Statistical Design and Analysis of Data from Functional Magnetic Resonance Imaging Experiments: Annotated Bibliography

Darcie Delzell, Richard F. Gunst, William R. Schucany
Department of Statistical Science
Southern Methodist University
Dallas, TX 75275-0332

April 10, 2008

1959

Box, G.E.P. and Lucas, H.L. (1959). "Design of Experiments in Non-linear Situations," *Biometrika*, 46, 77-90.

- Seminal paper on the statistical optimality of D-optimal designs for estimating parameters in nonlinear models; applicable to fMRI designs when the goal is to fit the 6-parameter two-gamma models for the functional form of the HRF

1971

Box, M. (1971). "Bias in Nonlinear Estimation," *Journal of the Royal Statistical Society*, Series B, 33, 171-201.

- Difficulties in calculating exact bias for nonlinear estimation are illustrated
- Method for calculating the exact bias in a class of nonlinear models is given
- Applies to Bayesian estimation with uniform priors

1972

Wynn, H.P. (1972). "Results in the Theory and Construction of D-Optimum Experimental Designs," *Journal of the Royal Statistical Society*, Series B, 34, 133-147.

- Provides an algorithm for generating D-optimal statistical designs
- Provides theorems guaranteeing optimality under specified conditions

1975

St. John, R.C. and Draper, N.R. (1975), "D-Optimality for Regression Designs: A Review,"*Technometrics*, 17, 15-23.

- Discusses and compares various design criteria. Provides an algorithm for finding D-optimal designs.

1994

Friston, K.J., Jezzard, P., Turner, R. (1994), "The Analysis of Functional MRI Time-Series," *Human Brain Mapping*, 2, 69-78.

- Characterizes the form of the HRF using least-squares deconvolution and a linear time invariant (LTI) model. Assumes a Poisson HRF, and uses the data to estimate the one parameter. They then use cross-correlation between the data and the convolved HRF to determine activation.
- In the framework of statistical parametric mapping.
- No colored noise is accounted for.

1995

Albert, P.S. and McShane, L.M. (1995). "A Generalized Estimating Equations Approach for Spatially Correlated Binary Data: Applications to the Analysis of Neuroimaging Data," *Biometrics*, 51, 627-638.

Friston, K.J. Frith, C.D. Frankowiak, R.S.J., and Turner, R. (1995a). "Characterizing Dynamic Brain Responses with fMRI: A Multivariate Approach." *NeuroImage*, 2, 166-172.

- Argues that fMRI data should be taken rapidly, second by second since the shape and duration of the HRF changes rapidly in a few seconds.
- Possible sources of nuisance trends (physiological biorhythms and stimulus correlated motion). Trends are usually approximated with low-order polynomials and long-period sinusoids.
- Uses MANCOVA to model all voxels simultaneously, getting one p-value. Implicitly accounts for spatial correlations without making any assumptions of form. Uses canonical variates to characterize the HRF without specifying a form.
- Scans are divided into epochs that correspond to a particular task or condition. A single multivariate response variable contains all the voxels in one scan.
- Low frequency artifacts and global effects removed by linear regression, data centered before eigenvectors and eigenvalues extracted. X* contains columns for corrected voxel values and rows are individual scans.
- Volumes (consisting of 10 transverse sections) acquired every 3s. 120 volumes acquired. Three conditions in blocks of 10 scans, each condition repeated 4 times, each time constituted a 30s epoch. Confounds are global activity and sine/cosine functions up to a max of 2.5 cycles/120 scans. Voxel values > 0.8 volume mean used to restrict to intracranial regions.
- An early approach using GLM with Fourier basis functions to model the time series and using canonical variates analysis to describe important features of the model. First introduction of basis functions to model "evoked responses in fMRI" (Friston et al. 1999). Advantage: can model voxel-specific forms of HRFs.
- The author mentions possible sources of nuisance trends (physiological biorhythms and stimulus correlated motion). Trends are usually approximated with low-order polynomials and long-period sinusoids.
- Fig 4 shows an immediate HRF response to a fixed schedule of stimuli and a delayed, with initial undershoot, HRF response to a random schedule of stimuli. Fig. 5, for another individual, shows a longer sustained HRF for the random stimuli but activations for both fixed and random stimuli begin immediately.

Friston, K.J., Holmes, A.P., Poline, J-B., Grasby, P.J., Williams, S.C.R., Frankowiak, R.S.J., and Turner, R. (1995b). "Analysis of fMRI Time-Series Revisited," *NeuroImage*, 2, 45-53.

- These authors consider a general linear model with convolution of an estimated HRF, and other effects. In the example they use a Gaussian kernel.
- Autocorrelations are considered almost completely accounted for by the convolution.

- Hypothesis testing is done with "effective degrees of freedom".

Worsley, K.J. and Friston, K.J. (1995). "Analysis of fMRI Time-Series Revisited – Again.," *NeuroImage,* 2, 173-181.

- Provides exact, correct linear model results to replace an earlier paper (Friston et al. 1995b) where heuristic arguments were given. Uses least squares on smoothed (i.e., convolved) responses represented by Kx. Corrected variances and test statistics. Cites well known generalized least squares results (e.g., Seber 1977) for effective degrees of freedom (usual trace functions).
- Did not apply correct generalized least squares estimation but argued by several means that the least squares results are unbiased (true) and the loss of efficiency is more than offset by the gain in robustness because generalized least squares requires the correct specification of K^{-1}.

1996

Boynton, G.M. Engel, S.A., Glover, G.H., and Heeger, D.J. (1996). "Linear Systems Analysis of Functional Resonance Imaging in Human V1", *The Journal of Neuroscience,* 16, 4207-4221.

- Experiments designed to evaluate the linear transform model of fMRI responses using visual stimulation, evaluates whether stimulus timing and contrast (visual) were separable, whether response to long stimuli could be predicted form responses to short stimuli, and whether noise was independent from stimulus contrast and temporal period.
- Two types of block designs (note: not event-related, shortest block = 3 sec.)
  - Periodic: alternating checkerboard and gray panels, moving left, with periods of 10, 15, 30, and 45 sec.; total stimulus duration 192 sec., first 12 sec. discarded; 120 images in record of 180 sec. (TR = 1.5 sec.)
  - Pulse: Full field checkerboard of duration 3, 6, 12, and 24 sec. followed by gray field for 24 sec.; 6 cycles repeated for the full scan – total duration depending on the pulse duration.
- fMRI response:
  - Periodic: amplitude of the sinusoid that best fits the series: amplitude and phase taken from the discrete FFT of the time series *at the stimulus temporal period*. Averaged across all voxels in the calcarine sulcus.
  - Pulse: average of the 6 blocks of stimuli times, averaged across all voxels in the calcarine sulcus.
- Confirms separability of contrast stimulus and stimulus timing.
- Confirmed linearity of pulse timing: response to stimulus of s sec. is the sum of s consecutive responses (each delayed by the stimulus timing) for 1 sec.
- A response of contrast magnitude *c* is *c* times the response to a contrast of magnitude 1.
- Uses a gamma HRF and a power curve for the contrast magnitude; models the stimulus-evoked fMRI response as a product of the contrast function and the HRF convolved with the stimulus.
- Reports that Menon et al.(1995) found an initial decrease in signal (attributed to initial deoxygenation) then increase due to oxygenated hemoglobin.

