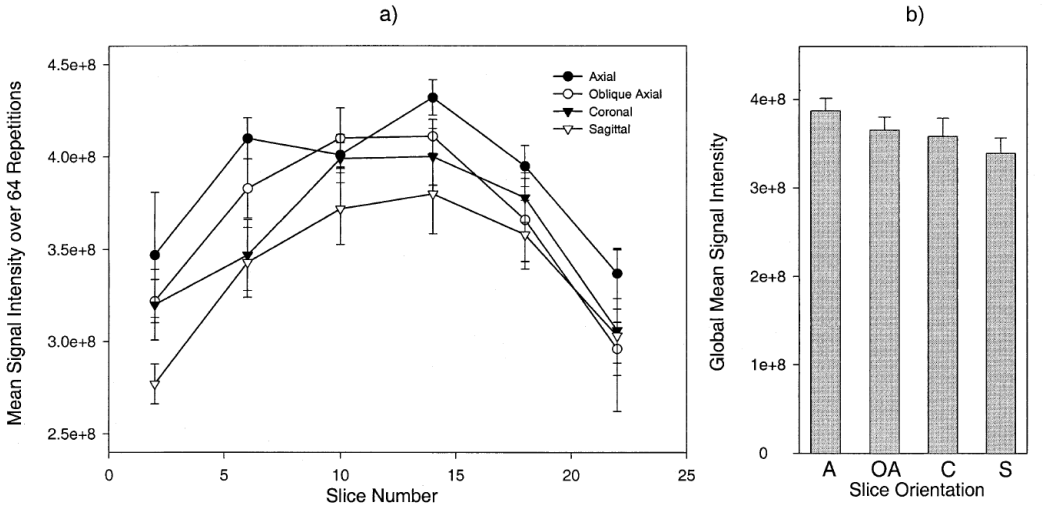

Fig. 3. Mean signal intensity over time (64 repetitions) for a 200 pixel ROI in 6 different slices through a spherical phantom (100 mm diameter) acquired in 4 different slice orientations. Error bars show the mean of the standard deviation in each slice. Error bars representing standard deviation of the mean would be entirely covered by the symbols. Part b) shows the global mean and standard deviation of the values in part a), for each slice orientation.

use of self-paced rather than paced finger opposition may have introduced spurious differences between the 4 orientations. This is unlikely to fully explain the inter-orientation differences observed since the order of acquisitions was randomised in each imaging session to minimise effects of habituation or fatigue. There are a number of other possible explanations for the observed orientation effect on motor cortex, which include differential degrees of head motion,

