NeuroImage **15,** 902–907 (2002)
doi:10.1006/nimg.2002.1053, available online at http://www.idealibrary.com on IDEAL®

Comparison of Detrending Methods for Optimal fMRI Preprocessing

Jody Tanabe,* David Miller,* Jason Tregellas,† Robert Freedman,†,‡ and Francois G. Meyer§

**Department of Radiology, ‡Department of Psychiatry, and †Department of Pharmacology, University of Colorado Health Sciences Center, Denver, Colorado 80262; and §Department of Electrical and Computer Engineering, University of Colorado, Boulder, Boulder, Colorado 80309*

Received July 9, 2001

Because of the inherently low signal to noise ratio (SNR) of fMRI data, removal of low frequency signal intensity drift is an important preprocessing step, particularly in those brain regions that weakly activate. Two known sources of drift are noise from the MR scanner and aliasing of physiological pulsations. However, the amount and direction of drift is difficult to predict, even between neighboring voxels. Further, there is no concensus on an optimal baseline drift removal algorithm. In this paper, five voxel-based detrending techniques were compared to each other and an auto-detrending algorithm, which automatically selected the optimal method for a given voxel time-series. For a significance level of $P < 10^{-6}$, linear and quadratic detrending moderately increased the percentage of activated voxels. Cubic detrending decreased activation, while a wavelet approach increased or decreased activation, depending on the dataset. Spline detrending was the best single algorithm. However, auto-detrending (selecting the best algorithm or none, if detrending is not useful) appears to be the most judicious choice, particularly for analyzing fMRI data with weak activations in the presence of baseline drift. © 2002 Elsevier Science (USA)

Key Words: **detrending; functional magnetic resonance imaging (fMRI); noise; low frequency drift.**

INTRODUCTION

The magnitude of signal intensity changes that underlie the detection of brain activation using MRI is small, requiring multiple repetitions to detect stimulus-correlated signal variations. This limits the flexibility of the paradigms that can be used for fMRI experiments. Therefore removing baseline drifts, whether emanating from the scanner or incurred by the aliasing of physiological pulsations, is crucial. Physiological noise can be reduced by straightforward measures that include gating and post-hoc filtering (Biswal *et al.*, 1996; Buonocore and Maddock, 1997) or more complicated corrections in k-space using navigator echoes or external monitoring and retrospective estimation (Le and Hu, 1996; Hu *et al.*, 1995). Sources other than cardiac and respiratory fluctuations are more difficult to eliminate. These include CSF flow and spontaneous low-frequency fluctuations, the latter hypothesized to be related to spontaneous neuronal activity (Biswal *et al.*, 1996; Kiviniemi *et al.*, 2000). The MR scanner itself is likely a substantial source of noise. Smith *et al.* found low frequency drifts in cadavers in regions of high spatial intensity changes, not only at the edges of the brain, but also the infoldings of cortex, precisely where fMRI activations are localized (Smith *et al.*, 1999). Thus, removing baseline drifts plays an important role in pre-processing data prior to statistical analysis for fMRI studies performed at both clinical and high (4T) field (Tegeler *et al.*, 1999).

Linear and higher order polynomial (Bandettini *et al.*, 1993; Mattay *et al.*, 1996; Skudlarski *et al.*, 1999) and spline (Genovese, 1997) algorithms have been variously used for post hoc removal of low frequency signal drifts. Wavelets have been used on several occasions for removing Rician noise from MR images (Nowak, 1999; Wood and Johnson, 1999) and fMRI data (Ngan *et al.*, 2000; LaConte *et al.*, 2000; Zaroubi and Goelman, 2000) and recently, Meyer has proposed using wavelets to model the drift in fMRI data (Meyer and McCarthy G, 2001). Few studies have compared these techniques *in vivo*. Predicting the amount, type, and magnitude of drift is difficult since the etiology is multifactorial and poorly understood. Rather than limit data analysis to one procedure, Lange *et al.* introduced a pluralistic and empirical strategy for analyzing fMRI data (Lange *et al.*, 1999). Similar to Lange *et al.*, we compared several detrending methods that are commonly used in fMRI data analysis, and extend this pluralistic approach to the voxel level. While some have looked at nonlinear and frequency domain filters (Mitra *et al.*, 1997; Marchini and Ripley, 2000), this paper focuses exclusively on a comparison of fairly easy to implement linear models.