Buckner, R. L., Bandettini, P.A., O'Craven, K.M., Savoy, R.L., Petersen, S.E., Raichle, M.E, and Rosen, B.R. (1996). "Detection of Cortical Activation During Averaged Single Trials of a Cognitive Task Using Functional Magnetic Resonance Imaging," *Proceedings of the National Academy of Sciences,* 93, 14878-14883.

- Authors investigate single trial paradigms and their ability to show activation. A block trial paradigm was used for comparison.
- They averaged the signal across runs, subjects, and blocks (or single trials).

Bullmore, E., Brammer, M., Williams, S.C.R., Rabe-Hesketh, S., Janot, N., David, A., Mellers, J., Howard, R., and Sham, P. (1996). "Statistical Methods of Estimation and Inference for Functional MR Image Analysis," *Brain Activation Mapping for fMR*I, 34, 261-277.

- Fits linear and sinusoidal (at the fundamental frequency of the boxcar stimulus function + first two harmonics) regression model terms to data from a block design with 30 sec. activation and 30 sec. rest over 5 complete cycles. Then estimates an AR(1) error term and refits using (estimated) generalized least squares and arima.mle from S-Plus. Fits are very close. Demonstrates the biased (too small) standard errors that OLS gives if the AR term is ignored.
- Uses randomization techniques to identify activated voxels.

1997

Dale, A. M. and Buckner, R. L. (1997). "Selective Averaging of Rapidly Presented Individual Trials Using fMRI," Human Brain Mapping, 5, 329-340.

- Major impediment to block trials: do not allow separate trials within task blocks to be distinguished. Must accept many repetitions of one trial type. Advocates single-trial or event-related designs.
- Goal: establish that hemodynamic response to rapidly presented isolated trials adds in a linear fashion and individual trials can serve as the basis for mixed trial designs.
- Compared single-stimulus fMRI responses with two- and three-stimulus responses. Two stimuli 5 sec. apart and three stimuli 2 sec. apart were examined. Concluded that responses were additive and that the individual stimuli responses could be obtained by subtraction.
- Showed that visual stimuli localized to one hemisphere could be detected using intermixed trial paradigms.

Holmes, A. P., Josephs, O., Buchel, C., and Friston, K.J. (1997). "Statistical Modeling of Low-Frequency Confounds in fMRI", *NeuroImage,* 5, S480. Poster.

- Use cosine functions with 7 periods $k\pi(s+24)/84$, where s is the scan number (1 – 84). Fastest period (7) has 3 ½ cycles; whereas the boxcar stimulus has 7 cycles. Long-periods (7) in a regression model with 7 columns acts as a high-pass filter (removes those cycles from the fit). Claims to remove aliased cardiac and respiratory effects. No proof that's what they are.
- Darcie: Also, it makes no sense to me to model cardiac and respiratory effects with LONG period functions. They should be faster.

Josephs, O., Turner, R., Friston, K. (1997). "Event-Related fMRI," *Human Brain Mapping*, 5, 243-248.

- Model a convolved HRF with an equally spaced event-related ISI using basis functions for the unconvolved HRF.
- Assume that each HRF is zero when a new stimulus is given – there is no overlap of HRF functions.
- Use Fourier series basis functions to create a GLM, different parameters for each voxel. Says that these sinusoidal basis functions account for phase shift in acquisition and that with other basis functions derivatives should be added to the model. This renders non-Fourier models less sensitive to artifactual phase differences." Inferences are then used to create a statistical parametric map.
- Assume a constant ISI and choose it in such a way as to sample the HRF at a higher frequency than the TR; states that block designs can be confounded with psychological factors such as attention span and habituation.
- Seem to be the first use of an SPM{F} map.

Lange, N. and Zeger, S.L. (1997). "Non-linear Fourier Time Series Analysis for Human Brain Mapping by Functional Magnetic Resonance Imaging," *Applied Statistics*, 46, 1-29.

- Use a gamma function with spatially varying shape and scale parameters to model fMRI data. No undershoot in their model. DFT estimation equations.

Zarahn, E., Aguirre, G.K., and D'Esposito, M.D. (1997). "Empirical Analyses of BOLD fMRI Statistics," *NeuroImage*, 5, 179-197.

- Observations made that suggest fMRI data might be autocorrelated (Friston et al. 1994, Boynton et al. 1996). Previously suggested that intrinsic autocorrelation might be negligible relative to autocorrelations introduced after smoothing (i.e., fitting linear models; Friston et al. 1995b).
- Temporal autocorrelations and spatial coherency (cross-correlations of voxel time series) were studied in human 'noise' datasets (no stimulus presentation) and also with water phantoms. Defined "global signal" as the average brain voxel time series: spatial coherence could be due to a stationary, continuously differentiable spatial autocovariance function or to other types of spatial smoothness. Included spatial global signal covariates to determine if the global signal was physiological or affected by the type of brain matter.
- Water Phantom: Experiment 1 assessed whether computers used for stimulus presentation contributed noise to fMRI time series and whether there were differences in computer noise. Experiment 2 assessed whether LCD panels contributed.
- Correlations were present in the water phantom, implying they were not physiological processes.
- 1/f and exponential models were considered in fitting the periodogram of the data for the 17 human "noise" subjects. See the Wiener-Khinchin Theorem and DFT inverse to understand the relationship between the power spectrum of the signal and autocorrelation. 1/f models fit the data but there was also a white-noise

component. The first 13 subjects had less noise than the last 4, suggesting that over the months of data-gathering additional noise from the computers affected the last 4 subject measurements.

- One subject compared to the phantom showed that the 1/f error component is not due to physiological processes.
- Even though the data were noise data, GLM results (with Fourier basis functions) were analysis to examine the effect of intrinsic autocorrelation on test statistics. Interestingly, activation was determined using R>.2.
- Overall Conclusions:
  - temporal autocorrelations for both human noise data and water phantom data are 1/f frequency error processes with white noise
  - the 1/f error component could not be explained by motion, equipment, or convolution of neuronal activity
  - when a GLM model was used to model intrinsic autocorrelations, the empirical model proved to be invalid, suggesting a spatially nonstationary temporal autocorrelation structure
  - spatial coherency was demonstrated, with greater coherence at lower frequencies; could not be explained by a continuously differentiable autocorrelation function

1998

Aguirre, G.K., Zarahan, E., and D'Esposito, M. (1998). "The Variability of Human, BOLD Hemodynamic Responses," *NeuroImage*, 8, 360-369.

- Previous work, going back to Boynton et al. (1996), used a single gamma model for all subjects; more recent work suggests individuals might vary in the parameterization of the HRF. This paper investigates this hypothesis.
- Event-related design, with ISI = 16 sec. Averages signal for each subject over the epochs after the stimulus applications. Previous studies (e.g., Dale and Buckner 1997) indicated return to baseline after 16 sec. TR = 2 sec gave 160 images per slice across a 320 sec (20 trials) scan.
- GLM using a Fourier basis set of 3 sines and 3 cosines at .0625 (16 sec/cycle or 1/16 cycle/sec), .125, .1875 Hz (Josephs et al. 1997). These 6 covariates + adjustment for Nyquist and testing frequencies provide a complete basis set for the 8 time points per activation.
- Significant subject variability, much less within-subject variability over different scans.