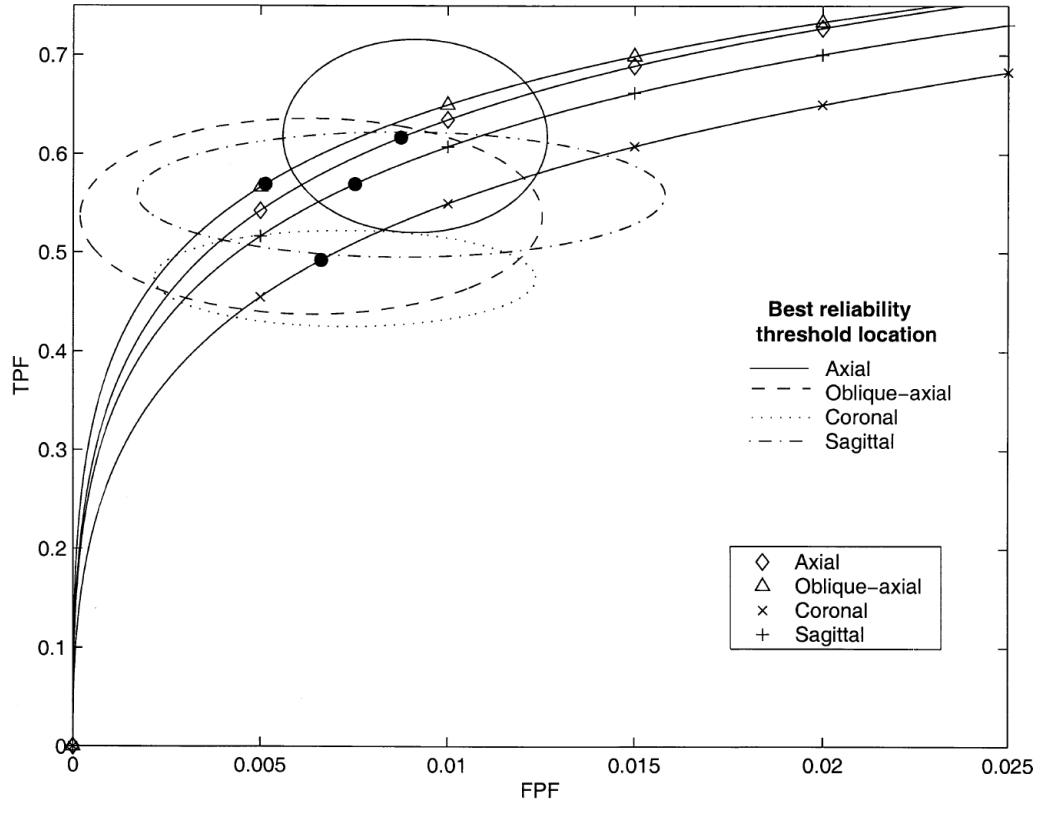

Fig. 4. Mean estimated ROC curves at low false positive fractions (FPFs) for axial, oblique axial, coronal and sagittal slice orientations. Solid black dots mark the point on the curve corresponding to a gradient of $(1 - \lambda)/\lambda$, where the trade-off between Type I and Type II error is maximised.

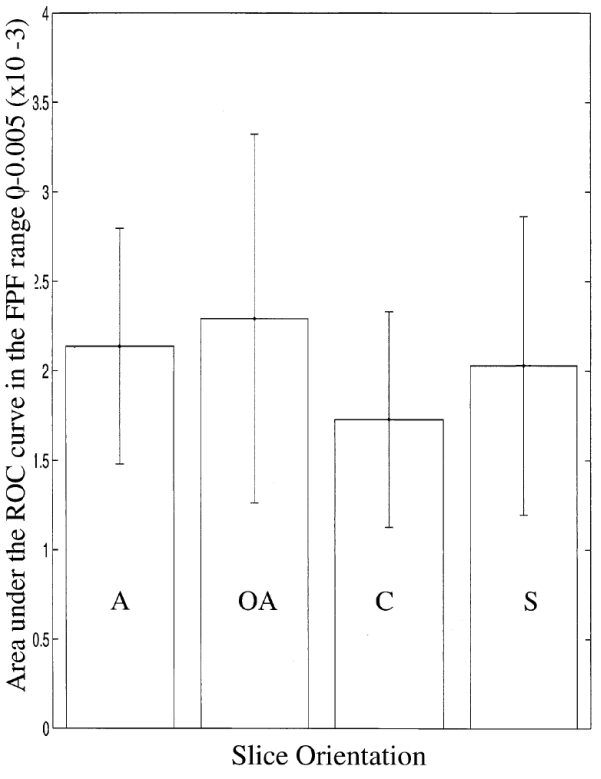

Fig. 5. Area under the ROC curve for the 4 slice orientations over a stringent range of false positive thresholds (0–0.005) common to fMRI analyses.

local variations in field homogeneity (shimming), and the effect of vessel orientation with respect to the main field.

We consider the most likely explanation is that partial volume effects resulting from vessel architecture in the strip-shaped motor cortex are responsible. Different slice orientations intersect this architecture at different angles, and blood vessels involved in the BOLD response occupy different proportions of a given voxel contributing to the image and therefore activation map. Partial volume effects due to resolution [13] and slice thickness can also affect the magnitude [12] and pattern [11] of fMRI activation. The fact that a significant effect of orientation was not also observed in the cerebellum further suggests that the results are related to the inherent strip shape of the motor cortex.

ROC curves probed the signal detection accuracy offered by each slice orientation using voxels across the whole brain. The finite mixture model assumes independence of voxel statistics and therefore these ROC curves were estimated from maps of t-statistic without the application of spatial smoothing. A degree of spatial dependence will exist nonetheless due to factors such as the point spread function of the scanner, and T_2^* decay. However, we have assumed this to be consistent between orientations, permitting the relativistic comparison presented here. ROC methods are beneficial because they are relatively insensitive to the distribution of the underlying data [15] and the concern as to whether noise distributions (physiologic, motion etc.) have been correctly modelled is less crucial.

The area under the ROC curves at stringent ($p < 0.005$) fMRI thresholds for Type I error control indicate that an incrementally higher signal detection accuracy is achieved using oblique axial slices compared to axial, sagittal and coronal. The lowest relative accuracy is given with coronal slices. The FPF corresponding to a gradient of $(1 - \lambda)/\lambda$ on each curve suggests that different orientations may require different false positive thresholds to achieve the optimal Type I/Type II error trade-off in detection accuracy. The lowest possible threshold is desirable to give the highest statistical significance, and is obtained using the oblique axial slice orientation. ROC methods could be used in conjunction with suprathreshold voxel analyses by providing a means of assigning an appropriate threshold for control of Type I error. ROC methods are also attractive because they consider both statistically significant and statistically insignificant voxels across the whole brain. Such analyses have revealed here incremental differences between orientations that were not clear from the analysis of suprathreshold voxels alone, suggesting that these inter-orientation differences are not restricted to the most statistically significant voxels.