METHODS

Acquisition and paradigm. MR data were acquired in two healthy individuals using a standard quadra-

1053-8119/02 $35.00

ture head coil on a 1.5T MR system (Magnetom VISION; Siemens AG, Iselin, NJ). Volunteers provided written, informed consent approved by the human IRB. We used a visual smooth pursuit eye tracking task (Radant and Hommer, 1992). We chose a task known to produce reliable activations (Berman *et al.*, 1999; Petit and Haxby, 1999), but not necessarily as robust as simple finger-tapping or flashing checkerboard. A white dot traversing across a black background was projected onto a screen, which the subject viewed using a mirror attached to the head coil. The dot moved at 16.7°/s and the movement subtended a visual angle of 28°. During rest the subject was instructed to look straight ahead at the black screen. For each run, a 10-s equilibration period was followed by alternating 25-s rest/25-s task for four cycles. Twenty volumes were acquired per cycle. A total of 80 gradient-echo EPI (TE50/TR2500) data sets were collected with a 64^2 matrix over a 240 mm^2 FOV. Each data set consisted of 20 axial 6-mm slices, 1-mm gap, angled parallel to the planum sphenoidale. Three runs were acquired.

Image analysis and detrending methods. Images were motion-corrected in SPM 99, and then analyzed using an in-house application written in IDL (Interactive Data Language; RSI, Inc.). The three runs were detrended separately, corrected for mean signal difference, concatenated, and finally analyzed using a *t* test. Five different detrending algorithms were tested, based on linear, quadratic, cubic, spline, and wavelet models of low frequency noise. Also, an automated detrending routine was examined, which, rather than apply one method to the entire data set, chose the optimal detrending algorithm for each voxel-based time series. The statistical maps generated using each method were modified to account for differences in degrees of freedom (DOF). For example, in the case of spline detrending the *P* value was calculated using a DOF of 221 (3 runs of 80 vol = 240—3 runs * 5 spline parameters per run—2 mean corrections for the second and third runs—1 for each group compared in the *t* test * 2 groups).

Linear, quadratic, and cubic detrending consisted of first modeling the voxel time-series using the appropriate polynomial. The detrended signal was obtained by subtracting the modeled signal from the original signal. Spline detrending was accomplished by first breaking up the voxel time series into an arbitrary number of segments, three in this case (a greater number of segments allows the spline to model higher frequency components of the signal). The mean for each segment was then calculated, followed by values for the beginning and end points, which were obtained from the mean of the first half of the first segment and the mean of the last half of the last segment. This resulted in five control points or knots, through which we passed a cubic-spline to obtain the modeled signal. Wavelet detrending was implemented based on a proposal by Meyer, where the fMRI signal is described as the sum of two effects: a smooth baseline drift and the response to the stimulus. In the wavelet domain, this model assumes that the trend belongs to a subspace spanned by large-scale wavelets. Wavelet scales can vary from 0 to $\log_2$ (number of data points), where scale 0 returns the mean of the signal and maximum scale returns the original signal. We found that a scale of 1 appropriately tracked low frequency noise. Then, a fast algorithm, based on a wavelet representation of the data, was applied to yield the detrended time-series (Meyer and McCarthy G, 2001). In the case of auto-detrending, the routine chose the detrending method that resulted in the lowest *P* value (DOF corrected) on a voxel-by-voxel basis.

No spatial or temporal smoothing was applied. All *t* maps were thresholded at $P < 10^{-6}$ (approximately Bonferroni correction). A 2-D clustering algorithm was applied to remove isolated voxels on the activation map (any two or more activated voxels within a 3×3 space were considered a cluster). The resulting number of activated voxels were totaled over the supratentorial brain (top 15 slices) and compared across algorithms.

In addition to comparing methods over all supratentorial voxels, we also wanted to determine which, if any, algorithm was selected if the analysis were restricted to specific regions of interest (ROI) and auto-detrending enabled. *A priori* ROIs included frontal eye field, parietal eye field, visual cortex, and supplementary eye field (SEF). Information obtained from the ROI included which detrending algorithms were selected (on a voxel-by-voxel basis), the averaged time series and the power spectrum.

Null dataset. A null dataset was collected on one subject using MR acquisition parameters and data analysis as above. There was no stimulus.

