8

Statistical Analysis for Recovery of Structure and Function from Brain Images

Michelle Yongmei Wang, Chunxiao Zhou and Jing Xia
University of Illinois at Urbana-Champaign
U.S.A.

1. Introduction

Brain imaging has the potential to advance our understanding of human health and to improve diagnosis and treatment of neurological diseases. Inspired by key questions in neuroscience and medicine, it becomes extremely important to develop statistical methods that can accurately and efficiently recover useful quantitative information from large amounts of brain images. The underlying computational issues are challenging and often hampered by uncertainties in imaging acquisition parameters, the variability of human anatomy and physiology, as well as the nature of the imaging data to be handled such as the presence of noise and correlation, and the sample and data sizes, and so on.

Structural and Functional MRI (sMRI and fMRI) Among the varieties of brain imaging modalities, magnetic resonance imaging (MRI) is primarily a noninvasive imaging technique used in radiology to visualize the brain's structure and function. Two main forms of MRI include: Structural MRI (sMRI) images the anatomy and strucure of the brain (Symms et al., 2004) and provides detailed pictures of the brain's size and shape; functional MRI (fMRI) identifies active regions, patterns of functional connectivities during either tasks specifically designed to study various aspects of brain fundtion or during the resting state (Martijn et al., 2010). The MRI machine is, in essence, a big magnet. As the subject lies in its magnetic field, invisible radio waves are released around the subject. This will result in harmless radio waves bouncing off the different substances that make up the brain. The radio waves are then detected by a computer, which transforms the data into images of the brain's structure and activity. In fMRI, as the subject lies in the MRI machine, simple tasks are given; the MRI then maps what parts of the brain are most active during those tasks compared with activity while the brain is at rest. This allows researchers to understand how the brain functions. This information is used together with the data from the sMRI data to reveal a comprehensive picture of brain structure and function that fit in the overall studies or to allow us to understand how the healthy brain works. The informaiton and fusion of structural and functional MRI can also improve our understanding and the treatment of neurodegenerative diseases and mental disorders such as Alzheimer's disease and schizophrenia.

Brain Morphometry Analysis with Hypothesis Testing from Structural MRI Structural MRI (sMRI), or simply called MRI, scans are usually stored in the format of three-dimensional (3D) voxels. There are several procedures for MRI post-processing, and the two

important ones are registration and segmentation. The registration maps an MRI scan to a pre-defined template (i.e. matches anatomical landmarks from different MRI images); this makes the exploration of group differences achievable. The segmentation classifies the voxels of an MRI scan as gray matter, white matter, cerebrospinal fluid, background, or region of interest (ROI); it serves as a foundation form for many analytical tools, including voxel-based morphometry, shape-based morphometry, and cortical thickness measuring, etc.

Volumetry analysis of the whole brain (Buckner et al., 2004) and ROIs (Jack et al., 1997; Wang et al., 2003) have been traditionally used to obtain the measurements of anatomical volumes and to investigate normal or abnormal tissue structure. However, pure volume measures of the brain or ROIs do not reveal the localized regional morphometry of brain structures. In addition, it is based on the definition of regions according to some a prior hypothesis, which, in practice, is not always available. Thus, in general, it limits the ability of a study to identify new and previously unexplored relationships between structural changes. The localization limitation of volumetry analysis can be overcome by methods generally referred to as high-dimensional morphologic analysis, such as voxel-based morphometry (VBM) (Ashburner and Friston, 2000; Chung et al., 2001; Davatzikos, et al., 2001), or surface-based (i.e. shape-based) morphometry (SBM) that examines the corresponding surface vertex locations or shape differences (Shen et al., 2005; Styner et al., 2005; Thompson et al., 2004). The outputs from these methods are statistical parametric maps of the 3D brain volume or the 3D surface of the ROIs, showing differences at each voxel (in VBM) or vertex (in SBM) between the comparison groups. Thus, the subsequent inference of differences among the groups is usually performed through hypothesis testing at each voxel or at each vertex.

The standard parametric test, such as *t*-test or *F*-test, could be used in brain morphometry analysis for simplicity with the assumption that the data to be tested are independent, identically, and normally distributed, for small or medium size samples. When the sample size is large enough, this assumption is not that strict any more. However, in practical neuroimage analysis, the distribution of the data is typically unknown and sample size is quite small, in which case, the nonparametric randomization or permutation tests can be applied for improved accuracy. Permutation tests obtain *p*-values from permutation distributions of a test statistic, rather than from parametric distributions. They belong to the nonparametric "distribution-free" category of hypothesis testing and are thus flexible, and have been used successfully in biomedical image analysis (Nichols & Holmes, 2001; Pantazis, et al., 2004; Zhou et al., 2009). One way to construct the permutation distribution is through exact permutation which enumerates all possible arrangements. Another way is to construct an approximate permutation distribution based on random sampling from all possible permutations (i.e. random permutation). The computational cost is the main disadvantage of exact permutation. Random permutation has the problem of replication and causes more Type I errors. When a large number of repeated tests are needed, it is also computationally expensive to achieve satisfactory *p*-value accuracy. In Section 2, we present our novel moments-based permutation methods, which take advantage of the parametric and nonparametric features for both efficiency and accuracy.

Brain Connectivity Analysis from Functional MRI fMRI is a powerful technique that noninvasively measures and characterizes brain functions in humans under various cognitive and behavioral tasks. One of the most common forms of fMRI is the Blood Oxygen Level-Dependent (BOLD) imaging (Ogawa et al., 1990), measuring the magnetic resonance

properties of the blood. As neurons do not have direct energy sources but only get energy from blood, more active neurons will need to be supplied with energy from the blood at a higher rate. Therefore, this BOLD contrast, is able to show which parts of the brain are more active. At a number of different time points over the course of an expeirment, fMRI provides a set of scans (at different depths through the brain) constituting a volume. fMRI data is a time-course of the BOLD intensity for each voxel in the brain.

During fMRI data acquisition, even a light move of a subject's head can cause severe irregularities within the acquired data. To account for these potential movements, a realignment or motion correction procedure needs to be performed on the data (Lindquist, 2008). This usually entails looking for six parameters - three rotations and three translations, that lead the volumes maximally aligned. The next pre-processing step is normalization: each complete set of volumes is normalized to a canonical brain, or the same stereo-tactic space. This is especially useful in multiple subjects studies to account for differences in brain size. Moreover, in order to improve the data signal to noise ratio, a spatial smoothing is often carried out by comvolving a Gaussin kernel with the fMRI data.

A number of analytic methods have been developed for detecting brain activity patterns and how these patterns change in patients with cognitive disorders (Calhoun et al., 2001; McIntosh & Lobaugh, 2004; Worsley & Friston, 1995). A thorough understanding of the neural mechanisms not only requires the accurate delineation of activation regions ("functional segregation or specification") but demands precise description of function in terms of the information flow across networks of areas ("functional integration"). That is, our brain is a newtork: it consistes of spatially distributed, but functionally linked regions that continuously share information with each other. Various approaches have been proposed to extract association information from fMRI datasets, most of which rely on either functional or effective connectivity (Horwitz, 2003). Functional connectivity has been identified as "temporal correlations between spatially remote neurophysiological events" (Friston et al., 1993). In Section 3, we present a novel and general statistical framework for robust and more complete estimation of functional connectivity or brain networks.

Overview In this chapter, we will present the statistical methods we have developed for the problems in the realms of brain morphometry and connectivity from analyzing structural and functional MRI data. The integration of the recovered structure and function from these imaging data may be able to provide complementary information and thus enhance our understanding of how the brain works and how its diseases occur. We will provide an explaination of the problem areas, a description of the statistical techniques involved and a demonstration of results on simulated and real imaging data using these statistical methods.

2. Brain shape morphometry analysis using novel permutation methods

There is increasing evidence that surface shape analysis of brain structures provides new information which is not available by conventional analysis. A critical issue in surface morphometry is the shape description and representation. Various strategies have been investigated recently in the literature, such as (Brechbühler et al., 1995; Thompson et al., 2004; Wang & Staib, 2000). The spherical harmonics (SPHARM) approach using spherical harmonics as basis functions for a parametric surface description was proposed in (Brechbühler et al., 1995). The correspondence across different surfaces is established by aligning the parameterizations via the first order ellipsoid. The present work employs the SPHARM-PDM shape description (Styner et al., 2006), which leads to corresponding

location vectors across all surfaces for our subsequent statistical analysis of surface shape. At each corresponding position on the surfaces, we test whether there is a significant mean vector difference between location vectors of two groups. If a hypothesis test leads to a p-value smaller than the pre-chosen α-level, we reject the null hypothesis and conclude that a significant shape difference exists at this surface location. In this chapter, we focus on the surface shape analysis for two groups, though our method can be extended to the multi-group case.

Since the distribution of the location vectors is unknown, only a limited number of subject samples are available, and the same tests are repeated on thousands of locations, we propose to use our hybrid or moments-based permutation approach to the brain shape analysis. This approach takes advantage of nonparametric permutation tests and parametric Pearson distribution approximation for both efficiency and accuracy/flexibility. Specifically, we employ a general theoretical method to derive moments of permutation distribution for any linear test statistics. Here, the term "linear test statistic" refers to a linear function of test statistic coefficients, instead of that of data. An extension of the method to the general weighted v-statistics has also been developed recently in (Zhou et al., 2009). The key idea is to separate the moments of permutation distribution into two parts, permutation of test statistic coefficients and function of the data. We can then obtain the moments without any permutation since the permutation of test statistic coefficients can be derived theoretically. Given the first four moments, the permutation distribution can be well fitted by Pearson distribution series. The p-values are then estimated without any real permutation. For multiple comparison of two-group difference, given the sample size $n_1 = 21$ and $n_2 = 21$, the number of tests is $m = 2000$. $m \times (n_1+n_2)!/\ n_1!/\ n_2! \approx 1.1 \times 10^{15}$ permutations are needed for an exact permutation test. Even for 20,000 random permutations per test, 4×10^7 permutations are still required. Alternatively, our hybrid or moments-based permutation method using Pearson distribution approximation only involves the calculation of analytically derived first four moments of exact permutation distributions while achieve high accuracy. Instead of calculating the test statistics in factorial scale with exact permutation, our permutation using mean difference test statistic only require O(n) computation cost, where $n = n_1+n_2$.