5. Conclusions

These preliminary findings suggest that suprathreshold cluster characteristics demonstrate high motor reproducibility across subjects and orientations. Subtle inter-orientation differences are highlighted in the ROC analyses, which are not obvious in the most statistically significant voxels; the oblique axial slice orientation offers the highest signal detection accuracy, whereas coronal slices give the lowest.

Acknowledgments

We thank Dr. Herchel Smith for the Endowment (SG, LDH). We thank the providers of the Technology Forsight grant which funds the magnet and staff positions (EJW).

References

[1] Rombouts SARB, Barkhof F, Hoogenraad FGC, Sprenger M, Scheltens P. Within-subject reproducibility of visual activation patterns with functional magnetic resonance imaging using multislice echo planar imaging. Magn Reson Imaging 1998;16:105–13.

[2] Noll DC, Genovese CR, Nystrom LE, Vazquez AL, Forman SD, Eddy WF, Cohen JD. Estimating test-retest reliability in functional MR imaging II: Application to motor and cognitive activation studies. Magn Reson Med 1997;38:508–17.

[3] Wexler BE, Fulbright RK, Lacadie CM, Skudlarski P, Kelz MB, Constable RT, Gore JC. An fMRI study of the human cortical motor system response to increasing functional demands. Magn Reson Imaging 1997;15:385–96.

[4] Ramsey NF, Tallent K, Gelderen P, Frank JA, Moonen CTW, Weinberger DR. Reproducibility of human 3D fMRI brain maps acquired during a motor task. Human Brain Mapping 1996;4:113–21.
[5] Moser E, Teichtmeister C, Diemling M. Reproducibility and postprocessing of gradient-echo functional MRI to improve localization of brain activity in the human visual cortex. Magn Reson Imaging 1996;14:567–79.
[6] Baumgartner R, Scarth G, Teichtmeister C, Somorjai R, Moser E. Fuzzy clustering of gradient-echo functional MRI in the human visual cortex. Part I: Reproducibility. J Magn Reson Imaging 1997;7:1094–101.
[7] Tegeler C, Strother SC, Anderson JR, Kim S-G. Reproducibility of BOLD-based functional MRI obtained at 4T. Human Brain Mapping 1999;7:267–83.
[8] Cohen MS, DuBois RM. Stability, repeatability, and the expression of signal magnitude in functional magnetic resonance imaging. J Magn Reson Imaging 1999;10:33–40.
[9] Casey BJ, Cohen JD, O'Craven K, Davidson RJ, Irwin W, Nelson CA, Noll DC, Hu X, Lowe MJ, Rosen BR, Truwitt CL, Turski PA. Reproducibility of fMRI results across four institutions using a spatial working memory task. Neuroimage 1998;8:249–61.
[10] Joliot M, Papathanassiou D, Mellet E, Quinton O, Mazoyer N, Courtheoux P, Mazoyer B. FMRI and PET of self-paced finger movement: Comparison of intersubject stereotaxic averaged data. Neuroimage 1999;10:430–47.
[11] Howseman AM, Grootoonk S, Porter DA, Ramdeen J, Holmes AP, Turner R. The effect of slice order and thickness on fMRI activation data using multislice echo-planar imaging. Neuroimage 1999;9:363–76.
[12] Frahm J, Merboldt K-D, Hanicke W. Functional MRI of human brain activation at high spatial resolution. Magn Reson Med 1993;29:139–44.
[13] Hyde JS, Biswal BB, Jesmanowicz A. Optimum voxel size in fMRI. In: Proceedings of the 8th Annual Meeting of the International Society for Magnetic Resonance in Medicine, Denver2000 S240.
[14] Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RSJ. Statistical parametric maps in functional imaging: A general linear approach. Human Brain Mapping 1995;2:189–210.
[15] Sorenson JA, Wang X. ROC methods for evaluation of fMRI techniques. Magn Reson Med 1996;36:737–44.
[16] Gruetter R. Automatic, localized *in vivo* adjustment of all first- and second-order shim coils. Magn Reson Med 1993;29:(6)804–11.
[17] Woods RP. Modelling for intergroup comparisons of imaging data. Neuroimage 1996;4:(3/3)S84–94.
[18] Skudlarski P, Constable RT, Gore JC. ROC analysis of statistical methods used in functional MRI: Individual subjects. Neuroimage 1999;9:311–29.
[19] Genovese CR, Noll DC, Eddy WF. Estimating test-retest reliability in functional MR imaging I: statistical methodology. Magn Reson Med 1997;38:497–507.
[20] Dorfman D, Alf E. Maximum likelihood estimation of parameters of signal-detection theory and determination of confidence intervals-rating method data. J Math Psychol 1969;6:487–96.
[21] Metz CE. ROC methodology in radiological imaging. Invest Radiol 1986;21:72–3.
[22] Day NE. Estimating the components of a mixture of normal distributions. Biometrica 1969;56:(3)463–74.
[23] Everitt BS, Hand DJ. Finite mixture distributions. 1st ed. London: Chapman and Hall, 1981.
[24] Redner RA, Walker HF. Mixture densities, maximum likelihood and the EM algorithm. Siam Review 1984;26:(2)195–239.
[25] Titterington EM, Smith AFM, Makov UE. Statistical analysis of finite mixture distributions. New York: John Wiley and Sons, 1985.
[26] Mclachlan GJ, Basford KE. Mixture models: inference and applications to clustering. New York: M. Dekker, 1988.
[27] Everitt BS, Bullmore ET. Mixture model mapping of brain activation in functional magnetic resonance images. Human Brain Mapping 1999;7:1–14.
[28] Singh LN, Higano S, Takahashi S, Abe Y, Sakamoto M, Kurihara N, Furuta S, Tamura H, Yanagawa I, Fujii T, Ishibashi T, Maruoka S, Yamada S. Functional MR imaging of cortical activation of the cerebral hemispheres during motor tasks. Am J Neuroradiol 1998;19:275–80.
[29] Desmond JE, Gabrieli JDE, Wagner AD, Ginier BL, Glover GH. Lobular patterns of cerebellar activation in verbal working memory and finger-tapping tasks as revealed by functional MRI. J Neuroscience 1997;17:9675–85.
[30] Constable RT, Skudlarski P, Mencl E, Pugh KR, Fulbright C, Lacadie S, Shaywitz SE, Shaywitz BA. Quantifying and comparing region-of-interest activation patterns in functional brain MR imaging: methodology considerations. Magn Reson Imaging 1998;16:289–300.
[31] Machielsen WCM, Rombouts SARB, Barkhof F, Scheltens P, Witter MP. FMRI of visual encoding: reproducibility of activation. Human Brain Mapping 2000;9:156–64.
[32] Weisskoff RM. Simple measurement of scanner stability for functional NMR imaging of activation in the brain. Magn Reson Med 1996;36:643–45.
[33] Jezzard P, Clare S. Sources of distortion in functional MRI data. Human Brain Mapping 1999;8:80–5.