RESULTS

Figure 1 shows that, compared to concatenation alone, correcting for the mean signal difference between runs resulted in a substantial increase in the activated voxels across the whole brain. For all subsequent analyses, the individual detrending methods were compared to data that was corrected for mean signal difference between the three runs (mean-corr). Compared to mean-corrected, linear (5–22%) and quadratic (13–31%) detrending moderately increased, cubic decreased (−1 to −12%), and wavelet (16 to −17%) methods either increased or decreased the number of activated voxels, depending on the dataset. Spline detrending resulted in the greatest increase by any one technique (19–50%). Because no single algorithm significantly outperformed others, the auto-detrending method further increased the number of activated voxels. For both subjects, the number of activated voxels

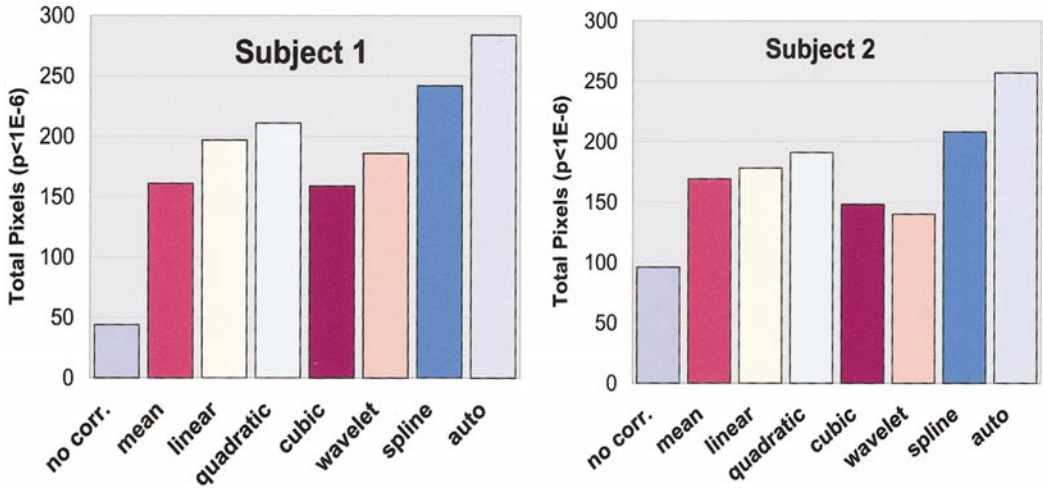

FIG. 1. Total activated voxels (whole brain) using different detrending methods in two subjects. (no corr., simple concatenation; mean, normalized to mean signal intensity for each run; auto, auto-detrending).

increased by about 150% compared to mean-correction. Both subjects showed activation in expected brain regions. Figure 2 demonstrates the increase in activated voxels in occipital cortex and dorsal cortical eye fields with the auto-detrending compared to mean-corrected algorithm.

In addition to whole brain analysis, we wanted to determine which of the detrending algorithms was selected among voxels limited to the regions of interest described earlier. Forty-two percent of voxels did not benefit from detrending, but among those voxels that did, spline detrending was most frequently selected (Fig. 3). Linear detrending accounted for only 1% of selected voxels.

Power spectra in Fig. 4 demonstrate that the auto-detrending method removed only low frequencies, at or below the task frequency of ~0.02 Hz. This was true for both an ROI (Figs. 4a and 4b) and for a slice containing the ROI (Figs. 4c and 4d). The purpose of analyzing the slice was to assess whether false positive activations were introduced since the risk of false positive activation is expected to be higher for a slice compared to an ROI. No component of the spectra increased after detrending.

To determine if there were any spatial patterns of low frequency drift, we calculated drift maps for the *in*