2.1 Hypothesis

Classical Hypothesis Given registered location vectors across all subjects, surface shape morphometry analysis becomes a two-sample test for equality of means at each surface location. The hypothesis is typically constructed as:

$$H_0: \ \underset{\sim}{\mu}_A = \underset{\sim}{\mu}_B \quad \text{vs.} \quad H_a: \ \underset{\sim}{\mu}_A \neq \underset{\sim}{\mu}_B \tag{1}$$

where $\underset{\sim}{\mu}_A = [\mu_A^{(x)}\ \mu_A^{(y)}\ \mu_A^{(z)}]^T$ and $\underset{\sim}{\mu}_B = [\mu_B^{(x)}\ \mu_B^{(y)}\ \mu_B^{(z)}]^T$ are three dimensional mean vectors of group A and B.

Bioequivalence Hypothesis In many applications, statistical significance is not equivalent to practical significance since smaller differences of two group location vectors can be more statistically significant than the larger ones. Statistical significance means that the observed difference is not a consequence of sampling error. Practical significance indicates whether the difference is large enough to be of value in a practical sense. Statistical significance does not necessarily indicate practical significance because extremely small and non-notable differences can be statistically significant. For example, there are two pairs of observed mean

location vectors $(\underset{\sim}{\mu}_{A1}, \underset{\sim}{\mu}_{B1})$ at location 1 and $(\underset{\sim}{\mu}_{A2}, \underset{\sim}{\mu}_{B2})$ at location 2, with $\underset{\sim}{\mu}_{A1} = [1,1,1]^T$, $\underset{\sim}{\mu}_{B1} = [0.999, 0.999, 0.999]^T$, $\underset{\sim}{\mu}_{A2} = [1,1,1]^T$, and $\underset{\sim}{\mu}_{B2} = [0.7, 0.7, 0.7]^T$. We assume that the variance of location vectors at location 2 is much larger than that at location 1, and their p-values of the observed mean differences are $p_1 = 0.001$ and $p_2 = 0.01$ respectively. The mean difference at location 1 is physically very small, although it is more statistical significant than the one at location 2. In this case, it is more reasonable to identify practical or physical shape difference at location 2 rather than at location 1. In order to achieve this, we propose to use the multivariate bioequivalence hypothesis test for our surface morphometry analysis:

$$\begin{aligned} H_0: &\quad \max\{\left|\mu_A^{(s)} - \mu_B^{(s)}\right|\} \le \Delta, s \in \{x, y, z\} \quad \text{bioequivalence} \\ H_a: &\quad \max\{\left|\mu_A^{(s)} - \mu_B^{(s)}\right|\} > \Delta, s \in \{x, y, z\} \quad \text{bioinequivalence} \end{aligned} \tag{2}$$

where Δ is the desired threshold. That is, the shape difference is detected as significant if the mean vector difference is large enough in either x, y or z direction. Bioequivalence tests were originally introduced in the pharmaceutical industry to determine the bioequivalence (Brown et al., 1997). Here, we employ bioequivalence concept though for detecting bioinequivalence as in Eq. (2) we constructed, instead of bioequivalence as in the standard pharmaceutical studies.

A permutation test is valid if the observations are exchangeable under the null hypothesis. However, the condition of exchangeability under null hypothesis is not satisfied in hypothesis Eq. (2). We thus propose to utilize a two-step permutation test.

Step 1: $H_0^{(1)}: \quad \mu_A^{(s)} = \mu_B^{(s)} + \Delta, s \in \{x, y, z\}$

$$H_a^{(1)}: \quad \mu_A^{(x)} > \mu_B^{(x)} + \Delta \ \ or \ \mu_A^{(y)} > \mu_B^{(y)} + \Delta \ \ or \ \mu_A^{(z)} > \mu_B^{(z)} + \Delta \tag{3}$$

Step 2: $H_0^{(2)}: \quad \mu_A^{(s)} = \mu_B^{(s)} - \Delta, \ s \in \{x, y, z\}$

$$H_a^{(2)}: \quad \mu_A^{(x)} < \mu_B^{(x)} - \Delta \ \ or \ \mu_A^{(y)} < \mu_B^{(y)} - \Delta \ \ or \ \mu_A^{(z)} < \mu_B^{(z)} - \Delta \tag{4}$$

If a hypothesis test of significance in step 1 (Eq. (3)) or in step 2 (Eq. (4)) gives a p-value lower than the $\alpha/2$-level, we reject the null hypothesis and significant shape difference exists. The total significance level in this case is still α due to the involved two steps in Eq. (3) and Eq. (4). Note that the classical hypothesis is a special case of the bioequivalence hypothesis when $\Delta = 0$. Classical hypothesis is used in applications where statistical and practical significances are consistent. Otherwise, bioequivalence test is preferred if there is any non-negligible difference between practical significance and statistical significance.

2.2 New Permutation Approach

Pearson Distribution Series The Pearson distribution series (Pearson I ~ VII) are a family of probability distributions that are more general than the normal distribution (Hubert, 1987). As shown in Fig. 1 (Hahn & Shapiro, 1967), it covers all distributions in the (β1, β2) plane including normal, beta, gamma, log-normal and etc., where distribution shape parameters

β1, β2 are the square of the standardized skewness and kurtosis measurements, respectively. Given the first four moments, the Pearson distribution series can be utilized to approximate the permutation distribution of the test statistic without conducting real permutation.

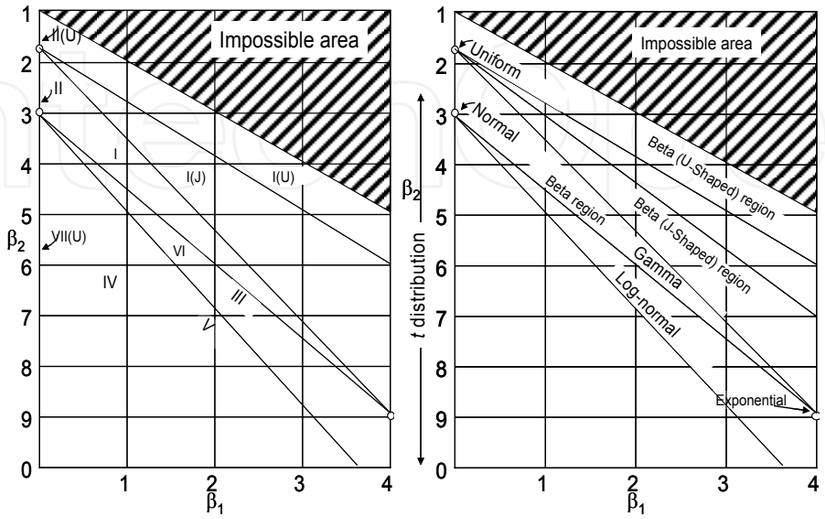

Fig. 1. Left: Regions in ($\beta1$, $\beta2$) plane for various types of Pearson distribution series; Right: Regions in ($\beta1$, $\beta2$) plane for various types of parametric distributions (Hahn & Shapiro, 1967).

Theoretical Derivation of Moments In order to approximate the permutation distribution with Pearson distribution, the moments of the exact permutation distribution need to be computed. Let $X = [x_1, x_2, \ldots, x_n]^T$ be the one dimensional data, and $T = C^TPX$ denotes the linear test statistic for permutation test. $C = [c_1 \cdots c_n]^T$ is the linear test statistic coefficients vector. The permutation matrix P is a matrix that has exactly one entry 1 in each row and each column and 0's elsewhere. A permutation matrix is a matrix obtained by permuting the rows of an identity matrix according to some permutation of the numbers 1 to n. For example, let $P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, $X = [x_1\ x_2\ x_3]^T$ and $C = [c_1\ c_2\ c_3]^T$, thus $PX = [x_3\ x_1\ x_2]^T$ and $T(X, \pi) = C^TPX = c_2x_1 + c_3x_2 + c_1x_3 = \sum_i c_{\pi(i)}x_i$, where $(\pi(1), \pi(2), \pi(3)) = (2, 3, 1)$ denotes a permutation of vector data by row. Then the linear test statistic can be denoted as $T(X, \pi) = C^TPX = \sum_i c_{\pi(i)}x_i$. The r-th conditional moment can be derived as:

$$E(T^r(X, \pi)\big|X) = E((C^TPX)^r) = \frac{1}{n!}\sum_{\pi}\sum_{i_1,\cdots i_r}(\prod_{k=1}^{r} c_{\pi(i_k)}x_{i_k})$$

$$= \frac{1}{n!}\sum_{\pi}\sum_{i_1,\cdots i_r}(\prod_{k=1}^{r} x_{i_k}\prod_{k=1}^{r} c_{\pi(i_k)}) = \sum_{i_1,\cdots i_r}(\prod_{k=1}^{r} x_{i_k}\frac{1}{n!}\sum_{\pi}\prod_{k=1}^{r} c_{\pi(i_k)}) \quad (5)$$