Appendix

Maximum Likelihood Estimator

The model is fitted to the observed data using a maximum likelihood estimator. The log-likelihood can be expressed as follows:

$$LL(\lambda, \delta) = \sum_{i=1}^{Nvoxels} \text{Log}\,(f(x_i, \lambda, \delta)) \tag{3}$$

where λ is as before (1) and δ is the non-centrality parameter. The pdfs under the null and alternative hypotheses for a given number of degrees of freedom can be numerically expressed in terms of gamma functions. Maximum likelihood estimates are obtained by maximising the log-likelihood (LL), which is achieved by differentiating LL and setting to zero. However the LL is a complicated function and a numerical maximisation procedure, the modified Newton method, is necessary.

The MLE is consistent and efficient (at least asymptotically), and tends to a normal distribution as the number of observations tends towards infinity. This is relevant to our experiment since the number of intra-cranial voxels is high. The variances of the parameters are therefore well approximated by the diagonal elements of the Fisher information matrix.

This is calculated using the inverse of the Hessian at the ML estimate of the parameters:

$$H = \begin{pmatrix} -\sum_i \left(\frac{f_{H_1}(x_i) - f_{H_0}(x_i)}{f(x_i)} \right)^2 & \sum_i \frac{f'_{H_1}(x_i) f_{H_0}(x_i)}{f^2(x_i)} \\ \sum_i \frac{f'_{H_1}(x_i) f_{H_0}(x_i)}{f^2(x_i)} & \sum_i \frac{\lambda f''_{H_1}(x_i) f(x_i) - \lambda^2 (f'_{H_1}(x_i)^2}{f^2(x_i)} \end{pmatrix} \tag{4}$$

The asymptotic gaussian property of the estimator gives also a very good test for comparing differences in parameters between datasets.