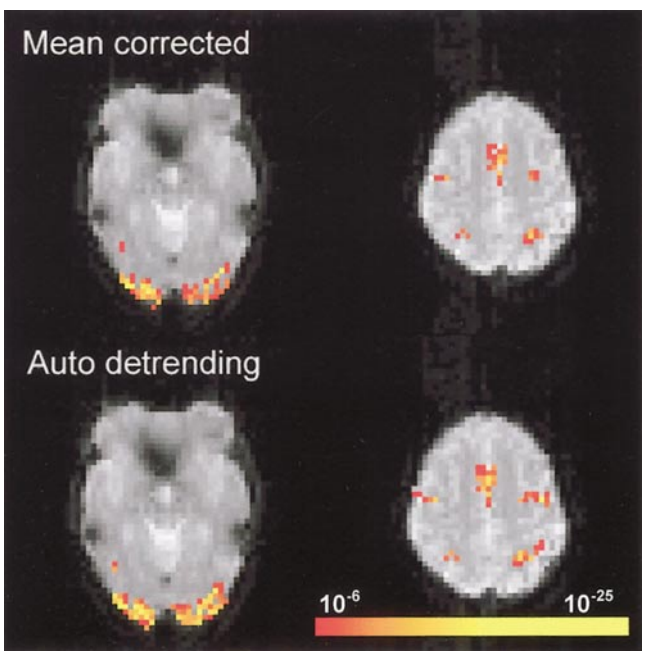

FIG. 2. Activation map of smooth pursuit eye movements demonstrates focal activation in frontal, parietal, and supplementary eye fields, and occipital cortex. Within these regions, auto-detrending revealed new activated voxels and lowered *P* values of already activated voxels. Top row: mean-corrected. Bottom row: auto-detrending. *P* values calculated using corrected degrees of freedom.

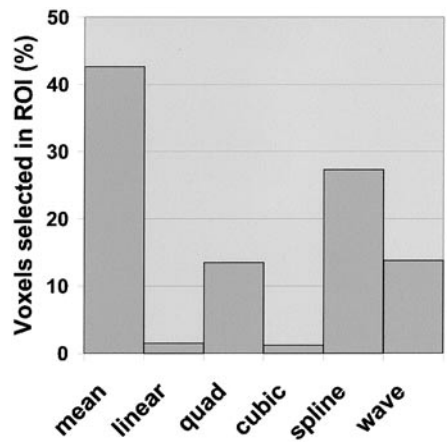

FIG. 3. The detrending algorithm automatically selected. 42% of voxels within combined ROIs did not improve with any detrending algorithm. Among remaining voxels, the spline was selected most often while linear and cubic detrending were selected least often. Both subjects are included.

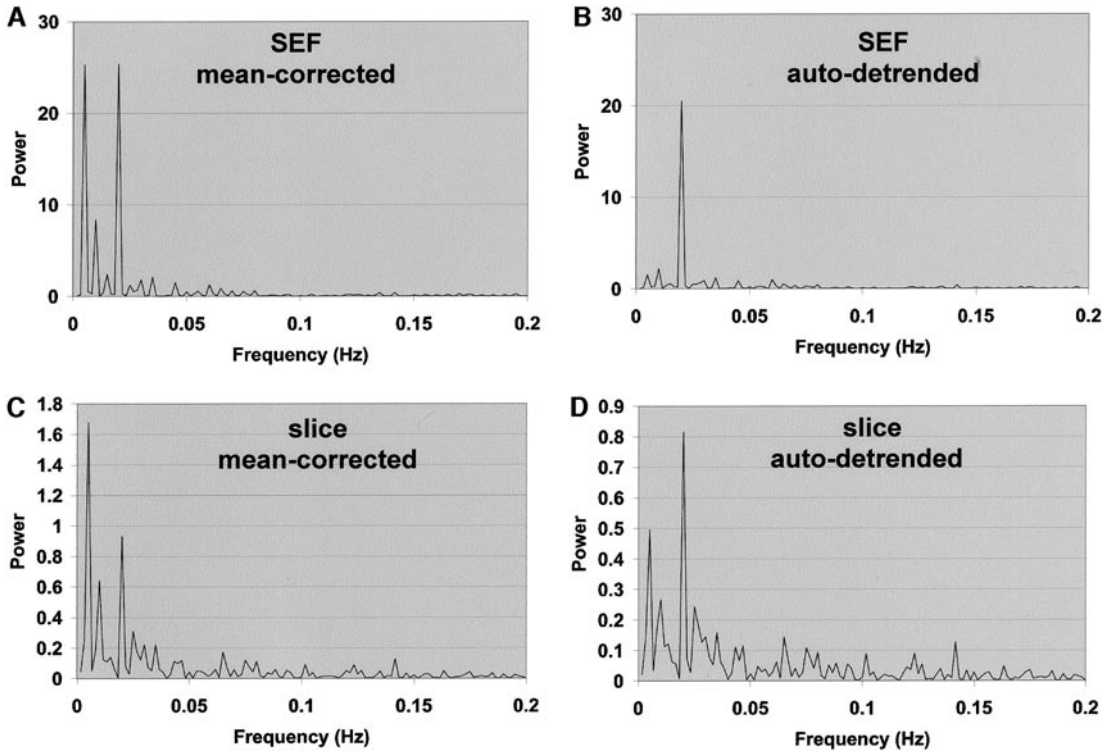

FIG. 4. Power spectra comparing mean-corrected (left) to auto-detrended (right). Top row, supplementary eye field (SEF); bottom row, slice containing the SEF. Note that only frequencies at or below task. (0.02 Hz) were dampened for both the ROI and a slice containing the ROI.

vivo data and for a spherical water phantom. Largest drifts (using the spline algorithm) were seen at brain edges, brain–csf interfaces, and the anterior skull base. Comparable relative drift was observed at the edge of the water phantom.