To compute $\frac{1}{n!}\sum_{\pi}\prod_{k=1}^{r} c_{\pi(i_k)}$, it is natural to partition the index space $U=\{1\,2\cdots n\}^r$ into $\bigcup_{(\lambda_1,\lambda_2,\cdots,\lambda_q)\in L} U^{(\lambda_1,\lambda_2,\cdots,\lambda_q)}$, where $L=\{(\lambda_1,\lambda_2,\cdots,\lambda_q):\lambda_1,\lambda_2,\cdots,\lambda_q\in Z^+;\ \lambda_1\le\lambda_2\le\cdots\le\lambda_q;\ \lambda_1+\lambda_2+\cdots+\lambda_q=r\}$. $U^{(\lambda_1,\lambda_2,\cdots,\lambda_q)}$ means that all r indices are permuted into q different numbers. Each number corresponds to λ_i indices. When $r=3$, $U=U^{(1,1,1)}\bigcup U^{(1,2)}\bigcup U^{(3)}$, where $U^{(1,1,1)}$ is the set of $\{i_1\neq i_2\ and\ i_1\neq i_3\ and\ i_2\neq i_3\}$ with $q=3$ and $\lambda_1=\lambda_2=\lambda_3=1$, $U^{(1,2)}$ is the set of $\{i_1=i_2\neq i_3\ or\ i_1=i_3\neq i_2\ or\ i_2=i_3\neq i_1\}$ with $q=2$ and $\lambda_1=1,\lambda_2=2$, and $U^{(3)}$ is the set of $\{i_1=i_2=i_3\}$ with $q=1$ and $\lambda_1=3$. Since permutation is equally related to all r indices, $\frac{1}{n!}\sum_{\pi}\prod_{k=1}^{r} c_{\pi(i_k)}$ is invariant in each category, we define $\frac{1}{n!}\sum_{\pi}\prod_{k=1}^{r} c_{\pi(i_k)}$ as moment coefficient, $a_{(\lambda_1,\lambda_2,\cdots,\lambda_q)}$, if $(i_1\,i_2\cdots i_r)\in U^{(\lambda_1,\lambda_2,\cdots,\lambda_q)}$. Eventually, the *r*-th conditional moment is:

$$E(T^r(X,\pi)\big|X)=\sum_{\lambda_1,\lambda_2,\cdots,\lambda_q} a_{(\lambda_1,\lambda_2,\cdots,\lambda_q)} \sum_{(i_1,\cdots i_r)\in U^{(\lambda_1,\lambda_2,\cdots,\lambda_q)}} \left(\prod_{k=1}^{r} x_{i_k}\right) \tag{6}$$

Eq. (6) separates the permutation from the data. To get the moments, we only need to derive the permutation of the coefficients of pre-chose test statistics and calculate the summation terms of data. Due to the simple pattern of the coefficients of test statistics which is the same for repeated tests, we can derive the moments of permutation distribution without permutation of the data. Alternatively, all a s can also be calculated by computer simulation without analytical derivation. In addition, the discussed approach can be easily extended to the multivariate case (Zhou & Wang, 2008).

Mean Difference Test Statistic In surface shape analysis, we use mean vector difference as test statistic, and $T = [TX, TY, TZ]^T = [CPX^*, CPY^*, CPZ^*]^T$, where the mean difference vector $C=[\frac{1}{n_1}1_{1\times n_1},\frac{-1}{n_2}1_{1\times n_2}]$, $X^*=X+[0_{1\times n_1},1_{1\times n_2}]^T\Delta$, $Y^*=Y+[0_{1\times n_1},1_{1\times n_2}]^T\Delta$, $Z^*=Z+[0_{1\times n_1},1_{1\times n_2}]^T\Delta$. Δ is the desired threshold for bioequivalence test, and is equal to zero in classical hypothesis test case. The detailed and complete formulas of corresponding a s are derived and listed in (Zhou & Wang, 2008). For the mean difference test statistic, the computation cost of data summation terms for the r-th moment in each index subspace can be reduced to $O(n)$ from $O(n^r)$.

Multiple Comparison via Adaptive ROI Constrained FDR Determining whether a location on the brain surface has significant group shape difference or not corresponds to performing a described hypothesis test at that position. Clearly, the location-wise *p*-values are spatially dependent. The significance rule, applied in the conventional False Discovery Rate (FDR) approach is defined as the expected proportion of false positives among the declared significant results. It is more powerful and less stringent than the Family-Wise-Error-Rate (FWER) approach. We adopt the adaptive concept of the FDR (Benjamini et al., 2006) and develop a ROI constrained adaptive FDR in (Zhou & Wang, 2008). This adaptive FDR

control is more powerful than the conventional one. It can find more significant areas while preserving the same desired FDR rate.

2.3 Experiments and results

Simulated Data In this experiment, we generated six different simulated data sets to evaluate our hybrid permutation tests. In case #1 and case #2, two group data are normal distributed with different mean and variance (Normal(0,1) vs. Normal(1,0.5)) in balanced design ($n_1 = n_2 = 10$) and unbalanced design ($n_1 = 6$, $n_2 = 18$), respectively. Each group has gamma distribution in case #3 (Gamma(3,3) vs. (Gamma(3,2), $n_1 = n_2 = 10$) and case #4 (Gamma(3,3) vs. (Gamma(3,2), $n_1 = 6$, $n_2 = 18$). In case #5 and case #6, two group data have beta distribution with different parameters (Beta(0.8,0.8) vs. Beta(0.1,0.1)) in balanced design ($n_1 = n_2 = 10$) and unbalanced design (n_1=6, n_2=18).

	Case #1	Case #2	Case #3	Case#4	Case#5	Case#6
t_HP	0.0113	0.0113	0.0123	0.0137	0.0172	0.0018
t_RP	1.1584	1.1438	1.1369	1.1250	1.1262	1.1384
t_EP	4.4389	4.2795	4.2983	4.3240	4.1320	4.2948
p_HP	0.0499	0.1314	0.0010	0.0249	0.0908	0.0805
p_RP	0.0495	0.1269	0.0012	0.0242	0.0889	0.0818
p_EP	0.0498	0.1301	0.0010	0.0250	0.0925	0.0803

Table 1. Comparison of computation costs and *p*-value accuracy for three permutation test methods. (HP: hybrid permutation; RP: random permutation; EP: exact permutation.). t_HP, t_RP and t_HP denote the respective computation time (in seconds) per test; p_HP, p_RP and p_EP are the respective *p*-value measurements by the three permutation methods.

Table 1 indicates the high accuracy of our hybrid permutation technique, especially for the tail area (Note: the exact permutation results are considered as ground truth.) Furthermore, comparing with exact permutation or random 20,000 permutations, the hybrid permutation tests reduce more than 99% computation cost and can further save computation time as the sample size increases. In order to demonstrate the robustness of our method, we repeated the simulation for 10 times in each case, and calculate the mean and variance of the absolute biases of *p*-values of both hybrid permutation and random permutation, treating the *p*-values of exact permutation as gold standard. In most cases, hybrid permutation is less biased and more stable than random permutation (Table 2), which demonstrates the robustness and accuracy of our method.

We also generated a synthetic dataset to demonstrate that bioequivalence test plays an important role in identifying practical significance. There are 12 surfaces in group A and 9 in group B, which were generated by adding two types of Gaussian noises to the two flat patches, a (5×5) top patch and a (21×21-5×5) bottom patch. For group A, Gaussian noise with mean zero and standard deviation $\sigma_b = 0.01$ was added to the bottom patch with $z = 0$; Gaussian noise with mean zero and standard deviation $\sigma_t = 0.09$ is added to the top patch with $z = 1$. The 9 surfaces in group B were generated with the same noise patterns as in group A but to different bottom patch $z = 0.01$ and top patch $z = 0.9$. Since the differences between the bottom patches of the two groups are very small ($z = 0$ vs. $z = 0.01$), the practical group differences should only occur on the top patch ($z = 1$ vs. $z = 0.9$). Fig. 2(c) shows that

	Case #1	Case #2	Case #3	Case #4	Case #5	Case #6
Mean_ABias_HP	8.79e-5	8.97e-6	9.54e-5	2.16e-4	6.79e-4	4.53e-4
Mean_ABias_RP	2.82e-4	6.64e-5	2.14e-4	1.30e-3	2.78e-4	5.99e-4
VAR_ABias_HP	5.99e-8	1.34e-7	2.10e-6	3.66e-7	9.55e-7	9.78e-6
VAR_ABias_RP	1.98e-6	1.42e-7	1.41e-6	5.34e-6	1.05e-5	1.00e-5

Table 2. Robustness and accuracy comparison of hybrid (moments-based) permutation and random permutation across 10 simulations, considering the *p*-values of exact permutation as gold standard. Mean_ABias_HP and VAR_ABias_HP are the mean and variance of the absolute biases of *p*-values of hybrid permutation; Mean_ABias_RP and VAR_ABias_RP are the mean and variance of the absolute biases of *p*-values of random permutation, respectively.

many significance locations on both the top and bottom patches are detected by the classical hypothesis tests. Also, the non-practical significances can not be revealed with a more conservative significance level (i.e., lower α) because not all p-values on the top patch are lower than those of the bottom patch (see Fig. 2(d)). Using bioequivalence tests, we are able to precisely identify the practical significances that occur at the top patch (Fig. 2(e)).