Burock, M, Buckner, R.L., Woldorff, M.G., Rosen, B.R. and Dale, A.M. (1998). "Randomized Event-related Experimental Designs Allow for Extremely Rapid Presentation Rates Using Functional fMRI," *NeuroReport*, 9, 3735-3739.

- Open question: how rapidly can individual trials be presented in event-related designs and still provide a powerful procedure for fMRI brain mapping.
- Difficulty is the delay in HRF response and the evolution of the HRF over 10-12 sec. On the other hand, longer ISIs restricts the choice of experimental paradigms and limits the number of trials available to detect signal changes.

- Solution: Randomized experimental designs allow one to overcome the overlap problem.
- Design: 250 ms flickering checkerboard (left hemi-field, right hemi-field, fixation), 20s time series after each stimulus; stimuli 16, 3, 1 s for fixed interval experiment; and random (yes, no) application with mean ISIs the same as for fixed length ISIs for the randomized interval experiment.
- Compares fixed ISI designs (1, 3, 16 sec) with randomly chosen ISIs that have the same mean ISI as the fixed designs; estimates the HRF function using non-parametric methods (selective averaging).
    - Fixed interval ISI designs deteriorate to constant HRFs and the HRF cannot be estimated as the interval decreases
    - Random interval ISIs have increased variability but do allow the HRF to be estimated for shorter durations, even an average ISI = 1 sec.
- The estimated HRF is compared to that empirically estimated by a 1 s checkerboard.
- Concludes that for rapid presentation rates, randomly presentated stimuli are better for HRF estimation than one fixed ISI because there is more variance in the time series. For a fixed ISI of 3 or 1, there is hardly variation in the time series.
- References a Masters Thesis and states that "in a separate study investigating the effect of such non-linearities on the hemodynamic response estimates using rapid presentation, randomized event-related designs and analysis, we found that the estimates were largely insensitive to the kinds of long time-scale non-linearities that have been observed."

Friston, K.J., Josephs, O., Rees, G., Holmes, A., Rugg, M.D., and Turner, R. (1998a). "Event-Related fMRI: Characterizing Differential Responses," *NeuroImage*, 7, 30-40.

- Models the HRF as Fourier basis functions or a two-gamma function (can test for differences in magnitude) and its derivative (can test for differences in latency). Includes them in a regression model with confounding effects as covariates.
- Uses GLM to calculate standard errors for any time *t* using the basis functions evaluated at *t.*
- Equally spaced ISI (16 sec.).
- Suggests an SPM{F} map for the combined effects of two different stimuli, and an SPM{t} map to contrast the stimuli.
- Acknowledge that nonlinear components occur in HRFs (Friston et al. 1998, see below), but that this occurs when ISIs are too small. With reasonably spaced ISIs, the nonlinearities can be discounted.

Friston, K.J., Fletcher, P., Josephs, O., Rees, G., Turner, R. (1998b), "Nonlinear Event-Related Responses in fMRI," *Magnetic Resonance in Medicine*, 39, 41-52.

- Examine the significance of nonlinear aspects of the HRF.
- Showed 'interactions' between stimuli spaced closely together (1-2s apart), resulting in reduced responsiveness at very small ISIs.
- Uses a Volterra series to model the times series nonlinearly. Still uses a prescribed 2-gamma HRF that was estimated from a previous experiment in one voxel.

- Comments that temporal HRF derivatives were used to accommodate multislice acquisition. Adding or subtracting the temporal derivative shifts the basis functions backwards or forwards in time.
- Their experiment was done in the left temporal superior gyrus with a passive word experiment, similar to ours.
- Notes a slight pre-undershoot.
- "suggests that if neuronal activity has been high in the past few seconds, then the hemodynamic response will be suppressed."
- The HRF deviates from a linear relationship at higher event frequencies.

McKeown, M.J, Makeig, S., Brown, G.G., Jung, Y-P, Kinderman, S.S., Bell, A.J., and Sejnowski, T.J. (1998). "Analysis of fMRI Data by Blind Separation Into Independent Spatial Components," *Human Brain Mapping*, 6, 160-188.

- Correlation techniques create a reference function by convolving the block design with a fixed HRF function. Those voxels whose signals correlate with the reference function above a threshold are designated as areas of activation.
- Since task-related signal changes are typically small (<10%), other "component processes" having separate time courses and spatial extent produce the bulk of the signals.
- ANOVA methods are based on the tenuous assumptions that 1) observations have known (e.g., Gaussian) distributions, 2) variances and covariances between repeated measurements are equal, 3) time courses of factors affecting the variances of the fMRI signal can be reliably estimated in advance, and 4) signals at different voxels are independent. None of the current modeling methods attempt to extract the intrinsic structure of the data.
- Current methods typically require grouping or averaging data over several task/control blocks. This reduced the ability to detect transient changes due to changes in strategy by the subject, learning or habituation of task performance, fatigue, etc.
- Principal components analysis (PCA) captures orthogonal spatial patterns or eigenimages that exhibit the greatest variability between pairs of voxels. However, if task-related fMRI changes are only a small portion of the total signal variance, retaining only those orthogonal eigenimages capturing the greatest variance in the data might reveal little information about task-related activations or other processes of interest. If voxels become simultaneously activated, methods based on voxel-pair associations might not capture the overall pattern of activation.
- Brain function is based on two principles. Localization implies that each psychomotor function is performed principally in a small set of brain areas. Connectionism posits that the brain regions involved with a given psychomotor function may be widely distributed and require functional integration of activity in multiple loci or distinct brain systems.
- The authors posit that brain areas activated by performance of a psychomotor task should be unrelated to brain areas whose signals are affected by artifacts such as head movements, machine noise, etc. – spatially-independent components.

- ICA models independent component time series of activities (activation, head movement, etc.) and a voxel's signal is a linear sum of the independent components. No a-priori assumptions need be made about the time courses of activation of the different components or the cause of any of the activations.
- The ICA algorithm is an iterative unsupervised neural network learning algorithm based in information-theoretic principles. The ICA algorithm determines an unmixing matrix W from which the component maps and time courses of activation can be computed. $C_{ij} = \sum_{k=1}^{N} W_{ik} X_{kj}$, where C_{ij} is the value of the jth voxel in the ith component map, X_{kj} is the kth time point in the jth voxel, and N is the number of time points. Equivalently, C = WX. The data can be reconstructed from $X^* = W^{-1}C$, where each column of W^{-1} is a time course of one of the components.
- The strength of this ICA approach is that it is completely nonparametric. The weakness is what does one have when one is finished? The independent component time courses still have to be identified. They still have to be shown to be consistent across groups of voxels. In the papers, either correlations with the stimulus vectors or various percentages of agreement are calculated. To the extent that high z values, correlations or percentages of agreement are achieved, the information is useful.
- Seriously deficient in the approach is any ability to properly use spatial correlations. That is actually cited as a positive because contiguous activations due to stimuli might not be in contiguous groups of clusters. On the other hand, we have already seen the increase in power that is obtainable using contiguous blocks of spatially correlated signals. We ultimately -- if we are to perform whole brain analyses -- need to be able to identify noncontiguous groups of activated clusters. Perhaps ICA gives us a starting point.