Null dataset. We analyzed the null dataset with mean-correction and auto-detrending enabled. There was no false positive activation in either case.

DISCUSSION

The main finding of this paper was that among several detrending methods commonly used for fMRI data analysis, a spline technique performed best. The improvement in detection of pixel activation exceeded that of linear and quadratic detrending. Using a voxel-based auto-detrending algorithm, we demonstrated that the spline technique resulted in the greatest increase in activated voxels whether the whole brain or a targeted ROI was analyzed. The concept of applying a plurality of analysis methods to fMRI data has been described previously in terms of model selection (Lange *et al.*, 1999) and concensus mapping (a summary image obtained by averaging various models) (Hansen *et al.*, 2001). Lange *et al.* suggest that when there is no reason *a priori* to favor one statistical model over another, the pluralistic analysis strategy could serve to strengthen the concordance between different methods as well as reveal new relationships between them (Lange *et al.*, 1999). Our study was a simplification of the Lange study in that our model was limited to the general linear model, but then extended it by testing several commonly used detrending algorithms to the voxel level.

In the MR literature, spline fitting is a tool used primarily in image registration, motion detection, and fitting NMR lineshapes (Suter and Chen, 2000; Raz *et al.*, 1997), but rarely used as a filter for fMRI time series data. We used a variant of cubic spline that divided the voxel time series into three arbitrary segments and used the mean of each segment as a knot. A small number of segments was more likely to remove low frequency fluctuations while more segments began to track the expected signal waveform. A potential explanation for the high performance of the spline model is that drift can change its shape rapidly in a manner not necessarily conforming to simple polynomial waveforms (Genovese, 1997). Our findings are consistent with those of Skudlarski et al. who found a progressive improvement in efficiency of fMRI preprocessing after applying linear and quadratic corrections, while a cubic drift subtraction did not help further. In fact, in our study, the number of activated voxels *decreased* after cubic and wavelet detrending compared

to a correction for mean signal (Fig. 1). Mattay *et al.* used a cubic polynomial to remove drift in their study of a motor task, but they did not systematically test this against other corrections (Mattay *et al.*, 1996). Wavelet transforms which involve simultaneous data analysis in frequency and time domains have become an important technique in image compression and noise suppression. Denoising of time series data is a more recent application of wavelets (Zaroubi and Goelman, 2000; LaConte *et al.*, 2000; Ngan *et al.*, 2000). Zaroubi and Goelman found a 63% increase in activation area after wavelet denoising was applied to a rat somatosensory stimulation model (Zaroubi and Goelman, 2000). However, we noted that occasionally wavelet detrending dampened the power spectrum peak at the task frequency. Tracking of the activation signal, followed by its removal during detrending, is likely the reason that the wavelet, as well as the cubic method, resulted in fewer activated voxels.

Analysis based on low *P* values must be carefully considered as it may introduce bias through exhaustive search and data exclusion. We tried to minimize bias by calculating the *P* value using degrees of freedom (DOF) specific for the detrending algorithm. At issue is that there is no single optimal index for comparing fMRI activations across techniques or subjects. Previous studies have used an ROC approach (Constable *et al.*, 1995), sum of *t* values (Constable *et al.*, 1998), pixel count (Bookheimer *et al.*, 2000), magnitude of *P* value (Buonocore and Maddock, 1997), and test–retest probability classifications (Genovese *et al.*, 1997). In addition to adjusting for differences in DOF, we also tried to avoid this bias with non parametric spectral analysis. The power spectra for both an ROI and an imaging slice demonstrated that only very low frequency components were removed with auto-detrending. Since the power spectrum was not amplified at the task frequency or above, it is highly doubtful that false-positives were introduced into the data. False-positive activations ought to have increased the amplitude at the task frequency. In fact, Fig. 4 shows that the amplitude at task frequency remained constant or decreased with detrending. Our results area consistent with experimental (Paus, 1996) and *in vivo* functional imaging studies of visual smooth pursuit task in humans (Berman *et al.*, 1999; Petit and Haxby, 1999). Therefore the improvement we observed within the ROIs is a reasonable qualitative indicator of whether this technique is beneficial. Finally, a null dataset did not yield any false-positive activation after auto-detrending was applied.