Real Data of the Human Brain The MRI hippocampi used in this experiment were semi-automatically segmented by human expert raters and manually grouped into 2 groups with 21 subjects in group A and 15 in group B (see the following site for details: *http://www.ia.unc.edu/dev/download/shapeAnalysis/*). This dataset serves as a testing dataset for methodology validation for all users of the SPHARM-PDM software. Evaluation of our bioequivalence test using hybrid permutation with the mean difference test statistics on this hippocampus dataset is shown in Fig. 3. It can be seen that the Pearson distribution approximation based on the first four moments leads to an accurate approximation to the real permutation distribution and thus *p*-values (Fig. 3(c) vs. Fig. 3(d)). At the same significance level $\alpha = 0.05$, there are differences (Fig. 3(e)) between the results of classical test

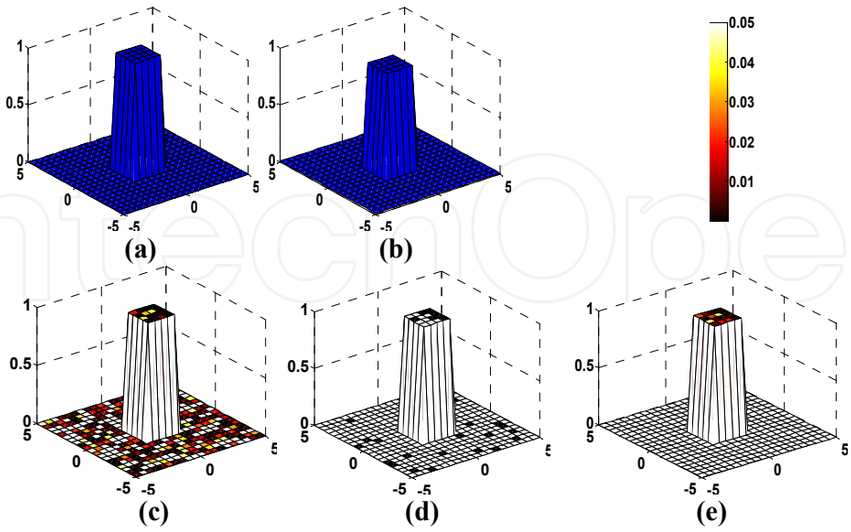

Fig. 2. (a): Mean shape of group A. (b): Mean shape of group B. (c) and (d): Results using conventional hypothesis tests with $\alpha = 0.05$ (c) and with $\alpha = 0.001$ (d). (e) Results using bioequivalence tests with $\alpha = 0.05$ and $\Delta = 0.025$.

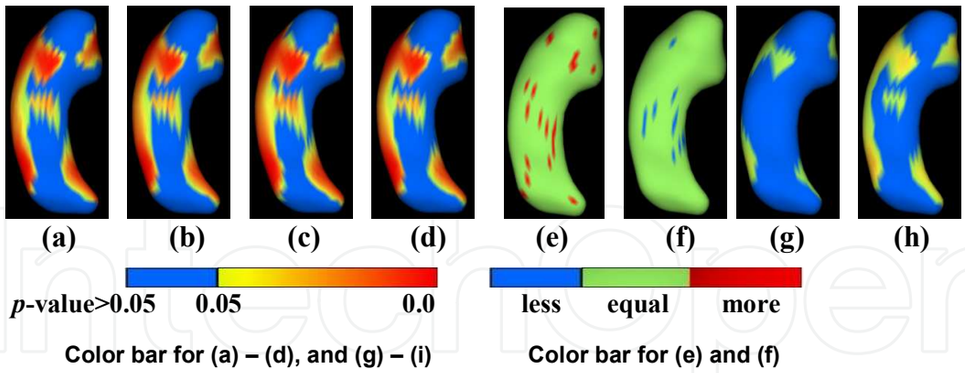

Fig. 3. (a) and (b): Raw *p*-value maps from classical hypothesis test using Pearson approximation, at $\alpha = 0.05$ (a), and at $\alpha = 0.035$ (b). (c) and (d): Raw *p*-value maps from bioequivalence test at $\alpha = 0.05$ (without correction), through real permutation ((c); number of permutations = 10,000), and using Pearson approximation (d). (e): The difference between (a) and (d), locations in red showing significances in (a) but not in (d). (f): The difference between (b) and (d), locations in blue showing significances in (d) but not in (b). (g): Our ROI constrained adaptive FDR corrected *p*-map for (d). (h): Our ROI constrained adaptive FDR corrected *p*-map for (a).

(Fig. 3(a)) and bioequivalence test (Fig. 3(d)). In addition, at a lower significance level, ($\alpha = 0.035$, Fig. 3(b)) both the non-notable differences and some practical significances that we would like to detect are shown as non-significant (see Fig. 3(f)). This indicates that simply decreasing the significance level in a standard hypothesis test may not lead to practical significances, which can only be achieved through the proposed bioequivalence test. The false positive error control results are shown in Fig. 3 (g) and (h).

3. Brain connectivity analysis from functional MRI

For functional connectivity studies, a common approach is to calculate the temporal correlation coefficients of a fMRI signal from a selected voxel or region (so called "seed", or "seed region") in a region of interest with all other voxels in the brain (Worsley et al., 2005). Each correlation map is resulting from the cross-correlation of only one seed region. However, when areas with quite different time series patterns are used as seed regions for brain connectivity inference, they should not be grouped as a single region; in some applications, functional co-activation to multiple seeds rather than a single one would be of particular interest. Multiple seeds can be chosen to calculate multiple correlation maps to separately discover the functional connectivity to different seeds. But how to reasonably integrate multiple connectivity maps for brain function inference is still unresolved and ambiguous. Furthermore, it is often unrealistic to examine all pair-wise correlations. Therefore, it is desirable to have a single correlation map resulting from the cross-correlation of two or more seed regions simultaneously.

Dynamic connections in fMRI are thought to be reflected by high temporal correlations of the time series. The strong correlation between the time series of each region in the network with that of another distant region implied by the functional interactions may be related to the spatially structured noise in fMRI (Cordes, et al., 2002). The spatial correlations of the

noise must therefore be taken into account when dealing with sensitive and reproducible estimation of the network. Furthermore, brain functional connectivity based on marginal correlation can be dominated by the stimulus-locked responses. For example, if visual and auditory stimuli are presented concurrently, the stimulus-locked neural responses would cause increases in the BOLD signal in the primary auditory cortex (A1) and the primary visual cortex (V1) simultaneously. Correlation between A1 and V1 would thus be high, though not due to any intrinsic task-induced functional couplings but due to the responses in both regions to externally driven stimuli. Partial correlation is the conditional correlation which estimates any remaining correlation between time series after taking into account the relationship of each to one or more reference time series. The stimulus-locked responses can then be accounted for by choosing the reference functions to model the external stimuli. This allows us to measure any additional task-induced, but not stimulus-locked relation over brain regions. Recently, methods using partial correlation (or coherence) have been proposed (Sun et al., 2004; Marrelec et al., 2006), though they are for pair-wise correlation (or coherence) analysis and not applicable to multiple seeds. How to apply the partial correlation concept to multiple seed regions to brain connectivity study while considering spatial partial correlations in noise to those seed regions is challenging. Here, we propose a novel procedure to achieve this as one of the goals of this work.

3.1 Functional connectivity using multiple correlations

Estimating Temporal/Sample Multiple Correlations Let T be the total number of time points of the fMRI data. The temporal or sample multiple correlation coefficient considers the fMRI time series correlation between a given voxel X and a combination of seed regions, $S_1, S_2, \cdots S_P$. Its estimation is based on the variance-covariance matrix:

$$\hat{\boldsymbol{\Sigma}}_{tem} = \begin{bmatrix} \hat{\text{var}}_X & \hat{\text{cov}}_{X,S_1} & \cdots & \hat{\text{cov}}_{X,S_P} \\ \hat{\text{cov}}_{S_1,X} & \hat{\text{var}}_{S_1} & \cdots & \hat{\text{cov}}_{S_1,S_P} \\ \vdots & \vdots & \vdots & \vdots \\ \hat{\text{cov}}_{S_P,X} & \hat{\text{cov}}_{S_P,S_1} & \cdots & \hat{\text{var}}_{S_P} \end{bmatrix} = \begin{bmatrix} \hat{\text{var}}_X & \hat{\mathbf{cov}}_{X\mathbf{S}}^{'} \\ \hat{\mathbf{cov}}_{X\mathbf{S}} & \hat{\boldsymbol{\Sigma}}_{\mathbf{SS}} \end{bmatrix},$$

where $\text{var}(X)$ and var_{S_p} are the time series variances for voxel X and seed S_p ($p = 1,2,\cdots P$), respectively; and cov_{X,S_p} is their covariance. The temporal multiple correlation coefficient $\hat{R}_{tem}$ between voxel X and the multiple seeds $[S_1, S_2, \cdots S_P]^{'}$ can be calculated as (Anderson, 2003):

$$\hat{R}_{tem} = \sqrt{\frac{\hat{\mathbf{cov}}_{X\mathbf{S}}^{'} \cdot \hat{\boldsymbol{\Sigma}}_{\mathbf{SS}}^{-1} \cdot \hat{\mathbf{cov}}_{X\mathbf{S}}}{\hat{\text{var}}_X}} .$$

Estimating Spatial Multiple Correlations in Noise The factors contributing to the spatial correlation of the noise include fMRI data preprocessing, the point spread function, which causes data from an individual voxel to contain some signal from the tissue around that voxel, an effect compounded by motion correction techniques, and the smoothness introduced by interpolation in motion correction (Woolrich et al., 2004). Despite the

strategies and efforts to reduce such structured noise (Wang et al., 2003; Wang, 2005), some residual and further corrections are still essential for robust fMRI data analysis.

