A Comparision of fMRI Activation Thresholding Methods.

B. R. Logan[1], D. B. Rowe[2]

[1]Department of Biostatistics, Medical College of Wisconsin, Milwaukee, WI, United States, [2]Department of Biophysics, Medical College of Wisconsin, Milwaukee, WI, United States

Summary:

Single-subject fMRI experiments identify active voxels by performing individual voxelwise tests of the hypothesis that the observed time course is not significantly related to an assigned reference function. The test statistics are thresholded so that those voxels whose test statistic exceeds the threshold are viewed as active. This study describes three error rates which may be used to formally set activation thresholds for fMRI data, reviews methods for determining these thresholding rules including incorporation of spatial correlation information, illustrates the application of these methods on a real dataset, and discusses relative performance of the three procedures in simulations.

Introduction:

Many fMRI experiments have a common objective of identifying active voxels or regions in a neuroimaging dataset. This is done in single subject experiments for example by performing individual voxelwise tests of the hypothesis that the observed time course is not significantly related to an assigned reference function. A voxel activation map is then constructed by applying a thresholding rule to the resulting t-statistics.

Error Rates:

A naive way to specify a thresholding rule is to base it on the significance level of the test, called the voxelwise error rate, which is the probability of incorrectly classifying an inactive voxel as active. However, this tends to result in too many inactive voxels being incorrectly classified as active (false positives). One alternative is to control the imagewise error rate, or the probability of finding at least one false positive. However, a threshold based on this criterion may be too stringent, as researchers may not care about a single incorrectly classified voxel. A third alternative is to control the false discovery rate, which is the expected proportion of truly inactive voxels which are declared as active over the total number of voxels declared active.

Thresholding Methods:

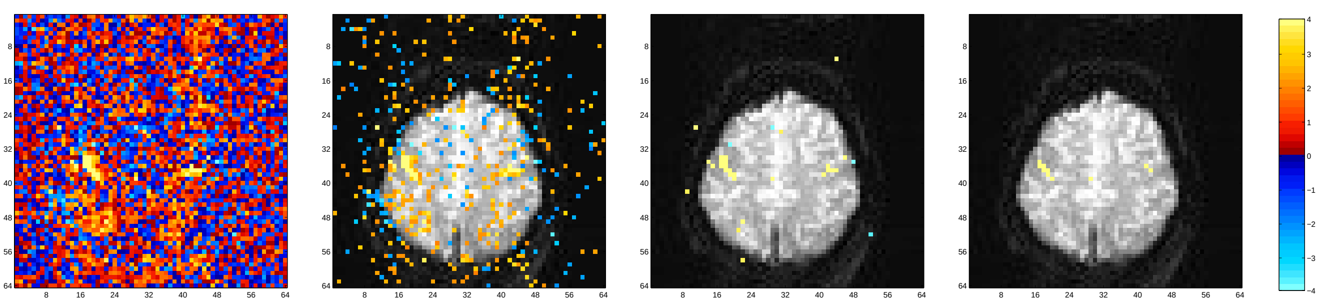
The Bonferroni method may be used to control the imagewise error rate, without any assumptions about the spatial correlation of the data. The procedure of Benjamini and Hochberg (BH, 1995) may be used to control the false discovery rate. This procedure requires that the voxelwise test statistics be independent or positively correlated. However, neither the Bonferroni method nor the BH method explicitly utilize the correlation present in the data to improve the power to detect active voxels while controlling their respective error rates. This is a potential limitation because fMRI experiments often exhibit spatial correlation among the voxel test statistics. Westfall and Young (WY, 1993) propose to incorporate correlation information into the Bonferroni procedure by simulating the distribution of the maximum (voxelwise) t-statistic and using the appropriate percentile of this max t distribution. Nonparametric methods (Holmes et al., 1996; Bullmore et al., 1996; Locascio et al., 1997) may be used to simulate this distribution. Alternatively, random field theory may be used to approximate this distribution (Worsley et al., 1996). For the FDR, Yekutieli and Benjamini (YB, 1999) propose a simulation based extension of the BH method which accounts for correlation in the data. The same nonparametric methods may be applied to estimate the FDR for any given threshold.

Example:

A bilateral finger tapping experiment was performed with a block design of eight epochs of 16s on and 16s off. Each voxel has dimensions in mm of 3.125 X 2.125 x 5, with TE = 27.2ms. Observations were taken every TR = 2000ms so that there are 128 in each voxel. Data from a single axial slice through the motor cortex was selected for analysis. A multiple regression model was fit to the data with an intercept, a time trend, and a reference function.

Results:

The unthresholded t-statistic map is given in the first panel below. Applying a voxelwise error rate threshold of 5% gives the second image, which picks up much of the activation as well as substantial “noise.” The BH method yields the activation region in the third image, while the Bonferroni method results in the fourth thresholded image. The permutation resampling FDR procedure (YB) results in the same thresholded image as the BH FDR procedure, and so is omitted. Likewise, the permutation resampling FWE procedure (WY) results in the same thresholded image as the Bonferroni procedure, and is also omitted.



Simulations:

Data was simulated for a 64×64 image with 128 data points and two 7×7 activation regions. Several covariance structures were considered: 1) A stationary, isotropic exponential covariogram such that each voxel is correlated with every other voxel according to ρ^d, where d = distance between voxels and $\rho = 0.0$, 0.7, and 0.95; 2) a Gaussian random field with FWHM = 5mm; and 3) Correlation structure generated from the residuals of the real dataset above. In each case the average power was computed for the voxels in the ROI’s. In all cases except for the exponential structure with $\rho = 0.95$ (extremely high spatial correlation), the permutation resampling FWE and FDR procedures performed very similarly to the simpler Bonferroni and BH procedures, respectively. This indicates that the simple procedures work well for the amount of spatial correlation or smoothness used in most single-subject fMRI datasets. As in the example above, some additional power can be obtained by using the FDR criterion rather than the FWE criterion, at the cost of some false positives.

References:

Benjamini, Y. and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. JRSS-B, 57: 289-300.

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. Magnetic Resonance in Medicine, 35: 261-277.

Holmes, A.P., Blair, R.C., Watson, J.D.G., and Ford, I. (1996). Nonparametric Analysis of Statistic Images from Functional Mapping Experiments. Journal of Cerebral Blood Flow and Metabolism, 16: 7-22.

Locascio, J.J., Jennings, P.J., Moore, C.I., and Corkin, S. (1997). Time Series Analysis in the Time Domain and Resampling Methods for Studies of Functional Magnetic Resonance Brain Imaging. HBM, 5: 168-193.

Westfall, P.H. and Young, S.S. (1993). Resampling Based Multiple Testing. Wiley, New York.

Worsley, K.J., Marrett, S., Neelin, P., Vandal, A.C., Friston, K.J., and Evans, A.C. (1996). A Unified Statistical Approach for Determining Significant Signals in Images of Cerebral Activation. Human Brain Mapping, 4: 58-73.

Yekutieli, D. and Benjamini, Y. (1999). Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics. JSPI, 82: 171-196.