Another finding was that the greatest drifts were detected at brain edges or brain–CSF interfaces. However, no systematic patterns emerged for one detrending algorithm compared to another (e.g., all the detrending techniques produced comparable drift maps). Our findings are consistent with others who observed low frequency drifts in regions of high spatial intensity gradient changes (Bandettini *et al.*, 1993), specifically at brain edges and interfaces with the CSF (Smith *et al.*, 1999). In addition we found relatively large drifts in the anterior skull base, which is an area of high static field inhomogeneity. This would be consistent with Smith *et al.* who suggested that local magnetic field instabilities contribute to areas of drift (Smith *et al.*, 1999).

A fundamental assumption of the auto-detrending method is that every voxel time-series may be treated individually. Global detrending is clearly more efficient, but there is no reason to believe that low frequency noise is distributed uniformly across voxels. On the contrary, if one assumes that local susceptibility dephasing contributes to the baseline drift, then adjacent voxels may display unrelated patterns of drift (Smith *et al.*, 1999). Further, the assumption is reasonable as long as the task frequency is sufficiently higher than the drift that is being removed (Skudlarski *et al.*, 1999).

In conclusion, this paper compares several linear detrending methods for analyzing fMRI time series data. Auto-detrending reveals new information especially in areas with baseline drift and weak activation, but spline detrending appears to be a good trade-off with respect to computation time. Comparison of these linear algorithms to non linear or frequency domain filters remains a subject for further investigation.

ACKNOWLEDGMENTS

This work was sponsored by the Research and Education fund for the Radiological Society of North American and a Whitaker Foundation Biomedical Engineering Research Grant.

REFERENCES

Bandettini, P. A., Jesmanowicz, A., Wong, E. C., and Hyde, J. S. 1993. Processing strategies for time-course data sets in functional MRI of the human brain. *Magn. Reson. Med.* **30:** 161–173.

Berman, R. A., Colby, C. L., Genovese, C. R., Voyvodic, J. T., Luna, B., Thulborn, K. R., and Sweeney, J. A. 1999. Cortical networks subserving pursuit and saccadic eye movements in humans: An FMRI study. *Hum. Brain Mapp.* **8:** 209–225.

Biswal, B., DeYoe, A. E., and Hyde, J. S. 1996. Reduction of physiological fluctuations in fMRI using digital filters. *Magn Reson. Med.* **35:** 107–113.

Bookheimer, S. Y., Strojwas, M. H., Cohen, M. S., Saunders, A. M., Pericak-Vance, M. A., Mazziotta, J. C., and Small, G. W. 2000. Patterns of brain activation in people at risk for Alzheimer's disease. *N. Engl. J. Med.* **343:** 450–456.

Buonocore, M. H., and Maddock, R. J. 1997. Noise suppression digital filter for functional magnetic resonance imaging based on image reference data. *Magn. Reson. Med.* **38:** 456–469.

Constable, R. T., Skudlarski, P., and Gore, J. C. 1995. An ROC approach for evaluating functional brain MR imaging and postprocessing protocols. *Magn. Reson. Med.* **34:** 57–64.

Constable, R. T., Skudlarski, P., Mencl, E., Pugh, K. R., Fulbright, R. K., Lacadie, C., Shaywitz, S. E., and Shaywitz, B. A. 1998.

Quantifying and comparing region-of-interest activation patterns in functional brain MR imaging: Methodology considerations. *Magn. Reson. Imaging* **16:** 289–300.

Genovese, C. R. 1997. A time-course model for fMRI data. Proc., ISMRM, 5th Annual Meeting. Vancouver, B. C. 1669. Abstract.

Genovese, C. R., Noll, D. C., and Eddy, W. F. 1997. Estimating test-retest reliability in functional MR imaging. I: Statistical methodology. *Magn. Reson. Med.* **38:** 497–507.

Hansen, L. K., Nielsen, F. A., Strother, S. C., and Lange, N. 2001. Consensus inference in neuroimaging. *Neuroimage* **13:** 1212–1218.