Voxel-based Spatial Correlograph of Noise. We assume the spatial noise is stationary and has a multivariate Gaussian distribution with variance-covariance matrix $\mathbf{\Sigma} = \left(\sigma_{i,j}\right)_{i,j=1}^{M}$, where M is the total number of voxels; σ_i and σ_j are positive standard deviations for voxels X_i and X_j. The spatial correlations in noise then depend only on the spatial distance between voxels: $\sigma_{ij} = \sigma_i \sigma_j \rho\left(\|i-j\|\right)$, where $\|i-j\|$ denotes the spatial distance or lag between X_i and X_j; and ρ is the spatial correlogram, a real-valued function that satisfies $\rho(0) = 1$ and is bounded by -1 and 1. Since it is unknown what voxels or regions are predominantly influenced by the noise, the entire set $D_h = \left\{\left(X_i, X_j\right) | \|i-j\| = h\right\}$ of pairs of voxels at lag h over the whole brain area is considered for the non-parametric estimate based on the median: $\hat{\rho}(h) = median\left\{r_{ij},\ \left(X_i, X_j\right) \in D_h\right\}$, , where r_{ij} is the Pearson's linear correlation between the time series of the two voxels. As in general the empirical estimator $\hat{\rho}$ of the correlogram does not provide a positive-definite correlation matrix, we focus on a parametric class of valid matrices, based on the empirical values $\hat{\rho}$ estimated from the fMRI data. The rational-quadratic model $\rho_{\boldsymbol{\theta}}(h)$ in (Cressie, 1993; Wang & Xia, 2009) is utilized for such purpose. The derived spatial correlogram of noise, $\rho_{\boldsymbol{\theta}}(h)$, decreases rapidly from a correlation level between nearby voxels, ρ_{0+}, towards an asymptotic correlation, ρ_∞. A critical distance h_∞^ε can be determined beyond which the correlogram is almost equal to the asymptote, with a tolerance of ε. The parameterization of the rational-quadratic model using $\left(\rho_{0+}, \rho_\infty, h_\infty^\varepsilon\right)$ is given as in (Bellec et al., 2006; Wang & Xia, 2009).

Spatial Multiple Correlations in Noise. The spatial multiple correlations of the noise consider the correlations between any voxel X and a combination of multiple seeds $S_1, S_2, \cdots S_P$. Suppose the distances between the voxel X and the seeds $S_1, S_2, \cdots S_P$ are respectively $h_1, h_2, \cdots h_P$, and the distances between any pair-wise seeds are h_{ij} (for S_i and S_j, $h_{ij} = h_{ji}$). The noise spatial correlation matrix for $\left[X, S_1, S_2, \cdots S_P\right]'$ can then be constructed as:

$$\mathbf{\Lambda}_{spa} = \begin{bmatrix} 1 & \rho_{\boldsymbol{\theta}}(h_1) & \cdots & \rho_{\boldsymbol{\theta}}(h_P) \\ \rho_{\boldsymbol{\theta}}(h_1) & 1 & & \rho_{\boldsymbol{\theta}}(h_{1P}) \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{\boldsymbol{\theta}}(h_P) & \rho_{\boldsymbol{\theta}}(h_{P1}) & \cdots & 1 \end{bmatrix} = \begin{bmatrix} 1 & \boldsymbol{\rho}_S{}' \\ \boldsymbol{\rho}_S & \mathbf{\Lambda}_{SS} \end{bmatrix}$$

where $\rho_{\boldsymbol{\theta}}(h)$ is the correlogram estimated above. Let σ_X^2 and $\sigma_{S_p}^2$ respectively denote the noise variance for voxel X, and seed S_p, $p = 1, 2, \cdots P$. Then the corresponding variance-covariance matrix, $\mathbf{\Sigma}_{spa}$, is:

$$\mathbf{\Sigma}_{spa} = \begin{bmatrix} \sigma_X^2 & \rho_{\boldsymbol{\theta}}(h_1) \cdot \sigma_X \cdot \sigma_{S_1} & \cdots & \rho_{\boldsymbol{\theta}}(h_P) \cdot \sigma_X \cdot \sigma_{S_p} \\ \rho_{\boldsymbol{\theta}}(h_1) \cdot \sigma_{S_1} \cdot \sigma_X & \sigma_{S_1}^2 & \cdots & \rho_{\boldsymbol{\theta}}(h_{1P}) \cdot \sigma_{S_1} \cdot \sigma_{S_p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{\boldsymbol{\theta}}(h_P) \cdot \sigma_{S_p} \cdot \sigma_X & \rho_{\boldsymbol{\theta}}(h_{P1}) \cdot \sigma_{S_p} \cdot \sigma_{S_1} & \cdots & \sigma_{S_p}^2 \end{bmatrix} = \begin{bmatrix} \sigma_X^2 & \boldsymbol{\sigma}_{XS}{}' \\ \boldsymbol{\sigma}_{XS} & \mathbf{\Sigma}_{SS} \end{bmatrix}.$$

The spatial multiple correlation coefficient of the noise between any voxel X and the seeds $\left[S_1, S_2, \cdots S_P\right]'$ is computed as: $R_{spa} = \sqrt{\frac{\boldsymbol{\sigma}_{XS}' \cdot \boldsymbol{\Sigma}_{SS}^{-1} \cdot \boldsymbol{\sigma}_{XS}}{\sigma_X^2}}$, and can be re-formulated to:

$$R_{spa} = \sqrt{\boldsymbol{\rho}_S' \cdot \boldsymbol{\Lambda}_{SS}^{-1} \cdot \boldsymbol{\rho}_S} \tag{7}$$

Identifying Functional Connectivity of Brain The factors Given the estimation of multiple correlations in noise, we use hypothesis testing to search for significant correlations between any voxel and the seed regions that are statistically unlikely to be due to noise.

Statistical Hypothesis Testing. We would like to test whether the temporal multiple correlation $\hat{R}_{tem}$ is likely to be found only by chance from the noise correlation. The hypothesis is:

$$H_0 : R_{tem} = R_{spa} \text{ vs. } H_1 : R_{tem} > R_{spa}.$$

Under the null hypothesis that the temporal multiple correlation, $\hat{R}_{tem}$, arises from a population whose multiple correlation equals the spatial multiple correlation of the noise, R_{spa}, the following quantity is a non-central F (Anderson, 2003) (pp. 153-154):

$$F = \frac{\hat{R}_{tem}^2}{1 - \hat{R}_{tem}^2} \cdot \frac{T - 1 - P}{P} \tag{8}$$

Here, the degrees of freedom are P and $T-1-P$, and the noncentrality parameter incorporating our re-formulated R_{spa} in Eq. (7) is $\frac{(T-1)\boldsymbol{\rho}_S' \boldsymbol{\Lambda}_{SS}^{-1} \boldsymbol{\Psi} \hat{\boldsymbol{\Sigma}}_{SS} \boldsymbol{\Psi} \boldsymbol{\Lambda}_{SS}^{-1} \boldsymbol{\rho}_S}{1 - \boldsymbol{\rho}_S' \cdot \boldsymbol{\Lambda}_{SS}^{-1} \cdot \boldsymbol{\rho}_S}$, where we condition on the seeds' time series, and $\boldsymbol{\Psi}$ is a $P \times P$ diagonal matrix with diagonal element $1/\sigma_{S_p}$, for $p = 1, 2, \ldots, P$. In this way, the p-value for each voxel can be calculated from this noncentral F distribution. A voxel shall be included in the functional connectivity network if the corresponding p-value is smaller than a pre-chosen type I error α (note: $\alpha = 0.05$ is used in this connectivity work). It can also be shown that under the null hypothesis of the population multiple correlation, R_{spa}, is zero (i.e. our hypothesis becomes: $H_0 : R_{tem} = 0$ vs. $H_1 : R_{tem} > 0$), the F in Eq. (8) is a central F (Anderson, 2003) (pp. 149-150), with P and $T-1-P$ degrees of freedom. In fact, this is equivalent to multiple correlation analysis of multi-seed functional connectivity but without taking the spatial correlations of the noise into consideration.

Effective Degrees of Freedom for Temporal Autocorrelation. A departure from the temporally i.i.d. (independent and identically distributed) assumption due to the temporal autocorrelation will result in a decrease in the degrees of freedom in the above hypothesis testing. To correct such possible bias, we estimate the effective degrees of freedom T_{eff}. This can be achieved through the context of the general linear model (Worsley & Friston, 1995). Note that the T_{eff} estimated this way assumes voxelwise spatial independence, which can be considered as an upper-bound estimation of our effective degrees of freedom. We approximate the effective degrees of freedom as T_{eff}, and use the estimated T_{eff} to replace the T-1 in the F statistic's calculation in Eq. (8).

Multiple Testing using Non-central F Random Field. We need to perform numerous tests equal to the total number of voxels over the brain area. In order to correct this multiple testing problem, different strategies can be potentially applied, such as Bonferroni correction, cluster-size thresholding, random field theory or false discovery rate control (Huettel et al., 2004; Benjamini & Hochbert, 1995; Logan & Rowe, 2004; Nicholos & Holmes, 2001; Zhou & Wang, 2008). The Random field theory (RFT) correction on the *t*-field, Hotelling's T^2 field, χ^2 field, central *F* field, and the correlation field has been developed by Worsley and colleagues (Cao & Worsley, 1999; Worsley, 1994). RFT estimates the number of independent statistical tests based upon the spatial correlation, or smoothness, of the experimental data. With even small to moderate amounts of smoothness in the data, the number of resels (resolution elements) will be much less than the original number of voxels. From the number of resels, one can estimate how many clusters of activity should be found by chance at a given statistical threshold. This number is known as the Euler characteristic of the data. RFT correction is less conservative than the Bonferroni correction. In this work, we use the non-central *F* RFT to correct the multiple comparison problems (Wang & Xia, 2009; Hayasaka et al., 2007).