Rosen, B. R., Buckner, R. L., and Dale, A. M. (1998). "Event-Related Functional MRI: Past, Present, and Future," *Proceedings of the National Academy of Sciences,* 95, 773-780.

- Survey of the history of fMRI studies.
- Cerebral blood flow increases within 2 sec. of stimulus, peaking approximately 5-7 sec. after stimulus.
- Details the use of event-related paradigms, especially studies proving the efficacy of event-related designs to detect changes in signal. Cited
    - Savoy et al. (1995) for demonstrating a detectable event-related signal from a stimulus of only 34 msec.
    - Boynton et al. (1996) for LTI responses.
    - Dale and Buckner (1997) for mixed trial paradigms with short ISIs.
    - Buckner et al. (1996) as evidence that high-level cognitive functioning could be detected by event-related fMRI.
    - Courtney et al, (1997) for intermixed trials that separated encoding of stimuli from maintaining the stimulus in working memory; analysis identified brain areas activated during both encoding and maintenance but showed differential participation in the two kinds of processes. See

also Cohen et al. (1997). Both separated prefrontal areas from posterior ones.

- Cautioned about nonlinearities in rapid ISI paradigms: unknown whether the HRF itself becomes nonlinear or the additivity breaks down.

Vazquez, A. L. and Noll, D.C. (1998). "Nonlinear Aspects of the BOLD Response in Functional MRI," *NeuroImage,* 7, 108-118.

- Linear time-invariant HRF satisfy scaling (amplitude of the input increased by a factor of c, amplitude of the output similarly affected) and superposition (2 1 sec. stimuli result in additive HRFs, with the second one beginning 1 sec. after the first one) conditions.
- Visual stimuli of 1, 2, 4, and 8 sec. (note: block design); stimulus contrast 10, 2, 40, and 80% at 4 sec. 10 trials per stimulus duration or contrast.
- The hemodynamic responses did show evidence of nonlinearity for stimulus durations < 4 sec and for contrasts less than 40%:
  - Short ISIs result in responses that are larger in amplitude and shorter in duration than predicted by a linear time-invariant (LTI) model; Boynton et al. (1996) reported similar results for the short duration (3 sec.) stimuli.
  - This suggests that a short, intense visual stimulus could give similar results to a longer, weaker stimulus.
  - Duration manipulation study: predicted amplitudes greater than observed.

1999

Dale, A.(1999). "Optimal Experimental Design for Event-related fMRI," *Human Brain Mapping*, 8, 109-114.

- Randomized ISI designs have better efficiency thatn fixed ISI designs for event-related designs.
- Widely reported that using ISIs of at least 15 s is optimal, shorter ISIs result in serious lack of power. Cites Cox and Bandettini 1998, Hutton et al. 1998. Other studies report success with ISIs as short as seconds. Cites Buckner et al. 1998, Burock et al. 1998, Clark et al 1998, Dele and Buckner 1997, Wagner et al. 1998.
- Explanation: Accuracy or efficiency is not determined by the mean ISI but by the entire distribution of ISIs.
- Authors use GLM methods with an empirically defined HRFs and selective averaging techniques to estimate the HRF. Efficiency of the estimates is examined for different mean ISIs; fixed vs variable.
- For a fixed design, it's more efficient to have long ISIs. For variable designs, the smaller mean ISI, the better efficiency. This is assuming LTI.
  - If ISI is jittered, efficiency improves monotonically with decreasing mean ISI
  - If ISI is kept stable, efficiency decreases with decreasing mean ISI
- Recommends generating a lot of stimulus vectors with the same mean ISI and picking the one with the best efficiency.

Eddy, W.F., Fitzgerald, M. Genovese, C., Lazar, N., Mockus, A., and Welling, J.(1999). "The Challenge of Functional Magnetic Resonance Imaging," *Journal of Computational and Graphical Statistics*, 8, 545-558.

- Survey of statistical and computational issues relevant to fMRI data modeling and analysis. Comments:
  - More oxygen in hemoglobin, smaller the magnetic field generated by the iron in the blood, thus less interference in the local magnetic field
  - Change in signal due to activity (1%) is less than the noise (2%)

Friston, K. J., Zarahan, Z., Josephs, O., Henson, R.N.A., and Dale, A.M. (1999). "Stochastic Designs in Event-Related fMRI," *NeuroImage,* 10, 607-619.

- Rapid stimulus onset asynchrony (SOA, ISI) allows for maintenance of a particular cognitive or attentional set, decreases the latitude subjects have for engaging alternative strategies or incidental processing, and allows the integration of event-related paradigms. Random SOAs ensure that preparatory or anticipatory factors do not confound event-relates responses and ensure a uniform presentation of stimuli.
- References:
  - Very short SOAs (Dale and Buckner 1997; Clark et al. 1998; Burock et al. 1998). SOAs of 1 sec. or less are commonplace.
  - Relatively long SOAs (Friston et al. 1998b). SOAs of several seconds or more.
- Concerned that rapid, fixed SOAs render the form of the HRF flat due to the extended time course of the HRF.
- Efficiency of estimation is inversely related to the inverse of the covariance matrix of the estimators (uses GLM), which is a function of and only of the design matrix and the error variance. Interprets X to be a convolution of the stimulus vector and the HRF, the latter characterized by a small set of basis functions.
- Wants to compare efficiency of short-and long-duration SOAs in multiple trial/event types. Notes that efficiency might depend on the type of response: individual event-related responses c' = (1,0) or contrasts of two responses c' = (-1, 1).
- Considered efficiency of both deterministic and stochastic methods of generating a stimulus vector assuming a GLM model with fixed HRF.
  - Gives an example (Fig. 1) for which a fixed-interval SOI is the least efficient and a block design in most efficient. Slowly varying (slow modulations, Fig. 1, dynamic stochastic slow) stimulus probabilities.
  - Experience is that there is always some nonstationary stochastic design that is substantially more sensitive (psychological benefits) than the equivalent stationary (probability of stimulus is constant over time) design.
- Considered nonlinearity issues for a minimum ISI.
  - LTI systems: Friston et al. 1994; Boynton et al. 1997
  - Nonlinear effects may predominate at very short SOAs: Vasquez and Noll 1998
  - Nonlinearities become important at < 2 sec SOAs (Fig. 2)

- Fig. 3 shows that the most efficient probability of occurrence was P = 0.5 and corresponded to a trial onset asynchrony (TOA) of 2 x SOA min = 2 sec.
- With more than one trial type very short TOAs are appropriate when comparisons are appropriate; for a 2 trial types but interest is in estimating one HRF, longer TOAs are needed (~ 16 sec.) when P = 0.5 and are ensured by null events in the design. This can be achieved by using P ~ 0.3, giving a mean TOA of ~ 3 sec.
- By making the probability of null events and all other events = 1/(N+1), where N = number of events, a mean TOA of (N+1) x SOA_{min} is optimal.

Glover, G.H. (1998). "Deconvolution of Impulse Response in Event-Related BOLD fMRI," NeuroImage, 9, 416-429.

- Cites undershoot (Kruger et al. 1996) and delay (Hu et al. 1997)
- Uses deconvolution to remove the effect of the impulse response from the measurements in order to more accurately depict the time course of the neuronal response. Goals:
  - Experiments to determine whether response to long stimuli could be predicted from measured responses to very short stimuli
  - Experiments to determine if different stimulus rates and episode repetitions would enable the signal to be extracted
  - Block and "repeated trial" designs for visual and auditory stimuli
  - Block trial analysis involved correlations of sine waves with signals after linear trend removal; repeated-trial data used quadratic trend removal, "time-locked" averaging of 30-35 time frames, and averaging of activated voxels in the cortex
  - Linearity assessed by convolving fitted 1 sec. finger tapping responses with rectangular impulse functions of 2, 4, 8, and 16 sec: neither the motor nor the auditory cortex system was linear – undershoot was not well predicted
  - Deconvolution of three 1 sec stimuli Ts sec (ISI) apart is not "well resolved" unless the Ts is 4 or more sec.; i.e, the deconvolution does not show three impulse responses
  - Used two-gamma convolved HRF function with n1 = 6, n2 = 12, t1 = 0.9, t2 = 0.9, and a2 = 0.35, cj = $[\max\{(t^{nj})\exp(-t/tj)\}]^{-1}$
  - Long-duration data looks flawed; 1 sec. data looks as the HRF would suggest

Kershaw, J., Ardekani, B.A., and Kanno. I. (1999). "Application of Bayesian Inference to fMRI Data Analysis, *IEEE Transactions on Medical Imaging*, 18, 1138-1153.

- fMRI data appears to contain confounding or nuisance trends (possibly sinusoidal); the precise nature of these trends is unknown but might include physiological biorhythms and stimulus correlated motion
  - Trends are usually approximated by making an ad-hoc choice for the confounding terms: popular choices are low-order polynomials and long-period sinusoids
  - Structure of the noise not well understood: data may be auto-correlated and models now routinely include first-order autoregressive terms
  - Analysis is unclear
- Goal: apply Bayesian modeling methods
  - No spatial or temporal smoothing used

- 3 models are considered using Bayesian methods
  - GLM with white noise
  - Linear time invariant nonlinear model
  - GLM with AR noise
- Comprehensive Bayesian derivation of posterior distributions and procedures for testing for activated voxels.

Kruggel, F. and von Cramon, D. Yves (1999). "Modeling the Hemodynamic Response in Single-trial Functional MRI Experiments, Magnetic Resonance in Medicine, 42, 787-797.

- HRF: within milliseconds, oxygen consumption elicits shortening of the T_2* time and consequent signal decrease; the active region is oversupplied by an inflow of oxygenated hemoglobin (HbO_2), leading to a signal increase 5-6 sec after stimulus onset; limitations of the vascular regulation and/or HbO_2 excess induce a dispersion of the signal increase in 3-4 sec; final negative dip is not yet interpreted.
- They used nonlinear regression to fit a Gaussian three-parameter HRF model. Modeled errors as an AR(1). Used suboptimal tests for Gaussian errors, Gaussian form for the HRF.

Quinn, B.G. (1999). "A Fast Efficient Technique for the Estimation of Frequency: Interpretation and Generalization," *Biometrika*, 86, 213-220.

- Expands on Quinn and Fernandez (1991) for estimating the frequency of a sinusoid in the presence of noise.

2000

Bandettini, P. A. and Cox, R. W. (2000). "Event-Related fMRI Contrast When Using Constant Interstimulus Interval: Theory and Experiment", *Magnetic Resonance in Medicine*, 43, 540-548.

- Constant ISIs:
  - More trials per unit time increase statistical power but slowness of the hemodynamic response causes signal overlap and possible saturation of the fMRI signal (increase in baseline level and subsequent attenuation of the amplitude change).
- Event-related studies employ stimulus durations (SDs) from 0.33 sec. to 2 sec and ISIs from 2 sec. to 30 sec.
- Noise is assumed stationary and white. *The stimuli are placed far enough apart that the signals do not overlap*.
  - $T_{opt} = 2\mu_2/\mu_1^2$ where μ_j is the jth raw moment of the HRF.
  - Examples: boxcar, tent, single gamma
- Optimal (in terms of minimum variance in the GLM model) ISI and stimulus duration (SD) were determined in the constant ISI case. (SD=2s; ISI=12-14s). The optimal repetition interval is T_{opt} = ISI + SD with ISI = 14 – SD for stimuli of 3 sec. or less; T_{opt} = 8 + 2SD = SD + ISI where ISI = 8 + SD for longer stimuli.
- *Experiments with ISI 8 sec. or less (SD = 2 sec.) show pre-undershoots that are a result of the previous HRF returning to baseline from the post-undershoot.*

  - Theoretical optimal ISI for SD = 2 is 10.3 sec. for a single-gamma HRF; experimentally, the optimal is 12 sec.
  - Comments that randomized ISIs require linearity assumption, preliminary work suggests optimal functional contrast achieved if on-off distribution is 50%-50% with ISI as short as desired within constraint of subject response time.
- Suggests that experiments imply that event-related signal amplitude is greater than predicted when assuming LTI models.
- Gives references for nonlinearity and pre-undershoot.

Buxton, R.B., Liu, T.T., Martinez, A., Frank, L.R., Luh, W.-M., and Wong, E.C. (2000). "Sorting Out Event-related Paradigms in fMRI: the Distinction Between Detecting and Activation and Estimating the Hemodynamic Response," *NeuroImage*, 11, S457.

- Demonstrate the estimation efficiency and detection power of event-related, block, and single-trial periodic designs.
- Conclude that the best compromise for maximizing both is a block with some randomness added, similar to Friston et al. (1999); c.f., Liu et al. (2001).

Friston, K.J., Josephs, O., Zarahn, E., Holmes, A.P., Rouquette, S., and Poline, J.-B. (2000). "To Smooth or Not to Smooth," *NeuroImage*, 12, 196-208.

- Note: Smoothing and Filtering are transformations y ➔Sy in a GLM framework
- Autocorrelation is present in fMRI time series but parametric modeling often misspecifies the structure. Serious bias can result from prewhitening but band-pass filtering, implicitly smoothing, can protect against serious bias.
- Neuronal noise is neurogenic signal not modeled by explanatory variables that occupies the same part of the frequency spectrum as the hemodynamic signal.
  - ISIs should be high frequency in event related designs
  - Physiological and nonphysiological, white and colored, typically low frequency (Holmes et al. 1997) or wide-band.
  - Superposition of these colored components induces error serial correlation
- Analysis requires 3 considerations: optimum experimental design, optimal filtering to obtain efficient parameter estimates, robustness of the statistical inference procedures. High efficiency and robustness require a variance – bias tradeoff that can be controlled by temporal filtering.
- Previous:
  - Worsley and Friston (1995), in order to avoid estimating serial correlations, convolved the data with a Gaussian kernel to impose an approximately known correlation structure – for robustness purposes
  - Bullmore et al. (1996) used the estimated autocorrelation structure to prewhiten the data prior to fitting a GLM with assumed uncorrelated errors – motivated by estimator efficiency. Validity (false positive rate no larger than the significance level of the test) and robustness assured because of the use of a randomization strategy.
  - Locascio et al. (1997) used ARMA models for each voxel, Purdon and Weiskoff (1998) used AR(1) and white noise, Aguirre et al. (1997) and Zarahn et al. (1997) used 1/f error models

- Experimental Design: Experimental effects can only be efficiently estimated if their frequency content survives convolution with the HRF.; i.e., experimental effects should have frequencies that have high power in the spectrum of the HRF. Noise (drift, sinusoids) has lower frequencies – experimental effects should avoid these frequencies, f < 1/64 Hz
    - Block designs desirable for DoX (but not good for other purposes) because they induce low frequency spectra (Note: all of this is under the context of filtering low frequency noise).
- Temporal Filtering: Minimum variance filters induce large bias, preference is on minimum bias filters. Suppressing low and high frequencies (high-pass filter with smoothing) is required to minimize bias. Note: Variance and bias of estimated standard errors.

Friston, K.J., Mechelli, A., Turner, R., and Price, C.J. (2000). "Nonlinear Responses in fMRI: The Balloon Model, Volterra Kernels, and other Hemodynamics," *NeuroImage*, 12, 466-477.

- Volterra basis functions approximate the nonlinear additivity of event-related HRFs. Enables the characterization of the dynamical behavior without having to define and measure state and input interactions in causing the response; however, results in no mechanistic explanation of how the response is mediated.
- Balloon model is a theoretical model of the dynamics of state variables on the response.:
    - Increases in blood flow inflate a venous "balloon" so that deoxygenated blood is diluted and expelled at a greater rate
    - Clearance of deoxygenated blood reduces intravoxel dephasing and engenders an increase in signal
    - Before the balloon has inflated sufficiently, the expulsion and dilution may be insufficient to counteract the increased delivery of deoxygenated blood to the venous compartment and an early dip in signal may occur
    - After the flow has peaked and the balloon has relaxed again reduced clearance an dilution contribute to the poststimulus undershoot
- The paper evaluates the Balloon model in relation to Volterra characteristics and real hemodynamics

Genovese, C.R. (2000). "A Bayesian Time-course Model for Functional Magnetic Resonance Imaging Data," *Journal of the American Statistical Association*, 95, 691-719. (with discussion).

- Comprehensive Bayesian analysis of fMRI data. Discusses the form of the HRF, selection of priors, inferences on trends rather than simple t tests, and computational issues. Discussion by several expert statisticians and neurologists.
- Includes baseline constant, drift component, HRF (multiparameter), and white noise model components
    - Points out that much of any serial correlation is likely filtered out by the fitted drift component
    - Presence of residual autocorrelation likely due to the inability of fitted models to adequately capture epoch-to-epoch variations in the response

Hopfinger, J.B., Buchel, C., Holmes, A.P., and Friston, K.J. (2000). "A Study of Analysis Parameters that Influence the Sensitivity of Event-related fMRI Analyses," *NeuroImage*, 11, 326-333.

- 36 analyses of a single fMRI data set, varying voxel size, spatial smoothing, temporal smoothing, and choice of basis sets. Used 16 brain regions
  - HRF and its temporal derivative was more sensitive to voxel activation than the HRF alone
  - 2 mm^3 resampled voxel size, 10 mm spatial smoothing, and 4 sec temporal smoothing recommended

2001

Ari, N. and Yen, Y. (2001). "Extraction of the Hemodynamic Response in Randomized Event-related Functional MRI," Proceedings of the 23rd EMBS International Conference, Istanbul, Turkey.

- Cites Dale (1999) for greater efficiency of randomized event-related designs over fixed ISI designs.
- Linearity has been demonstrated for ISIs in the 2-15sec range but not always for short ISIs.
- Time-window averaging is unreliable. Cepstral analysis is unstable when noise is added to the signal. The conjugate gradient method worked best when applied in the frequency domain ($Y(f) = S(f)H(f)$).

Bullmore, E., Long, C., Suckling, J., Fadili, J., Calvert, G., Zelaya, F., Carpenter, T.A., and Brammer, M. (2001). "Colored Noise and Computational Inference in Neurophysiologcal (fMRI) Time Series Analysis: Resampling Methods in Time and Wavelet Domains" *Biometrics*, 57, 554-562.

- Discuss pre-whitening and pre-coloring options for validating statistical methods when the errors are 1/f. Introduce wavelet transformations and Resampling in the wavelet domain because of the uncorrelated nature of wavelet coefficients.
  - Pre-whitening is more efficient but may be biased if autocorrelation is misspecified
  - Pre-coloring is less efficient but less to bias provided the coloring matrix is robust enough to impose its predicted form on the residual autocorrelation
  - Time series methods assume stationarity; resampling methods assume exchangeability – wavelet methods achieve this approximately through the whiteningproperty of the discrete wavelet transform
- Point out other studies that suggest colored noise is not only due to HRF-convolved neuronal or instrumental white noise, aliased cardiorespiratory pulsation, uncorrected head movement, and incomplete specification of the design matrix, but also physical effects because of the presence of colored noise in measurements on cadavers and phantoms.

Gossl, C., Auer, D.P., and Fahrmeir, L. (2001). "Bayesian Spatiotemporal Inference in Functional Magnetic Resonance Imaging," *Biometrics*, 57, 554-562.

- Hierarchical Bayes formulation allows temporal and spatial interdependence of voxel time series to be included in the model formulation. Use parametric and semiparametric spatial and spatiotemporal models.

Lazar, N.A., Eddy, W.F., Genovese, C.R., and Welling, J. (2001). "Statistical Issues in fMRI for Brain Imaging," *International Statistical Review*, 69, 105-127.

- Excellent review of the basics of fMRI, Fourier space acquisition, preprocessing, and statistical modeling and analysis.

Liu, T.T., Frank, L.R., Wong, E.C., and Buxton, R.B. (2001). "Detection Power, Estimation Efficiency, and Predictability in Event-related fMRI," *NeuroImage*, 759-773.

- Randomized designs: maximum estimation efficiency, poor detection power (Dale 1999)
- Block Designs: good detection power, minimum estimation efficiency
- Periodic single-trial experiments: poor on both criteria
- ➢ Above conclusions from Buxton et al. (2000)
- Semirandom designs: tradeoff between power and efficiency (Friston et al. 1999), achieve both at the cost of increasing the experiment by less than a factor of 2
- Predictability: ability to circumvent confounds such as habituation and anticipation
    - ➢ Small increases in predictability can offer gains in power with only a minor decrease in efficiency

Naiman, D.Q. and Priebe, C.E. (2001). "Computing Scan Statistic p Values Using Importance Sampling, with Applications to Genetics and Medical Image Analysis," *Journal of Computational and Graphical Statistics*, 10, 296-328.

- Correction to Bonferroni p-values

Ollinger, J.M., Corbetta, M., and Shulman, G.L. (2001). "Separating Processes within a Trial in Event-related Functional MRII. The Method," *NeuroImag*e, 13, 210-217.

- HRF was estimated without assuming a form but required partial trials
- Estimated the mean for each time point after averaging across similar events: 14 equations in 14 unknowns
- Mix compound trials with partial single-event trials

Ollinger, J.M., Corbetta, M., and Shulman, G.L. (2001). "Separating Processes within a Trial in Event-related Functional MRI: II. Analysis," *NeuroImag*e, 13, 218-229.

- BOLD HRF curves can be estimated easily from block designs because the long task and control periods enable square waves to approximate the resulting curve. Modeling as a linear system enables rise time, delay, and fall time to be estimated (Boynton 1996).
- Event related designs make estimating the shape of the BOLD response more difficult and the estimated shape can depend on the experimental paradigm. In a companion paper, the HRF was estimated without assuming a form but required partial trials
- Statistical maps using F statistics from extra sums of squares principle or t statistics from cross correlations

Purdon, P.L., Solo, V., Weisskoff, R.M., and Brown, E.N. (2001). "Locally Regularized Spatiotemporal Modeling and Model Comparison for Functional MRI," NeuroImage, 14, 912-923.

- Model includes a physiologically derived low-frequency noise component.
- Spatiotemporal model, estimation, and model comparison.

Woolrich, M.W., Ripley, B.D., Brady, M., and Smith, S.M. (2001). "Temporal Autocorrelation in Univariate Linear Modeling of fMRI Data, *NeuroImage*, 14, 1370-1386.

- Excellent discussion of tapering as a component of a pre-whitening process for transforming data prior to a GLM; tapering removes edge effects and the influence of high lags in estimating autocorrelations – smooths spectral density estimates
- Used to accommodate 1/f error processes

2002

Aguirre, G.K., Detre, J.A., Zarahn, E., and Alsop, D.C. (2002). "Experimental Design and the Relative Sensitivity of BOLD and Perfusion fMRI," *NeuroImage*, 15, 488-500.

- Arterial Spin Labeling (ASL): noninvasive, quantification of brain tissue perfusion using labeled inflowing arterial protons as endogenous tracers, absolute units (cc blood / 100 g of tissue / min) – change in blood flow itself, as opposed to BOLD
- Subtraction of labeled and unlabeled image pairs
- BOLD fMRI demonstrates greater power at low frequencies, characterized by 1/f modeling: boxcar design (e.g., 60 sec) will have reduced sensitivity because of the presence of greater noise (1/f) at low frequencies
- ASL will differ from BOLD in that the subtraction will difference 1/f noise, leading to data that behave more like they are uncorrelated
    - Experimental designs with concentrated power at low frequencies are feasible; e.g., seeking a slow, continuously changing signal over time

Birn, R.M., Cox, R.W., and Bandettini, P.A. (2002). "Detection versus Estimation in Event-related fMRI: Choosing the Optimal Stimulus Timing," *NeuroImage*, 15, 252-264.

- Estimation and detection are fundamentally different and require different stimulus times
    - Estimating the HRF requires frequent stimulus times that vary between active and control states
    - Activation detection is optimized by block designs
- Maximum detectability or accuracy for a stimulus pattern with a varying ISI occurs when the number of time points in the task and control states are equal
    - Significant decrease in estimation accuracy if the stimulus and control periods are forced to vary on a coarser time scale than the TR (see also Bandettini et al. 2000)

Duan, J-R, Jung, T-P, Kuo, W-J, Yeh, T-C, Makeig, S., Hsei, J-C, and Sejnowski, T.J. (2002). "Single-trial Variability in Event-related BOLD Signals." *NeuroImage*, 15, 823-835.

- Model-based methods assume the form and all parameters of the HRF except amplitudes are known a-priori. They also assume similar HRFs across stimulus or task events, brain areas, stimulus parameters, and subjects.
- Independent components analysis (ICA) decomposes the entire fMRI data set into component activities associated with fixed spatial distributions. Variability is modeled as a sum of deterministic processes with maximally independent spatial distributions. $M = WX$ or $X = W^{-1}M$, where X is the T (times) x V (voxels) BOLD signals, W is an "unmixing" matrix, and M contains spatially independent component activations.
- McKeown et al. (1998) contains details.

Friman, O., Borga, M., Lundberg, P., and Knutsson, H. (2002). "Detection of Neural Activity in fMRI Using Maximum Correlation Modeling," NeuroImage, 15, 386-395.

- Detecting of neuronal activity is achieved by calculating and determining the significance of maximum correlations between a spatially varying (only the scale parameter) HRF and the fMRI time course.
- Local voxel time courses are averaged to increase power.
- Inference is carried out via Monte Carlo simulations

Shen, X., Huang, H-C, and Cressie, N. (2002). "Nonparametric Hypothesis Testing for a Spatial Signal," Journal of the American Statistical Association, 97, 1122-1140.

- Procedure for controlling the false discovery rate using spatial wavelets, based on the premise that large wavelet coefficients tend to cluster spatially and noise wavelet coefficients are approximately uncorrelated
  - Define clusters of neighboring wavelet coefficients using a distance measure for 4 neighboring voxels at the same wavelet scale or adjacent locations for wavelets at different scales – selects b = 11 locations
  - Select L* hypotheses tests corresponding to the L* clusters with the largest cluster-maximum wavelet coefficients; eliminate all remaining hypotheses
  - Apply standard FDR procedures to the maximum coefficients
  - Details for selecting L* are given in Section 3.2

Worsley, K.J., Liao, C.H., Aston, J., Petre, V., Duncan, G.H., Morales, F., and Evans, A.C. (2002). "A General Statistical Analysis for fMRI Data," *NeuroImage*, 15, 1-15.

- Uses estimated autocorrelation parameters, GLM, and the EM algorithm to circumvent estimating small numbers of degrees of freedom

2003

Cressie, N. and Kornak, J. (2003). "Spatial Statistics in the Presence of Location error with an Application to Remote Sensing of the Environment," *Statistical Science*, 18, 436-456.

- Seminal paper on the application of spatial modeling when the locations of the measurements are not exactly known; i.e., they are measured with error.

- Potential for application to the brain, where locations are spatially normalized to accommodate movement and morphed to a common template to accommodate different brain sizes and geometries.

Pfeuffer, J., McCullough, J.C., Van de Moortele, P-F, Ugurbil, K., Hu, X. (2003). "Spatial Dependence of the Nonlinear BOLD Response at Short Stimulus Times," *NeuroImage*, 18, 990-1000.

- Using very short visual presentation times (< 2 sec), deviations from the linear model in the measured BOLD data was found for the response integral, amplitude and width.

Smith, M., Putz, Benno, Auuer, Dorothee, and Fahrmeir, L. (2003). Assessing Brain Activity through Spatial Bayesian Variable Selection. *NeuroImage*, 20, 802-815.

- Predecessor to Smith et al. (2003) in which the prior for activation indicator variables is introduced. Spatial correlation is directly modeled for activation probabilities and indirectly for regression coefficients for covariates. Shows superior edge-preserving properties and fast computing.

2004

Friman, O., Borga, M., Lundberg, P., and Knutsson, H. (2004). "Detection and Detrending in fMRI Data Analysis," *NeuroImage*, 22, 645-655.

- Event-related designs benefit more from pre-whitening than block designs for the detection of activated voxels when colored noise is present.
- Canonical correlation analysis (Friman 2002) is used to remove drift

Kim, H-J. and Boos, D.D. (2004). "Variance Estimation in Spatial regression Using a Non-parametric Semivariogram Based on Residuals," *Scandinavian Journal of Statistics*, 31, 387-401.

- Demonstrate the well-known negative bias in sample semivariograms for large lags when mean parameters have to be estimated. While predictions are known to be little affected, estimated standard errors can be seriously negatively biased.
- Propose a monotonic averaging of neighboring sample semivariogram values when values decrease with increasing lag distance.
- Propose a bias-adjustment to the estimated prediction variance.

Long, C., Brown, E.N., Manoach, D., and Solo, V. (2004). "Spatiotemporal Wavelet Analysis for Functional MRI," *NeuroImage*, 23, 500-516.

- Spatiotemporal wavelet procedure that uses a stimulus-convolved hemodynamic signal plus correlated noise model

Spence, J.S., Carmack, P.S., Gunst, R.F., Schucany, W.R, Woodward, W.A., and Haley, R.W. (2004). "Improved Agreement Between Talairach and MNI Coordinate Spaces in Deep Brain Regions," *NeuroImage*, 22, 367-371.

Talavage, T.M. and Edmister, W.B. (2004). "Nonlinearity of fMRI Responses in Human Auditory Cortex," *Human Brain Mapping*, 22, 216-228.

- Intense noise at acoustic frequencies caused by the scanner causes activity in the auditory cortex that competes with responses to a presented stimulus.
- Nonlinear additivity of BOLD responses is the dominant source of reductions in the measurements of both amplitude and spatial extent of auditory cortex activation.

Woolrich, M.W., Behrens, T.E.J., and Smith S.M. (2004). "Constrained Linear Basis Sets for HRF Modeling Using Variational Bayes," *NeuroImage*, 21, 1748-1761.

- Large portion of subspaces spanned by basis functions produce nonsensical HRF shapes
- Assuming a linear time-invariant system, priors are placed on members of a basis set to give higher probability to components that represent reasonable shapes for the HRF
- Demonstrate far improved detection of voxel activation

2005

Christensen, W.F. and Yetkin, F.Z. (2005). "Spatio-temporal Analysis of Auditory Cortex Activation as Detected with Silent Event Related fMRI," *Statistics in Medicine*, 24, 2539-2556.

- Focus is in scanner noise contributing to the signal responses in the auditory cortex
- Uses a spatiotemporal, measurement-error-free kriging model to spatially smooth the data and increase activation sensitivity

Friman, O. and Westin, C-F. (2005). "Resampling fMRI Time Series," *NeuroImage*, 25, 859-867.

- When determining thresholds for activation, BOLD responses bias temporal autocorrelations, leading to biased thresholds
- Fourier and wavelet Resampling methods may lead to erroneous thresholds
- Resampling based on a pre-whitening transform, driven by an explicit AR(1) error model, fitted by usual methods

Penny, W.D., Trujillo-Barreto, N.J., and Friston, K.J. (2006). "Bayesian fMRI Time Series Analysis with Spatial Priors," *NeuroImage*, 24, 350-362.

- Bayesian priors are placed on regression coefficients in a GLM
- Variational Bayes is used to let the data determine the optimal amount of smoothing

2006

Calhoun, V.D. and Adali, T. (2006). "Unmixing fMRI with Independent Component Analysis," *IEEE Engineering in Medicine and Biology Magazine*, March/April, 79-90.

- Survey of ICA with special reference to fMRI.

Heller, R., Stanley, D., Yekutieli, D., Rubin, N, and Benjamini, Y. (2006). "Cluster-based Analysis of fMRI Data," *NeuroImage*, 33, 599-608.

- Introduces a clustering algorithm for spatial data that produces small clusters using the maximum correlations of neighboring voxels.
- Tests individual voxels in a cluster and then averages the significant voxels in the cluster to form a cluster average
- Cluster tests are the usual SPM tests for clusters

Pavlicova, M., Cressie, N., and Santner, T.J. (2006). "Testing for Activation in Data from fMRI Experiments," *Journal of Data Science*, 4, 275-289.

- Block designs: periods of transition between activation and rest are ignored, two-sample *t*-tests and Wilcoxon rank tests are compared to determine activated voxels; a modified Wilcoxon test performs best

Spence, J.S., Carmack, P.S., Gunst, R.F., Schucany, W.R, Woodward, W.A., and Haley, R.W. (2006). "Using a White Matter Reference to Remove the Dependency of Global Signal on Experimental Conditions in SPECT Analyses," *NeuroImage,* 32, 49-53.

Strother, S.C. (2006). "Evaluating fMRI Preprocessing Pipelines," *IEEE Engineering in Medicine and Biology Magazine,* 25 (2), 27-41.

- Reviews and evaluates the optimization of preprocessing steps for BOLD fMRI
- Provides 147 references

2007

Benjamini, Y. and Heller, R. (2007). "False Discovery Rates for Spatial Signals," *Journal of the American Statistical Association,* 102, 1271-1281.

- Generalizes Benjamini and Hochberg (1995) to clusters of observations. Can be applied to spatial data. Clusters need to be previously defined as in Heller et al. (2006).

Bowman, F.D. (2007). Spatiotemporal Models for Region of Interest Analyses of Functional Neuroimaging Data. *Journal of the American Statistical Association,* 102, 442-453.

- Discusses spatial correlations among spatially distant brain locations and nondecreasing spatial correlations as a function of increasing spatial distances due to connectivity issues between different structures in the brain. Defines a spatial correlation model based on functional similarity, not geographic proximity.
- Very good discussion of spatiotemporal models and estimation. Distance is measured using previous estimates of activity profiles in each voxel.

Smith, M. and Fahrmeir, L. (2007). Spatial Bayesian Variable Selection with Application to Functional Magnetic Resonance Imaging. J*ournal of the American Statistical Association*, 102, 417-431.

- Use Ising priors to smooth spatially indicator variables representing whether a covariate is included in the model, indirectly smoothing the regression coefficients. Spatially smooths activation maps from regression models of blood oxygenation. Uses a single-site sampling scheme to rapidly evaluate posterior activation maps and activation amplitudes. Maps are superior to Bayesian approaches using continuous Markov random fields.

- For baseline trend components of the model, the authors use 1, t, t^2, t^3, $\sin(2\pi t/T)$, $\cos(2\pi t/T$, $\sin(3\pi t/T)$, and $\cos(3\pi t/T)$.

Spence, J.S., Carmack, P.S., Gunst, R.F., Schucany, W.R., Woodward, W.A., and Haley, R.W. (2007). Accounting for Spatial Dependence in the Analysis of SPECT Brain Imaging Data, *Journal of the American Statistical Association*, 102, 464-473.

- Spatially models distance-dependent correlations among voxels in each of several deep brain structures using spatial semivariograms.
- Calculates block and structure averages using weights obtained from fitted spatial semivariograms; standard errors of the averages also calculated.
- Use general linear model (GLM) analyses to compare SPECT measurements of Gulf War syndrome groups with a control group. Determined that there were baseline difference in some structures between syndrome and control groups. Also found treatment differences after administration of physostigmine to the subjects.