Hu, X., Le, T. H., Parrish, T., and Erhard, P. 1995. Retrospective estimation and correction of physiological fluctuation in functional MRI. *Magn. Reson. Med.* **34:** 201–212.

Kiviniemi, V., Jauhiainen, J., Tervonen, O., Paakko, E., Oikarinen, J., Vainionpaa, V., Rantala, H., and Biswal, B. 2000. Slow vasomotor fluctuation in fMRI of anesthetized child brain. *Magn. Reson. Med.* **44:** 373–378.

LaConte, S. M., Ngan, S. C., and Hu, X. 2000. Wavelet transform-based Wiener filtering of event-related fMRI data. *Magn. Reson. Med.* **44:** 746–757.

Lange, N., Strother, S. C., Anderson, J. R., Nielsen, F. A., Holmes, A. P., Kolenda, T., Savoy, R., and Hansen, L. K. 1999. Plurality and resemblance in fMRI data analysis. *Neuroimage* **10:** 282–303.

Le, T. H., and Hu, X. 1996. Retrospective estimation and correction of physiological artifacts in fMRI by direct extraction of physiological activity from MR data. *Magn. Reson. Med.* **35:** 290–298.

Marchini, J. L., and Ripley, B. D. 2000. A new statistical approach to detecting significant activation in functional MRI. *Neuroimage* **12:** 366–380.

Mattay, V. S., Frank, J. A., Santha, A. K., Pekar, J. J., Duyn, J. H., McLaughlin, A. C., and Weinberger, D. R. 1996. Whole-brain functional mapping with isotropic MR imaging. *Radiology* **201:** 399–404.

Meyer, F. G., and McCarthy, G. 2001. Wavelet based estimation of baseline drifts in fMRI. XVIIth International Conference on Information Processing in Medical Imaging, Davis, CA. Abstract.

Mitra, P. P., Ogawa, S., Hu, X., and Ugurbil, K. 1997. The nature of spatiotemporal changes in cerebral hemodynamics as manifested in functional magnetic resonance imaging. *Magn. Reson. Med.* **37:** 511–518.

Ngan, S. C., LaConte, S. M., and Hu, X. 2000. Temporal filtering of event-related fMRI data using cross-validation. *Neuroimage* **11:** 797–804.

Nowak, R. D. 1999. Wavelet-based Rician noise removal for magnetic resonance imaging. *IEEE Trans. Image Process.* **8:** 1408–1419.

Paus, T. 1996. Location and function of the human frontal eye-field: A selective review. *Neuropsychologia* **34:** 475–483.

Petit, L. and Haxby, J. V. 1999. Functional anatomy of pursuit eye movements in humans as revealed by fMRI. *J. Neurophysiol.* **82:** 463–471.

Radant, A. D. and Hommer, D. W. 1992. A quantitative analysis of saccades and smooth pursuit during visual pursuit tracking. A comparison of schizophrenics with normals and substance abusing controls. *Schizophren. Res.* **6:** 225–235.

Raz, J., Fernandez, E. J., and Gillespie, J. 1997. Modeling NMR lineshapes using logspline density functions. *J. Magn. Reson.* **127:** 173–183.

Skudlarski, P., Constable, R. T., and Gore, J. C. 1999. ROC analysis of statistical methods used in functional MRI: Individual subjects. *Neuroimage* **9:** 311–329.

Smith, A. M., Lewis, B. K., Ruttimann, U. E., Ye, F. Q., Sinnwell, T. M., Yang, Y., Duyn, J. H., and Frank, J. A. 1999. Investigation of low frequency drift in fMRI signal. *Neuroimage* **9:** 526–533.

Suter, D. and Chen, F. 2000. Left ventricular motion reconstruction based on elastic vector splines. *IEEE Trans. Med. Imag.* **19:** 295–305.

Tegeler, C., Strother, S. C., Anderson, J. R., and Kim, S. G. 1999. Reproducibility of BOLD-based functional MRI obtained at 4 *T. Hum. Brain Mapp.* **7:** 267–283.

Wood, J. C. and Johnson, K. M. 1999. Wavelet packet denoising of magnetic resonance images: Importance of Rician noise at low SNR. *Magn. Reson. Med.* **41:** 631–635.

Zaroubi, S. and Goelman, G. 2000. Complex denoising of MR data via wavelet analysis: Application for functional MRI. *Magn. Reson. Imaging* **18:** 59–68.