3.2 Functional connectivity using partial multiple correlations

Estimating Temporal Partial Multiple Correlations The temporal partial multiple correlation coefficient considers the fMRI time series correlation between a given voxel *X* and a combination of seed regions $S_1, S_2, \cdots S_P$ conditioned on fixed stimuli (experimental paradigms or reference functions, i.e. the convolution functions of the hemodynamic response with the 0-1 boxcar stimulus functions in the case of the block-design experiments), $V_1, V_2, \cdots V_N$. Its estimation is based on the matrix:

$$\begin{bmatrix} \hat{\text{var}}_X & \hat{\text{cov}}_{X,S_1} & \cdots & \hat{\text{cov}}_{X,S_P} & \hat{\text{cov}}_{X,V_1} & \cdots & \hat{\text{cov}}_{X,V_N} \\ \hat{\text{cov}}_{S_1,X} & \hat{\text{var}}_{S_1} & & \hat{\text{cov}}_{S_1,S_P} & \hat{\text{cov}}_{S_1,V_1} & \cdots & \hat{\text{cov}}_{S_1,V_N} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \hat{\text{cov}}_{S_P,X} & \hat{\text{cov}}_{S_P,S_1} & \cdots & \hat{\text{var}}_{S_P} & \hat{\text{cov}}_{S_P,V_1} & \cdots & \hat{\text{cov}}_{S_P,V_N} \\ \hat{\text{cov}}_{V_1,X} & \hat{\text{cov}}_{V_1,S_1} & \cdots & \hat{\text{cov}}_{V_1,S_P} & \text{var}_{V_1} & \cdots & \text{cov}_{V_1,V_N} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \hat{\text{cov}}_{V_N,X} & \hat{\text{cov}}_{V_N,S_1} & \cdots & \hat{\text{cov}}_{V_N,S_P} & \text{cov}_{V_N,V_1} & \cdots & \text{var}_{V_N} \end{bmatrix} = \begin{bmatrix} \hat{\text{var}}_X & \hat{\mathbf{cov}}_{X\mathbf{S}}' & \hat{\mathbf{cov}}_{X\mathbf{V}}' \\ \hat{\mathbf{cov}}_{X\mathbf{S}} & \hat{\boldsymbol{\Sigma}}_{\mathbf{SS}} & \hat{\boldsymbol{\Sigma}}_{\mathbf{SV}}' \\ \hat{\mathbf{cov}}_{X\mathbf{V}} & \hat{\boldsymbol{\Sigma}}_{\mathbf{SV}} & \boldsymbol{\Sigma}_{\mathbf{VV}} \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{\Sigma}}_{\{X\mathbf{S}\}\{X\mathbf{S}\}} & \hat{\boldsymbol{\Sigma}}_{\{X\mathbf{S}\}\mathbf{V}}' \\ \hat{\boldsymbol{\Sigma}}_{\{X\mathbf{S}\}\mathbf{V}} & \boldsymbol{\Sigma}_{\mathbf{VV}} \end{bmatrix},$$

where var_{V_n} is the time series variances of the stimulus V_n $(n = 1, 2, \cdots N)$; cov_{X,V_n} is the covariance between voxel *X* and V_n, and cov_{S_p,V_n} is the covariance between the seed S_p and V_n. Their estimation can be achieved through time series and reference function samples of size *T*, though here they are not technically variances and covariances because the V_n are fixed stimuli. With the assumption that the conditional distribution $\left(X, S_1, \cdots, S_P \middle| V_1 = v_1, V_2 = v_2, \cdots V_N = v_N\right)'$ is a multi-normal distribution (Anderson, 2008), its variance-covariance matrix can be calculated as:

$$\hat{\boldsymbol{\Sigma}}_{tem}^{*} = \hat{\boldsymbol{\Sigma}}_{\{X\mathbf{S}\}\{X\mathbf{S}\}} - \hat{\boldsymbol{\Sigma}}_{\{X\mathbf{S}\}\mathbf{V}}' \cdot \boldsymbol{\Sigma}_{\mathbf{VV}}^{-1} \cdot \hat{\boldsymbol{\Sigma}}_{\{X\mathbf{S}\}\mathbf{V}} \cdot$$

Let the components of $\hat{\boldsymbol{\Sigma}}^{*}_{tem}$ be divided into four groups, $\begin{bmatrix} \hat{\text{var}}^{*}_{X} & \hat{\mathbf{cov}}^{*}_{X\mathbf{S}}{}' \\ \hat{\mathbf{cov}}^{*}_{X\mathbf{S}} & \hat{\boldsymbol{\Sigma}}^{*}_{\mathbf{SS}} \end{bmatrix}$, where $\hat{\text{var}}^{*}_{X}$ and $\hat{\boldsymbol{\Sigma}}^{*}_{\mathbf{SS}}$ are variances of voxel X and the seeds holding the reference functions (stimuli) fixed, $\hat{\mathbf{cov}}^{*}_{X\mathbf{S}}$ is their corresponding covariance under the same condition. The temporal partial multiple correlation is:

$$\hat{R}_{tem\bullet\mathbf{V}} = \sqrt{\frac{\hat{\mathbf{cov}}^{*}_{X\mathbf{S}}{}' \cdot \hat{\boldsymbol{\Sigma}}^{*-1}_{\mathbf{SS}} \cdot \hat{\mathbf{cov}}^{*}_{X\mathbf{S}}}{\hat{\text{var}}^{*}_{X}}}$$

Estimating Spatial Partial Multiple Correlations in Noise In Section 3.1, we estimate the voxel-based spatial correlograph of noise using the median of Pearson's linear correlation, i.e. marginal correlation. Here, we take the similar approach but replace the marginal correlation with partial correlation because the stimuli are now considered to be fixed, i.e., $\hat{\rho}^{*}(h) = median\left\{ r_{ij\cdot\mathbf{V}},\ \left(X_i, X_j\right) \in D_h \right\}$, , where $r_{ij\cdot\mathbf{V}}$ is the partial correlation coefficient between voxels X_i and X_j holding $V_1, V_2, \cdots V_N$ fixed. Let the variance-covariance matrix of $\left(X_i, X_j, V_1, \cdots, V_N\right)'$ be $\begin{bmatrix} \hat{\boldsymbol{\Sigma}}_{\mathbf{XX}} & \hat{\boldsymbol{\Sigma}}_{\mathbf{XV}}{}' \\ \hat{\boldsymbol{\Sigma}}_{\mathbf{XV}} & \boldsymbol{\Sigma}_{\mathbf{VV}} \end{bmatrix}$. $\left(X_i, X_j \middle| V_1 = v_1, V_2 = v_2, \cdots V_N = v_N\right)'$ is assumed to be multi-normal; its variance-covariance matrix is calculated as:

$$\hat{\boldsymbol{\Sigma}}_{\mathbf{X}\cdot\mathbf{V}} = \hat{\boldsymbol{\Sigma}}_{\mathbf{XX}} - \hat{\boldsymbol{\Sigma}}_{\mathbf{XV}}{}' \cdot \boldsymbol{\Sigma}^{-1}_{\mathbf{VV}} \cdot \hat{\boldsymbol{\Sigma}}_{\mathbf{VX}} = \begin{bmatrix} \hat{\sigma}^{2}_{i\cdot\mathbf{V}} & \hat{\sigma}'_{ij\cdot\mathbf{V}} \\ \hat{\sigma}_{ij\cdot\mathbf{V}} & \hat{\sigma}^{2}_{j\cdot\mathbf{V}} \end{bmatrix}.$$

The partial correlation coefficient is thus given by (Anderson, 2008):

$$r_{ij\cdot\mathbf{V}} = \frac{\sigma_{ij\cdot\mathbf{V}}}{\sigma_{i\cdot\mathbf{V}} \cdot \sigma_{j\cdot\mathbf{V}}}.$$

The noise spatial partial correlation matrix for $\left[X, S_1, S_2, \cdots S_P\right]'$ holding $V_1, V_2, \cdots V_N$ fixed can be constructed as:

$$\boldsymbol{\Lambda}^{*}_{spa} = \begin{bmatrix} 1 & \rho^{*}_{\boldsymbol{\theta}}(h_1) & \cdots & \rho^{*}_{\boldsymbol{\theta}}(h_P) \\ \rho^{*}_{\boldsymbol{\theta}}(h_1) & 1 & & \rho^{*}_{\boldsymbol{\theta}}(h_{1P}) \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{*}_{\boldsymbol{\theta}}(h_P) & \rho^{*}_{\boldsymbol{\theta}}(h_{P1}) & \cdots & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{1} & \boldsymbol{\rho}^{*}_{\mathbf{S}}{}' \\ \boldsymbol{\rho}^{*}_{\mathbf{S}} & \boldsymbol{\Lambda}^{*}_{\mathbf{SS}} \end{bmatrix}$$

where $\rho^{*}_{\boldsymbol{\theta}}(h)$ is the correlogram. Let $\sigma^{*}_{S_p}$ denote the residual standard deviation of the noise for seed S_p, $p = 1, 2, \cdots P$, holding $V_1, V_2, \cdots V_N$ fixed (see Appendix in [Wang & Xia, 2009]

for its estimation). The spatial partial multiple correlation coefficient of the noise between any voxel *X* and the seeds $\left[S_1, S_2, \cdots S_P\right]'$ holding the stimuli fixed is computed as:

$$R_{spa\cdot \mathbf{v}} = \sqrt{\boldsymbol{\rho}_{\mathbf{S}}^{*\prime} \cdot \boldsymbol{\Lambda}_{\mathbf{SS}}^{*-1} \cdot \boldsymbol{\rho}_{\mathbf{S}}^{*}}\ .$$

Identifying Conditional Functional Connectivity of Brain We would like to test whether the temporal partial multiple correlation $\hat{R}_{tem\cdot\mathbf{v}}$ is likely to be found only by change from the noise correlation. The hypothesis is:

$$H_0 : R_{tem.\mathbf{v}} = R_{spa\cdot\mathbf{v}} \text{ vs. } H_1 : R_{tem.\mathbf{v}} > R_{spa\cdot\mathbf{v}}\ .$$

Here, we can show that the following quantity is a non-central *F*:

$$\frac{\hat{R}^2_{tem.\mathbf{v}}}{1-\hat{R}^2_{tem.\mathbf{v}}} \cdot \frac{(T-1)-N-P}{P} \tag{9}$$

with the degrees of freedom P and $T-1-N-P$, and the noncentrality parameter $\frac{(T-1)\boldsymbol{\rho}_{\mathbf{S}}^{*\prime}\boldsymbol{\Lambda}_{\mathbf{SS}}^{*-1\prime}\boldsymbol{\Psi}^{*}\hat{\boldsymbol{\Sigma}}_{\mathbf{SS}}^{*}\boldsymbol{\Psi}^{*}\boldsymbol{\Lambda}_{\mathbf{SS}}^{*-1}\boldsymbol{\rho}_{\mathbf{S}}^{*}}{1-\boldsymbol{\rho}_{\mathbf{S}}^{*\prime}\boldsymbol{\Lambda}_{\mathbf{SS}}^{*-1}\boldsymbol{\rho}_{\mathbf{S}}^{*}}$, where we condition on the seeds' time series; $\boldsymbol{\psi}^{*}$ is a $P\times P$ diagonal matrix with diagonal element $1/\sigma^{*}_{S_p}$, for $p=1,2,\cdots P$. Note that the temporal autocorrelation can be handled in a similar way as in the multiple correlation case by computing the associated effective degrees of freedom; the multiple testing correction is similarly based on non-central *F* random field theory (see Section 3.1 for details).

Similar to the multiple correlation case, under the null hypothesis of the population partial multiple correlation, $R_{spa\cdot\mathbf{v}}$, is zero (i.e. our hypothesis becomes: $H_0 : R_{tem.\mathbf{v}} = 0$ vs. $H_1 : R_{tem.\mathbf{v}} > 0$), the *F* in Eq. (9) is a central F, with P and $T-1-N-P$ degrees of freedom. This is then equivalent to partial multiple correlation analysis of multi-seed functional connectivity but without taking the spatial partial correlations of the noise into consideration.

3.3 Experiments on real fMRI data

The real fMRI data (single-subject) was obtained from the SPM data site (*http://www.fil.ion.ucl.ac.uk/spm/data/attention*) with the detailed description in (Buchel & Friston, 1997). The subject was scanned during four runs, each lasting 5 min 22 s. One hundred image volumes were acquired and the first ten was discarded in each run. Each condition lasted 32.2 s, giving 10 multislice volumes per condition. The fMRI data size was $53\times 63\times 46\times 360$. Four conditions – 'fixation', 'attention', 'no attention', and 'stationary' – were used. Electrophysiological and neuroimaging studies have shown that attention to

visual motion can increase the responsiveness of the motion-selective cortical area V5 and some other areas, and an occipito-parieto-frontal network is involved in the visual pathway modulation by attention. The structural model for the dorsal visual pathway includes primary visual cortex (V1), visual cortical area MT (V5), posterior parietal cortex (PP), and modulatory interaction term involving dorsolateral prefrontal cortex (PFC). The activation regions were identified by categorical comparisons using the SPM5 software package, contrasting "attention" and "no attention" and contrasting "no attention" and "stationary". Here we examine the functional interactions by using the different seed regions: V1, or V5 or both V1 and V5, and by using different methods.

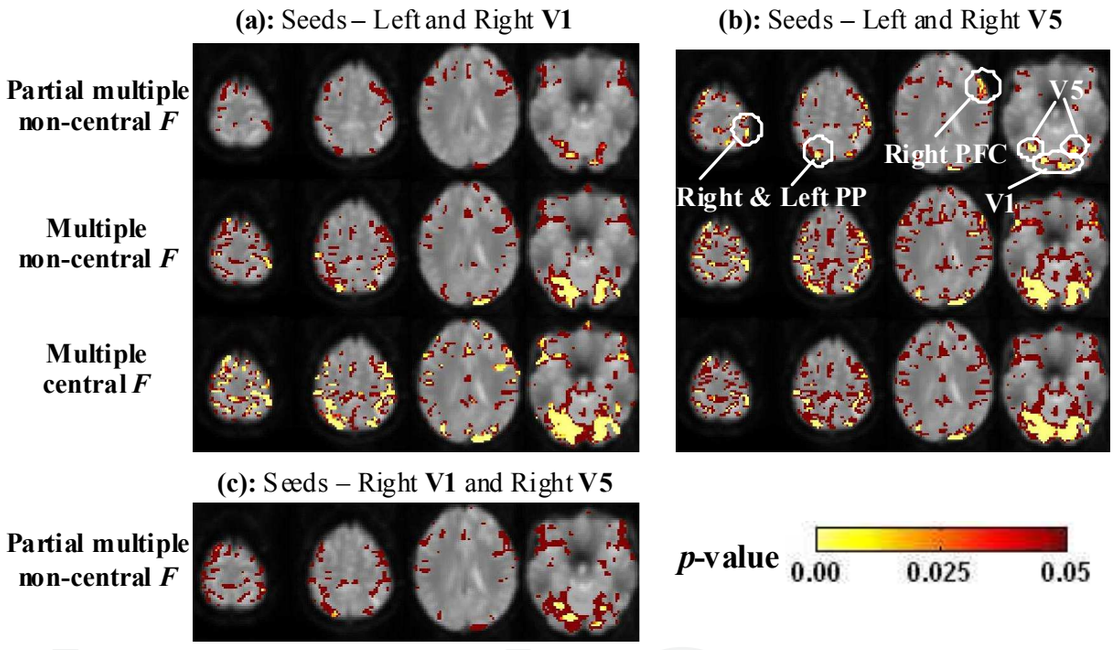

Fig. 4. Comparison and results of functional interaction maps for the real fMRI data.

Partial Correlation Effects – Multiple vs. Partial Multiple Correlations From Fig. 4 (a) and (b), we can see that using multiple correlation (2nd row), both stimulus-locked and task-induced networks are identified, with all the dorsal visual pathway involved regions shown as highly significant (yellow). However, using partial multiple correlation (1st row), since the stimulus-locked effects are accounted for, the network regions and their sizes are considerably reduced. Specifically: i) with V1 as seed regions (Fig. 4(a), 1st row), mainly the low level visual network is identified as highly significant implying task-induced coupling among the visual areas, such as V1 to V5; ii) withV5 as seed regions (Fig. 4(b), 1st row), the two PP and the right PFC regions are still shown as highly significant, suggesting the involved task-induced coupling of attention to motion modulation, after taking account of the stimulus-locked effects.

Effects of Multiple Seeds – Comparison using Both V1 and V5 as Seeds The partial multiple correlation results using V1 or V5 (Fig. 4 (a) and (b), 1st row) as seed regions have been illustrated in the above. With a combination of a V1 and a V5 as seed regions (Fig. 4(c)), using partial multiple correlation, we can not only detect the highly significant low level visual network regions (as in Fig. 4 (a) and (b), 1st row, last slice), but also identify the highly significant attention to motion modulation PP areas (as in Fig. 4(b), 1st row; also compare to Fig. 4(a), 1st row), achieving the combined effects of multiple seed regions involving both V1 and V5. Note that since only one V5 is used here, the region sizes and significance levels for the PP and right PFC are not as large as the ones using both V5 regions in Fig. 4(b) (1st row).

Effects of Spatial Noise – Comparison of Non-central and Central *F*-tests The results without taking the spatial noise correlations into consideration (central *F*-test) are shown in Fig. 4 (a) and (b), 3rd row, with many unjustified areas identified as functionally correlated with the V1 and/or V5 seed regions due to the spatial structured noise in the fMRI data.

4. Conclusions

A new statistical surface morphometry analysis method is presented and developed by using our moments-based permutation tests where the permutation distributions are accurately approximated through Pearson distributions for considerably reduced computation cost. The proposed hybrid strategy takes advantage of nonparametric permutation tests and parametric Pearson distribution approximation to achieve both accuracy/flexibility and efficiency. In addition, hybrid permutation schemes for both the conventional and bioequivalence tests are provided. Compared with the classical hypothesis tests, bioequivalence tests can screen out the non-notable differences and accurately locate the practical or physical significances. In real applications, either the standard or the bioequivalence hypothesis tests can be chosen, depending on the specific problems, i.e. whether the statistical and practical significance differences are negligible or not.

This chapter also presents a novel and general statistical framework for sensitive and reproducible estimation of brain networks from fMRI based on multiple and partial multiple correlation analyses and multiple seed regions, with the standard single-seed region analysis as the degenerate and a special case. Compared with using only a single seed, using multiple seeds can not only lead to more robust estimation of functional connectivity, but also more sensitive identification of functional co-activation networks or regions to multiple seeds that may not be detected in the single-seed method. The use of the partial multiple correlation has the interesting features of providing a convenient summary of conditional independences and hence of being more closely related to the direct functional interactions (i.e. effective connectivity) of the brain than marginal correlation.

The statistical and computational data analysis methods presented in this chapter can lead to precise and efficient recovery of structure and function of the human brain. The discovery of relationship between brain structure and function through combination of different modalities (e.g. sMRI and fMRI) will be a future direction, which may provide a unique perspective (Jiang et al., 2008) and thus further enhance our understanding of the complex system of the human brain.

5. Acknowledgments

This work was supported in part by the National Institute of Health under grant 1K25AG033725-01 from the National Institute of Aging awarded to Dr. Wang; Email: <ymw@illinois.edu>.

6. References

Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis*. 3rd edition, John Wiley and Sons Inc., New Jersey.

Ashburner, J. & Friston, K. J. (2000). Voxel-based morphometry: The methods. *Neuroimage*, 11: 805-821.

Bellec, P.; Perlbarg, V.; Jbabdi, S.; Pelegrini-Issac, M.; Anton, J.; Doyon, J. & Benali, H. (2006). Identification of large-scale networks in the brain using fMRI. *Neuroimage*, 29: 1231-1243.

Benjamini, Y. & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, 57: 289-300.

Benjamini, Y.; Krieger, A. M. & Yekutieli, D. (2006). Adaptive linear step-up false discovery rate controlling procedures. *Biometrika*. *93*: 491-507.

Brechbühler, C.; Gerig, G. & Kübler, O. (1995). Parametrization of closed surfaces for 3-D shape description. *Computer Vision, Graphics, and Image Processing,* 61: 154–170.

Brown, L. D.; Hwang, J. T. G. & Munk, A. (1997). An unbiased test for the bioequivalence problem. Ann. Stat. 25: 2345-2367.

Buchel, C. & Friston, K. J. (1997). Modulation of connectivity in visual pathways by attention: Cortical inferences evaluated with structural equation modeling and fMRI. *Cerebral Cortex*, 7: 768-778.

Buckner, R. L.; Head, D.; Parker, J.; Fotenos, A. F.; Marcus, D.; Morris, J. C. & Snyder, A. Z. (2004). A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume. *Neuroimage*, 23: 724-738.

Calhoun, V. D.; Adali, T.; Pearison, G. D. & Pekar, J. J. (2001). Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms. *Human Brain Mapping*, 13: 43-53.

Cao, J. & Worsley, K. J. (1999). The geometry of correlation fields with an application to functional connectivity of the brain. *Annals of Applied Probability*, 9: 1021-1057.

Chung, M. K.; Worsley, K. J.; Paus, T.; Cherif, C.; Collins, D. L.; Giedd, J. N.; Rapoport, J. L. & Evans, A. C. (2001). A unified statistical approach to deformation-based morphometry. *Neuroimage*, 14: 595-606.

Cressie, N. (1993). *Statistics for Spatial Data*. John Wiley and Sons, New York.

Cordes, D.; Haughton, V.; Carew, J. D.; Arfanakis, K. & Maravilla, K. (2002). Hierarchical clustering to measure connectivity in fMRI resting-state data. *Magnetic Resonance Imaging*, 20: 305–317.

Davatzikos, C.; Gene, A.; Xu, D. & Resnick, S. M. (2001). Voxel-based morphometry using the RAVENS maps: Methods and validation using simulated longitudinal atrophy. *Neuroimage*, 14: 1361-1369.

Friston, K. J.; Frith, C. D.; Liddle, P. F. & Frackowiak, R. S. J. (1993). Functional connectivity: The principal-component analysis of large (PET) data sets. *Journal of Cerebral Blood Flow & Metabolism*, 13: 5-14.

Hahn, J. & Shapiro, S. S. (1967). *Statistical Models in Engineering*. John Wiley. New York.

Hayasaka, S.; Peiffer, A. M.; Hugenschmidt, C. E. & Laurienti, P. (2007). Power and sample size calculation for neuroimaging studies by non-central random field theory. *Neuroimage*, 37: 721-730.

Horwitz, B. (2003). The elusive concept of brain connectivity. *Neuroimage*, 19: 466-470.

Hubert, L. (1987). *Assignment Methods in Combinatorial Data Analysis,* Marcel Dekker, Inc.

Huettel, S. A. ; Song, A. W. & McCarthy, G. (2004). *Functional Magnetic Resonance Imaging*. Sinauer Associates, Inc. Publisher.

Jack, C. R.; Petersen, R. C.; O'Brien, P. C. & Tangalos, E. G. (1997). MR-based hippocampal volumetry in the diagnosis of Alzheimer's disease. *Neurology*, 42: 183-188.

Jiang, T.; Liu, Y.; Shi, F.; Shu, N.; Liu, B.; Jiang J. & Zhou, Y. (2008). Multimodal magnetic resonance imaging for brain disorders: Advances and perspectives. *Brain Imaging and Behavior*, 2: 249-257.

Lindquist, M. A. (2008). The statisticsl analysis of fMRI data. *Statistical Science*, 23: 439-464.

Logan, B. R. & Rowe, D. B. (2004). An evaluation of thresholding techniques in fMRI analysis," *Neuroimage*, 22: 95-108.

Marrelec, G.; Krainik, A.; Duffau, H.; Pelegrini-Issac, M.; Lehericy, S.; Doyon, J. & Benali, H. (2006). Partial correlation for functional brain interactivity investigation in functional MRI. *Neuroimage*, 32: 228-237.

Martijn, P.; van den Heuvel, P. & Hulshoff Pol, H. E. (2010). Exploring the brain network: A review on resting-state fMRI functional connectivity. *European Neuropsychopharmacology*, 20: 519-534.

McIntosh, A. R. & Lobaugh, N. J. (2004). Partial least squares analysis of neuroimaging data: Applications and advances. *Neuroimage*, 23: S250-S263.

Nichols, T. E. & Holmes, A. P. (2001). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*, 15: 1–25.

Ogawa, S.; Lee, T. M.; Nayak, A. S. & Glynn, P. (1990). Oxygenation-sensitive contrast in magnetic resonance imaging of rodent brain of high magnetic fields. *Magnetic Resonance in Medicine*, 14: 68-78.

Pantazis, D., Leahy, R. M.; Nichols, T. E. & Styner, M. (2004). Statistical surface-based morphometry using a non-parametric approach. *IEEE International Symposium on Biomedical Imaging*, pp. 1283-1286.

Shen, L.; Saykin, A.; McHugh, T.; West, J.; Rabin, L.; Wishart, H.; Chung, M. & Makedon, F. (2005). Morphometric analysis of 3D surfaces: Application to hippocampal shape in mild cognitive impairment. *6th Int. Conf. on Computer Vision, Pattern Recognition and Image Processing in conjunction with 8th Joint Conference on Information Sciences*, pp. 699-702.

Styner, M.; Lieberman, J. A.; McClure, R. K.; Weinberger, D. R.; Jones, D. W. & Gerig, G. (2005). Morphometric analysis of lateral ventricles in schizophrenia and healthy controls regarding genetic and disease-specific factors. *Proc. Natl. Acad. Sci. U. S. A.*, 102: 4872-4877.

Styner, M.; Oguz, I.; Xu, S.; Brechbuehler, C.; Pantazis, D.; Levitt, J. J.; Shenton, M. E. & Gerig, G. (2006). Framework for the statistical shape analysis of brain structures using SPHARM-PDM, *Open Science Workshop at MICCAI 2006, Insight Journal*.

Symms, M.; Jager, H. R.; Schmierer, K. & Yoursry, T. A. (2004). A review of structural magnetic resonance neuroimaging. *Journal of Neurology, Neurosurgery & Psychiatry*, 75: 1235-1244.

Sun, F. T.; Miller, L. M. & D'Esposito, M. (2004). Measuring interregional functional connectivity using coherence and partial coherence analyses of fMRI data. *Neuroimage*, 21: 647-658.

Thompson, P. M.; Hayashi, K. M.; de Zubicaray, G. I.; Janke, A. L.; Rose, S. E.; Semple, J.; Hong, M. S.; Herman, D. H.; Gravano, D.; Doddrell, D. M. & Toga, A. W. (2004). Mapping hippocampal and ventricular change in Alzheimer disease. *Neuroimage*, 22: 1754-1766.

Wang, L.; Swank, J. S.; Glick, I. E.; Gado, M. H.; Miller, M. I.; Morris, J. C. & Csernansky, J. G. (2003). Changes in hippocampal volume and shape across time distinguish dementia of the Alzheimer type from healthy aging. *Neuroimage*, 20: 667-682.

Wang, Y. M. (2005). Modeling and nonlinear analysis in fMRI via statistical learning. In: *Advanced Image Processing in Magnetic Resonance Imaging*, Marcel Dekker International Publisher, L. Landini, V. Positano, and M. F. Santarelli, Eds., pp. 565-586.

Wang, Y. M.; Schultz, R. T.; Constable, R. T. & Staib, L. H. (2003). Nonlinear estimation and modeling of fMRI data using spatio-temporal support vector regression. In: *Proceedings of the 18th International Conference on Information Processing in Medical Imaging*, Lecture Notes in Computer Science, vol. 2732, pp. 647-659.

Wang, Y. & Staib, L. H. (2000). Boundary finding with prior shape and smoothness models. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22: 738-743.

Wang, Y. M. & Xia, J. (2009). Unified framework for robust estimation of brain networks from fMRI using temporal and spatial correlation analyses. *IEEE Transactions on Medical Imaging*, 28: 1296-1307.

Woolrich, M.; Jenkinson, M.; Brady, M. & Smith, S. M. (2004). Fully Bayesian spatio-temporal modeling of fMRI data. *IEEE Transactions on Medical Imaging*, 23:213-230.

Worsley, K. J. (1994). Local maxima and the expected Euler characteristic of excursion sets of χ^2, *F* and *t* fields. *Advances in Applied Probability*, 26: 13-42.

Worsley, K. J. & Friston, K. J. (1995). Analysis of fMRI time-series revisited – again. *Neuroimage*, 2: 173-181.

Worsley, K. J.; Chen, J-I; Lerch, J. & Evans, A. C. (2005). Comparing connectivity via thresholding correlations and SVD. *Philosophical Transactions of the Royal Society*, 360: 913-920.

Zhou, C. & Wang, Y. M. (2008). Hybrid permutation test with application to surface shape analysis. *Statistica Sinica*, 18(4): 1553-1568.

Zhou, C.; Wang, H. & Wang, Y. M. (2009). Efficient Moments-based Permutation Tests. *Advances in Neural Information Processing Systems*, 22: 2277-2285.

Biomedical Engineering, Trends, Research and Technologies

Edited by Dr. Sylwia Olsztynska

ISBN 978-953-307-514-3
Hard cover, 644 pages
Publisher InTech
Published online 08, January, 2011
Published in print edition January, 2011

This book is addressed to scientists and professionals working in the wide area of biomedical engineering, from biochemistry and pharmacy to medicine and clinical engineering. The panorama of problems presented in this volume may be of special interest for young scientists, looking for innovative technologies and new trends in biomedical engineering.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen