University of Michigan School of Public Health
The University of Michigan Department of Biostatistics Working Paper Series

Year 2008 *Paper* 76

Cluster Mass Inference Method via Random Field Theory

Hui Zhang* Thomas E. Nichols†
Timothy D. Johnson‡

*University of Michigan Biostatistics
†University of Michigan Biostatistics
‡University of Michigan Biostatistics, tdjtdj@umich.edu

http://biostats.bepress.com/umichbiostat/paper76

Cluster Mass Inference Method via Random Field Theory

Hui Zhang, Thomas E. Nichols, and Timothy D. Johnson

Abstract

Cluster extent and voxel intensity are two widely used statistics in neuroimaging inference. Cluster extent is sensitive to spatially extended signals while voxel intensity is better for intense but focal signals. In order to leverage strength from both statistics, several nonparametric permutation methods have been proposed to combine the two methods. Simulation studies have shown that of the different cluster permutation methods, the cluster mass statistic is generally the best. However, to date, there is no parametric cluster mass inference available. In this paper, we propose a cluster mass inference method based on random field theory (RFT). We develop this method for Gaussian images, extend it to Student's t-statistic images and investigate its statistical properties via simulation studies and real data. Simulation results show that the method is valid under the null hypothesis and demonstrate that it can be more powerful than the cluster extent inference method. Further, analyses with a single-subject and a group fMRI dataset demonstrate better power than traditional cluster size inference, and good accuracy relative to a gold-standard permutation test.

Cluster Mass Inference Method via Random Field Theory

Hui Zhang[1], Thomas E. Nichols[1,2,3], Timothy D. Johnson[1]

[1]*Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109-2029,USA*

[2]*Clinical Imaging Centre, GlaxoSmithKline, London, W12 0NN, United Kingdom*

[3]*Department of Clinical Neurology, FMRIB Centre, University of Oxford, Headington, OX3 9DU, United Kingdom*

December 20, 2007

Running title: RFT Cluster Mass Inference

Abstract

Cluster extent and voxel intensity are two widely used statistics in neuroimaging inference. Cluster extent is sensitive to spatially extended signals while voxel intensity is better for intense but focal signals. In order to leverage strength from both statistics, several nonparametric permutation methods have been proposed to combine the two methods. Simulation studies have shown that of the different cluster permutation methods, the cluster mass statistic is generally the best. However, to date, there is no parametric cluster mass inference available. In this paper, we propose a cluster mass inference method based on random field theory (RFT). We develop this method for Gaussian images, extend it to Student's *t*-statistic images and investigate its statistical properties via simulation studies and real data. Simulation results show that the method is valid under the null hypothesis and demonstrate that it can be more powerful than the cluster extent inference method. Further, analyses with a single-subject and a group fMRI dataset demonstrate better power than traditional cluster size inference, and good accuracy relative to a gold-standard permutation test.

Keywords: cluster mass, random field theory, Gaussian field, Gaussianized *t* image

1 Introduction

Cluster extent and voxel intensity are two widely used statistics in neuroimaging inference. Cluster extent is sensitive to spatially extended signals [6, 23], while voxel intensity is sensitive to focal, intense signals [5, 24]. Both can suffer from a lack of power for signals of moderate extent and intensity [9]. Furthermore, one does not generally know, a priori, whether the generated signal is large in extent, intensity or both. While some practitioners simply select the statistic that gives the most statistically significant test, this embodies a multiple testing problem and will result in inflated false positive error rates. An ideal test statistic would combine spatial extent and peak height intensity and would be sensitive to both without increasing the number of tests considered.

Poline *et al.* [23] (henceforth referred to as PWEF) develop a method which combines extent and intensity based on Gaussian random field theory (RFT). They derive the joint distribution of cluster extent and voxel-wise peak height intensity and make inference on minimum P value of a cluster extent test and a local maximum intensity test. However, their method is only applicable to Gaussian or approximately Gaussian images (e.g. a very large group analysis, or a single subject fMRI analysis).

Cluster mass, the integral of suprathreshold intensities within a cluster, naturally combines both signal extent and signal intensity. Initially suggested by Holmes [11], Bullmore *et al.* [2] used permutation to obtain cluster mass P values. Currently the cluster mass is default test statistic in the BAMM[1] and CAMBA[2] software, and is implemented in FSL's randomise[3] tool and in the SnPM[4] toolbox for SPM[5].

[1] http://www-bmu.psychiatry.cam.ac.uk/BAMM
[2] http://www-bmu.psychiatry.cam.ac.uk/software/
[3] http://www.fmrib.ox.ac.uk/fsl/randomise
[4] http://www.sph.umich.edu/ni-stat/SnPM
[5] http://www.fil.ion.ucl.ac.uk/spm

Hayasaka & Nichols [9] study the statistical properties of cluster mass along with a variety of other "combining methods" in the permutation testing framework. Among the combining methods they study are Tippet's method [15, 22] (minimum P values, used by PWEF) and Fisher's method (-2 $\times$ sum of $\ln$ P values). They conclude, through simulation studies and analyses of real data, that the nonparametric cluster mass method is generally more powerful than the other methods they investigate.

A strength of nonparametric inference methods is that they rely on fewer assumptions about the distributional form of the data. However, they require additional computational effort and are not very flexible. For example, the precise permutation scheme used depends on the experimental design and cannot be trivially determined from a design matrix. Nuisance covariates cannot be accommodated in general, as they induce null-hypothesis structure which violates exchangeability. Also, nonparametric methods cannot be used directly for single subject data analysis as a parametric autocorrelation model or wavelet transformation is needed to whiten the data. For all of these reasons, a parametric cluster mass inference method that can operate with a general linear model and deal with single subject analyses would be of great value.

In this paper we develop a theoretical distribution for the cluster mass statistic via Gaussian RFT. We generalize the work of PWEF, deriving the cluster mass statistic, extending the method to Gaussianized t data. We study the statistical size and power of our test on Gaussian and Gaussianized t image data through simulations and illustrate the method on two real data example, a single subject fMRI dataset and a group level fMRI data analysis with low degrees of freedom.

2 Materials and Methods

2.1 Cluster mass test theory

In a mass univariate data analysis, a general linear regression model (GLM)

$$Y_i = X\beta_i + \varepsilon_i \tag{1}$$

is fit for each voxel $i = 1, \ldots, I$, where Y_i is an $N \times 1$ vector of responses, X is a common $N \times q$ design matrix of predictors, β_i is a $q \times 1$ vector of unknown parameters and ε_i is a $N \times 1$ vector of random errors. Typically, at each voxel, errors are assumed to be independent and identically distributed $N(0, \sigma_i^2)$ random variates, though dependent errors can be accommodated [16]. The ordinary least squares estimator of β_i is $\widehat{\beta}_i = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}Y_i$, and of σ_i^2 is $\widehat{\sigma}_i^2 = e_i^{\mathrm{T}}e_i/\eta$, where $e_i = Y_i - X\widehat{\beta}_i$ and where η is the error degrees of freedom. Then the Student's *t*-statistic at voxel i is

$$T_i = c\widehat{\beta}_i \left(c(X^{\mathrm{T}}X)^{-1}c^{\mathrm{T}}\widehat{\sigma_i^2}\right)^{-1/2} \tag{2}$$

where c is a contrast of interest (row vector). We write the t-statistic image as $\mathbf{T} = \{T_i\}_{i=1}^I$.

Given cluster-forming threshold $u_c > 0$, the set of suprathreshold statistics $\{T_i : T_i > u_c\}_{i=1}^I$ is used to define clusters. Contiguous clusters are defined by a neighborhood scheme, typically 18 connectivity scheme on a three dimensional image.

Let L be the number of clusters found, with cluster ℓ having S_ℓ voxels (i.e. the cluster extent), $\ell = 1, 2, \ldots, L$. Further let I_ℓ be the set of voxel indices corresponding to cluster ℓ. The cluster mass, M_ℓ, of cluster ℓ is the summation of the suprathreshold intensities:

$$M_\ell = \sum_{i \in I_\ell} H_i \tag{3}$$

where $H_i = T_i - u_c$. Note that $M_\ell = S_\ell \bar{H}_\ell$ where $\bar{H}_\ell = \sum_{i \in I_\ell} H_i / S_\ell$ is the average suprathreshold intensity of cluster ℓ, showing cluster mass to be the product of the cluster extent and the average

suprathreshold intensity.

To use Random Field Theory results, we begin by assuming that the standardized error images, called the *component fields*, are discrete samplings of a continuous, smooth, stationary Gaussian random process. The component field for scan j is $\{\varepsilon_{ij}/\sigma_i\}_i$, where ε_{ij} is the error for scan j at voxel i. The component fields are assumed to follow a mean zero, unit variance multivariate Gaussian distribution. Stationarity implies that the spatial correlation is determined by an auto-correlation function that is homogeneous over space. The process is regarded as "smooth" if the autocorrelation function has two derivatives at the origin. Based on these assumptions, t image defined by (2) defines a Student's t random field.

While any univariate random variable can be transformed into a Gaussian variate, or *Gaussianized*, a Gaussianized t image may not resemble a realization of Gaussian random field. Randomness in $\hat{\sigma}_i^2$ reduces the smoothness of the statistic image relative to the component fields [24], as reviewed in Appendix B.1. However, Worsley et al [25] argues that when the t degrees of freedom exceed 120, the Gaussianized *t*-statistic can be regarded as a Gaussian Random Field. Hence we proceed by deriving results assuming $\mathbf{T}$ is a Gaussian image, but return to the issue of Gaussianization below.

The full derivation our the null distribution of the cluster mass statistic is given in Appendix B, but we sketch an overview of the result here. We begin by approximating the statistic image about a local maximum as a parabaloid, which allows cluster mass to be obtained a function of cluster extent, S_ℓ, and suprathreshold peak intensity, $H_\ell = \max\{H_i : j \in I_\ell\}$,

$$M_\ell \approx 2/(D+2) \times S_\ell \times H_\ell \tag{4}$$

where D is the dimension of the image. By assuming that the autocorrelation function of the image is proportional to a Gaussian probability density function, the distribution of H_ℓ conditional on H_ℓ can be found. At this point PWEF made a small excursion assumption, replacing peak height $u_c + H_\ell$ with u_c. We also make this assumption and create what we denote the $\mathcal{U}$ result, but also remove this assumption, deriving the $\mathcal{Z}$ result.

Finding the joint distribution and integrating out H_ℓ yields the final result, an expression for $\mathsf{P}(M_\ell > m)$, the uncorrected P-value for an observed cluster mass value of m. This requires two numerical integrations, one dependent on u_c, and one on m. In practice, for any given dataset, P-values for a grid m values can be pre-computed and interpolation used to find the P-value for an arbitrary value of m.

Note that the tail probability $\mathsf{P}(M_\ell > m)$ is an uncorrected P-value which does not account for searching over all clusters in the image. Uncorrected P-values are only appropriate for a single cluster that can be pre-identified before observing the data [4], a situation that rarely arises in practice. As detailed in Appendix B, the uncorrected P-values can be transformed into familywise-error corrected P-values which accounts for the chance of one or more false positive clusters anywhere in the image.

2.1.1 Student's *t*-statistic image

When the degrees of freedom are small, a Gaussian random field will not provide a good approximation for a Student's *t*-statistic image. In this case, a *t*-to-*z* transformation is performed via the probability integral transform. The transformed image, however, will be rougher than the component field, and so the roughness parameter is adjusted according to the degrees of freedom of the *t*-statistic image. Subsequently the above results are applied to find the uncorrected and corrected P values for the cluster mass statistic.

2.2 Simulation

To evaluate the accuracy of our cluster mass result, Equation (4), both 2D (256×256) and 3D ($64 \times 64 \times 30$) Gaussian noise images are simulated. In order to understand the influence of image roughness on the proposed statistic, each of the 10,000 independent Gaussian noise images are convolved with different isotropic Gaussian smoothing kernels. Kernel sizes 2, 4, 8, 10, and 12 voxels full width at half maximum (FWHM[6]) are used, and these sizes then directly determine $|\Lambda|$, the image roughness parameter. Two cluster forming thresholds are investigated ($u_c = 2.326$ and $u_c = 3.090$, corresponding to uncorrected $P = 0.01$ and $P = 0.001$, respectively). A nominal significance level of 0.05 is used for all inferences.

To evaluate the method on Gaussianized *t*-statistic images, 15 Gaussian noise images are simulated, mean-centered and divided by the voxel-wise standard error to produce 14 degrees-of-freedom t images. A *t*-to-*z* transformation is then applied to generate Gaussianized *t* images (see Appendix B.1).

To assess the power of our method, a spherically shaped signal (radius 1, 3, 5, 7, 10mm) with various uniform intensities (0.25, 0.5, 0.75, 1, 1.5, 2) is added to the center of Gaussian noise images. Power is measured as the probability of a true positive cluster, defined a significant cluster that contains one or more non-null voxels. The cluster extent inference methods are those from RFT [1] implemented in the Statistical Parametric Mapping (SPM2) [21] software.

One objective of the evaluations is to determine whether the $\mathcal{U}$ result, based on the small excursion approximation, or the $\mathcal{Z}$ result is more accurate. Since the derivation depends on the joint distribution of cluster mass and peak height, we examine the approximation accuracy of our results for this bivariate distribution with simulation. In addition to visualizing images of the predicted and

[6] Kernel standard deviation = FWHM / $\sqrt{8 \ln 2} \approx 0.4247$ FWHM

simulated densities for the $\mathcal{Z}$ and $\mathcal{U}$ results, we compute the Kullback-Leibler divergences [14], a measure of distance between two distributions. This allows a quantitative comparison between the two results.

The ultimate accuracy of the method depends on the marginal distribution of cluster mass. We compare the specificity and validity of the mass test statistic for the $\mathcal{U}$ and $\mathcal{Z}$ results, as well as cluster size P-values found with our derived cluster extent distribution and cluster extent P-values produced by SPM. We present results for both uncorrected and corrected P values to understand the performance of the method, though only the corrected P-values are of practical interest. The specificity and validity is gauged with plots of theory-based P-values versus Monte Carlo ("true") P-values, called P-P plots.

2.3 Applications

We demonstrate our cluster mass inference method on two fMRI data sets, one single subject and one group dataset

2.3.1 FIAC data

The first example is the Functional Imaging Analysis Contest (FIAC) example [8]. The experiment uses a sentence listening task, considering effects of different or same speakers and different or same sentences. We only consider the sentence effect "Different Sentence vs. Same Sentence": In each block, six sentences are read; in the "Different" condition six different sentences are read, while in "Same" condition the same sentence is repeated six times. For complete details see [8].

We use subject 3 ("func4"), block design data with 6mm FWHM smoothing, fit with a GLM

which produces a *t* statistic image with 179 degrees-of-freedom. Here we can assume that the *t* image reasonably approximates a Gaussian image and use the method directly on the t image. The cluster forming threshold is $P = 0.001$ uncorrected.

2.3.2 Working Memory Data

We also use a group level analysis with 12 subjects from a working memory experiment. Since the degrees of freedom is rather small (11), we perform a *t*-to-*z* transformation to generate a Gaussianized *t* image.

While the experiment considers different aspects of working memory, we only use the item recognition task. In the item recognition condition subjects are shown a set of five letters and, after a 2 second delay, shown a probe, to which respond "Y" if it was in the set, or "N" otherwise; in a control condition five "X"s are shown and the probe is just "Y" or "N" indicating the required response. For full details see Marshuetz *et al* [17].

A one-sample t-test is used to model the data. We use t-to-z transformation and a cluster defining threshold of $P = 0.01$ uncorrected ($t_{11} = 4.02$ or $z = 3.09$). The roughness parameter is adjusted by 1.3891 [11, 24] to account for increased roughness of the Gaussianized t statistic. In addition to parametric results in SPM, we also use SnPM to obtain nonparametric cluster extent and mass results. With 12 subjects there are $2^{12} = 4096$ possible sign flips of the contrast data to create a permutation distribution.

3 Results

3.1 Simulations

For the simulation studies, we only show results for a smoothness parameter of FWHM = 8 voxels, as the results are similar for the other smoothness parameters.

3.1.1 Accuracy of derived joint distribution

The top row of Figure 1 shows the true (simulated) joint distribution of cluster mass and peak height intensity, the $\mathcal{Z}$ result and the $\mathcal{U}$ result for 3D Gaussian noise images. The bottom row shows difference images of true and derived distributions for the $\mathcal{Z}$ and $\mathcal{U}$ results. The distributions are qualitatively similar, though for very small cluster masses and cluster height around 0.5 to 1.0, the two results tend to underestimate the truth; while for cluster mass between 0 and 50 and cluster heights between 0 and 0.5, the results can overestimate the truth. The Kullback-Leibler divergences are 1.285 for the $\mathcal{Z}$ result and 1.610 for the $\mathcal{U}$ result.

Figure 2 displays corresponding results for 3D Gaussianized *t* image. Again, there is little difference between the true distribution and the two results, and again the Kullback-Leibler divergence between the true distribution and the $\mathcal{Z}$ result is smaller than that between the true distribution and the $\mathcal{U}$ result (1.701 vs. 2.338). Thus, for both Gaussian images and Gaussianized images, the $\mathcal{Z}$ result appears to be superior to the $\mathcal{U}$ result.

[Figure 1, 2 are about here]

3.1.2 Accuracy of derived cluster mass null distribution

Figure 3 shows the P-P plots for 3D Gaussian null simulated data and Figure 4 3D Gaussianized t-statistic null simulated data. Both cluster mass (dot-dashed lines) and cluster size results (solid lines) are shown. For all of our derived methods, the $\mathcal{U}$ results are more conservative (the null will be rejected less often than nominal) than the $\mathcal{Z}$ results. The SPM cluster size results are also more conservative than the $\mathcal{Z}$ results for Gaussian null simulated data and the $\mathcal{U}$ results for Gaussianized t-statistic null simulated data. While our $\mathcal{Z}$ result for cluster size exhibits some anticonservativeness, overall the $\mathcal{Z}$ result of cluster mass is the least conservative method, while maintaining validity over most of the range of probabilities included in this simulation study.

[Figure 3, 4 are about here]

Figure 5 shows the Type I error rates for a 3D Gaussianized t image with 14 degrees of freedom with various smoothness parameters (FWHM) and cluster defining thresholds. The figure shows that the $\mathcal{Z}$ cluster mass result provides better results for high thresholds and large FWHM than for low threshold and low FWHM. For corrected P values, this result is valid for all levels of smoothing studied, whereas the $\mathcal{Z}$ result of cluster extent is, by and large, invalid. Furthermore, the $\mathcal{Z}$ cluster mass corrected P-values—those that are used in practice—are always closer to the nominal significance level when correcting for multiple comparisons.

[Figure 5 is about here]

3.1.3 Power comparisons

Having found our own cluster extent result to be invalid, we compare the power of our $\mathcal{Z}$ cluster mass result to SPM's cluster extent result. Table 1 lists simulated power for the cluster extent (SPM) and cluster mass ($\mathcal{Z}$). As expected, for a given intensity, the power increases with signal intensity, and, for a given radius, power increases as the signal intensity increases. When the image smoothness is low (FWHM ≤ 4 voxels), SPM cluster extent generally provides better power than the $\mathcal{Z}$ mass result. However, for greater smoothness (FWHM ≥ 8 voxels), the $\mathcal{Z}$ result is more powerful than SPM, regardless of signal extent or signal intensity.

[Table 1 is about here]

3.2 Real Data Evaluations

The FIAC data results show the method's performance at high degrees-of-freedom, while the working memory data assess the method using Gaussianization of the t image.

3.2.1 FIAC data

The estimated smoothness of the component fields based on the residuals is [2.4964 2.3599 1.7525] voxel FWHM with 27,862 $3.0 \times 3.0 \times 4.0\text{mm}^3$ voxels. Figure 6 shows the maximum

intensity projection of the all clusters found with a $P = 0.001$ threshold. Table 2 provides the values of cluster extent, suprathreshold peak height intensity and cluster mass for each cluster, as well as the P-values, all sorted by peak height. The first three clusters have corrected significance with cluster mass, while peak height and cluster extent only find one cluster significant each. The uncorrected significances show that if a cluster is significant by any of the three methods, it is significant by cluster mass. Again, while we do not advocate use of uncorrected inferences, this demonstrates the relative sensitivity of the method.

[Figure 6, Table 2 are about here]

3.2.2 Working Memory Data

The estimated smoothness is [4.8611 6.4326 6.6156] voxel FWHM with 122,659 $2.0 \times 2.0 \times 2.0$ voxels. Figure 7 shows the all of the clusters found with a $P = 0.001$ cluster-forming threshold. Table 3 compares our RFT cluster mass results to an equivalent permutation method. Our RFT method finds the five largest clusters significant, as does the RFT cluster size statistic. Notable is the close correspondence between the RFT P-values and the permutation P-values.

[Figure 7, Table 3 are about here]

4 Discussion

Although cluster mass inference with nonparametric permutation has been found to be a quite sensitive inference method for neuroimaging data [9], permutation is computationally intensive, not a very flexible modeling framework. We propose a new theoretical cluster mass inference method for Gaussian images and Student's *t*-statistic images, based on Gaussian RFT. Our simulation studies show that our derived null distribution is accurate, and performs well not only for Gaussian images, but also for Student's *t*-statistic image. Like other RFT methods, our results depend only on the smoothness and the volume of the image. While we did not find close form results for the P-value

for an arbitrary mass value, these are quickly pre-computed on a grid of mass values which can then be interpolated.

Our evaluations of the test's specificity reveal that the proposed cluster mass inference method works best when the image is sufficiently smoothed, at least 4 voxel FWHM, and ideally for larger smoothness parameters (FWHM $\geq$ 8 voxels). However, our real data evaluations found it to perform as good or better than parametric cluster size inference, even though image smoothness was only about 2 voxels FWHM in the single subject dataset.

Consistent with its competitor, the nonparametric cluster mass inference method, our theoretical cluster mass inference statistic generally has better power than either the cluster extent inference statistic or the voxel intensity statistic, alone. This is especially true when the cluster extent and the suprathreshold peak height intensity are moderately sized. More remarkable, is that despite a large number of assumptions and a sequence of approximations, our RFT cluster mass P-values are so close to the permutation results which have very few assumptions.

The Gaussianization of t images is a shortcoming of the method, but it is not an uncommon strategy. The FSL [20] software has always (as of version 4.0) used Gaussianization of t and F images. While the SPM software has abandoned Gaussianization for voxel-wise inference ever since SPM99, its cluster extent inference has always (as of SPM5) used Gaussian and not t random field results cluster extent P-values and currently neglects the smoothness adjustment described in Appendix B.1.

Although the proposed cluster mass inference method has many good statistical properties, it has its limitations. When we derive the formulas for the marginal distribution of cluster mass, we assume that the shape of a cluster above a certain threshold is approximated by a parabaloid. This

assumption is rational for a Gaussian image that has been convolved with a Gaussian smoothing kernel. However, for real data, this assumption may be too strong, even after smoothing the data. For example, we may have a large flat cluster with only one voxel of high intensity. The activated regions may also have other shapes that are not well approximated by a parabaloid. In addition, we use a Gaussian shaped correlation function to simplify the variance in the derivation. We also assume that we have stationary fields, though an extension to accommodate local variation in smoothness [10] may be possible.

Finally we note that, while both real data examples were fMRI, the method makes no assumptions about the modality and should operate well with PET and other types of imaging data.

5 Acknowledgments

This work is funded by the US NIH: grant number 5 R01 MH069326-04. The authors would like to thank Dr. Christy Marshuetz and the FIAC group for providing the data sets used in this paper.

References

[1] R.J. Adler. *The Geometry of Random Fields*. Wiley, New York, 1981.

[2] E.T. Bullmore, J. Suckling, S. Overmeyer, S. Rabe-Hesketh, E. Taylor, and M.J. Brammer. Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain. *IEEE Trans. Med. Imaging*, 18:32–42, 1999.

[3] J. Cao and K.J. Worsley. *Applications of random fields in human brain mapping, in: Moore, M.(Ed), Spatial Statistics: Methodological Aspects and Applications, Springer lecture Notes in Statistics*, volume 159, pages 169–182. Springer, New York, 2001.

[4] K.J. Friston. Testing for anatomically specified regional effects. *Human Brain Mapping*, 5:133–136, 1997.

[5] K.J. Friston, C.D. Frith, P.F. Liddle, and R.S.J. Frackowiak. Comparing functional (PET) images: The assessment of significant change. *J. Cereb. Blood Flow Metab.*, 11:690–699, 1991.

[6] K.J. Friston, A. Holmes, J.-B Poline, C.J. Price, and C.D. Frith. Detecting activations in pet and fmri: levels of inference and power. *NeuroImage*, 4:223–235, 1996.

[7] RD Gordon. Values of mill's ratio of area to bounding ordinate of the nor mal probability integral for large values of the argument. *Annals of Mathematical Statistics*, 12:364–366, 1941.

[8] Madic Group. Functional Imaging Analysis Contest (FIAC). `http://www.madic.org/fiac/how_to_participate.html`, 2005.

[9] S. Hayasaka and T.E. Nichols. Combining voxel intensity and cluster extent with permutaion test framework. *Neuroimage*, 23:54–63, 2004.

[10] Satoru Hayasaka, K. Luan Phan, Israel Liberzon, Keith J. Worsley, and Thomas E. Nichols. Nonstationary cluster-size inference with random field and permutation methods. *Neuroimage*, 22(2):676–687, 2004.

[11] Andrew Holmes. *Statistica Issues in functional Brain Mapping*. PhD thesis, University of Glasgow, 1994.

[12] A.P. Holmes, R.C. Blair, J.D.G. Watson, and I. Ford. Nonparametric analysis of statistic images from functional mapping experiments. *J. Cereb. Blood Flow Metab.*, 16(1):7–22, 1996.

[13] S Kiebel, J.B. Poline, K.J. Friston, A Holmes, and K.J. Worsley. Robust smoothness estimation in statistical parametric maps using standardized residuals from the general linear model. *Neuroimage*, 10:756–766, 1999.

[14] S. Kullback and R. A. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22:79–86, 1951.

[15] N.A. lazar, B. Luna, J.A. Sweeney, and W.F. Eddy. Combining brains: a survey of methods for statitical pooling of information. *Neuroimage*, 16:538–550, 2002.

[16] W.-L. Luo and T.E. Nichols. Diagnosis and exploration of massively univariate fmri models. *NeuroImage*, 19(3):1014–1–32, 2003.

[17] C. Marshuetz, E.E. Smith, J. Jonides, J. DeGutis, and T.L. Chenevert. Order information in working memory:fMRI evidence for partietal and prefrontal mechanism. *J. Cogn. Neurosci.*, 12(S2):130–144, 2000.

[18] T.E. Nichols and S. Hayasaka. Controlling the familywise error rate in functional neuroimaging: A comparative review. *Statistical Methods in Medical Research*, 12(5):419–446, 2003.

[19] T.E. Nichols and A.P. Holmes. Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human Brain Mapping*, 15:1–25, 2002.

[20] Analysis Group of Functional Magnetic Resonance Imaging of the Brain(FMRIB). Fmrib software library. `http://www.fmrib.ox.ac.uk/fsl/`. Oxford, London, UK.

[21] Wellcome Department of Imaging Neuroscience. Statistical parametric mapping. `http://www.fil.ion.ucl.ac.uk/spm/`. University of College London, London, UK.

[22] F. Pesarin. *Multivariate Permutation Tests*. Wiley, New York, 2001.

[23] J.B. Poline, K.J. Worsley, A.C. Evans, and K.J. Friston. Combining spatial extent and peak intensity to test for activations in functional imaging. *Neuroimage*, 5:83–96, 1997.

[24] K.J. Worsley, A.C. Evans, S. Marrett, and P. Neelin. Three-dimensional statistical analysis for CBF activation studies in human brain. *J. Cereb. Blood Flow Metab.*, 12:900–918, 1992.

[25] K.J. Worsley, S. Marrett, P. Neelin, A.C. Vandal, K.J. Friston, and A.C. Evans. A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*, 4:58–73, 1996.

Appendix

A P values

For permutation test analyses of fMRI data [12, 19], there are two types of P values, uncorrected P values (not adjusted for multiple comparisons) and corrected P values (adjusted for multiple comparisons). Uncorrected voxel-wise permutation P-values are just those from a standard univariate permutation test. Uncorrected cluster-wise permutation P-values require additional assumptions. Based on an assumption of stationarity, that the distribution of cluster statistics (e.g. size, mass, local peak height, etc) does not vary with space, cluster statistics can be pooled over space. For each permutation, the set of observed cluster statistics are added to the permutation distribution, creating a distribution with many more entries than the number of permutations. Uncorrected cluster inferences are then obtained by reference to this distribution.

For familywise error (FWE) corrected P-values, the distribution of the maximal cluster statistic, searched over the image, is created. This process produces one (maximal) cluster statistic per permutation, yielding a permutation with the standard number of elements. The corrected P-value is the proportion of permutation elements as large or larger than that the cluster statistic considered. No assumption of stationarity is required, though if there is severe non-stationarity, the smoothest regions of the image will contain the greatest risk of false positives, and rough regions will lack sensitivity to detect small clusters (which, relative to the local smoothness, may be unusually large). For more on the issue of spatially varying smoothness, see [10].

If cluster statistics are marked as significant only when FWE-significant at 0.05, there is then 95% confidence of no false positive clusters anywhere in the image. For more on FWE see [18].

B Derivation of Null Distribution of Cluster Mass

Our derivation of the distribution of cluster mass follows that of Poline *et al.* [23] (PWEF) with several departures. A rough outline of the derivation is as follows:

1. A second order Taylor series approximates the statistic image at a local maximum as a parabaloid, determined by peak height and curvature about the maximum.

2. The geometry of a parabaloid gives cluster extent and mass as a function of peak height and the curvature (Jacobian determinant).

3. Distribution of the curvature, conditional on peak height, is found using an assumption of a Gaussian autocorrelation function.

4. Combining two previous results relates extent and mass, conditional on peak height, to a χ^2 distribution. A bias correction is made using the expected Euler characteristic.

5. PWEF proceeds with a small excursion assumption; we produce a pair of results, with and without this assumption.

6. Joint distribution of mass and height are found and marginalized to produce final mass result.

Let $\mathbf{Z}(\mathbf{x})$ be a D-dimensional Gaussian image, with

$$\begin{aligned} \mathsf{E}(\mathbf{Z}(\mathbf{x})) &= 0, \\ \mathsf{Var}(\mathbf{Z}(\mathbf{x})) &= 1, \\ \mathsf{Var}(\nabla \mathbf{Z}(\mathbf{x})) &= \Lambda \end{aligned}$$

for all $\mathbf{x} \in \Omega \subset \Re^D$ in the image volume, where ∇ is the gradient operator and Λ is the $D \times D$ matrix which parameterizes roughness. We assume the process is smooth, in that $\nabla^2 \rho(\mathbf{0})$ exists, where $\rho(\cdot)$ is the autocorrelation function and ∇^2 is the Hessian operator.

Without loss of generality, suppose there exists a local maximum at $\mathbf{x} = \mathbf{0}$, and consider the approximating parabaloid from a second order Taylor series about $\mathbf{x} = \mathbf{0}$

$$\mathbf{W}(\mathbf{0}) = \mathbf{Z}(\mathbf{0}) + \mathbf{x}^{\mathbf{T}}(\nabla^2\mathbf{Z}(\mathbf{0}))\mathbf{x}^{\mathbf{T}}/2$$

Suppressing the spatial index, let $Z \equiv \mathbf{Z}(\mathbf{0})$, and denote $J = |-\nabla^2\mathbf{Z}(\mathbf{0})|$ the negative Jacobian determinant.

For a cluster-defining threshold u_c, let $H = Z - u_c$ be the suprathreshold magnitude (note that we suppress the ℓ subscript used in the body of the paper). Then the geometry of the approximating parabaloid gives cluster extent as

$$S = a2^{D/2}H^{D/2}J^{-1/2} \tag{5}$$

where $a = \pi^{D/2}/\Gamma(D/2+1)$ is the volume of the unit sphere, and mass as

$$M = 2SH/(D+2). \tag{6}$$

Conditional on H, PWEF shows that another Taylor series yields

$$\ln J|H \approx \ln|\Lambda| + D\ln(H+u_c) + \eta, \tag{7}$$

where η is mean zero Gaussian with variance[7]

$$\mathsf{Var}(\eta|H) \;=\; \left[\mathrm{tr}\left((\Lambda^{-1}\otimes\Lambda^{-1})\rho^{(4)}(\mathbf{0})\right) - D^2\right]/(H+u_c).$$

While this expression is quite involved, if we assume that ρ is proportional to a Gaussian pdf, it simplifies to $\mathsf{Var}(\eta|Z) = 2D/(H+u_c)^2$. Subsequently we will need $J^{-1/2}$, and so write the exponentiated and powered equation (7) as $J^{-1/2} \approx |\Lambda|^{-1/2}(H+u_c)^{-D/2}\exp(\eta/2)^{-1}$. However, as in PWEF, we find that numerical evaluations of the final result are poor when η is assumed to be Gaussian (results not shown). We instead linearize the exponential,

$$J^{-1/2}|H \;\approx\; |\Lambda|^{-1/2}(H+u_c)^{-D/2}(1+\eta/2)^{-1} \tag{8}$$

[7]Note there is a typo in the PWEF paper's equation (8), where $2Z$ should in fact be just Z, or $H+u_c$ as we have written.

and approximate $1+\eta/2$ with η', where $\nu\eta'$ is χ^2_ν variate. Matching the second moments of $1+\eta/2$ and η' gives $\nu = 4(H+u_c)^2/D$. Combining with Equations (8) and (5) yields

$$S|H \approx a2^{D/2}|\Lambda|^{-1/2}(H+u_c)^{-D/2}H^{D/2}\eta'^{-1}. \tag{9}$$

B.0.3 The $\mathcal{U}$ result

PWEF proceed by using a small excursion approximation, that H is small relative to u_c, replacing $H+u_c$ with u_c. With this change, and marginalizing out H, the expected cluster extent can be found as

$$\mathsf{E}_{\mathcal{U}}(S) = (2\pi)^{D/2}\,|\Lambda|^{-1/2}\,u_c^{-D}. \tag{10}$$

However, accurate results using the expected Euler Characteristic [1] give

$$\mathsf{E}_{\mathrm{EC}}(S) = (2\pi)^{D/2}\,|\Lambda|^{-1/2}\,u_c^{-(D-1)}\,(1-\Phi(u_c))/\phi(u_c) \tag{11}$$

where Φ is the standard Gaussian CDF and ϕ is the standard Gaussian PDF. Hence, the approximation for $S|H$ is scaled by

$$c_{\mathcal{U}} = \frac{\mathsf{E}_{\mathrm{EC}}(S)}{\mathsf{E}_{\mathcal{U}}(S)} = u_c(1-\Phi(u_c))/\phi(u_c). \tag{12}$$

As a side note, this is Mill's ratio [7] scaled by u_c, which will have $c_{\mathcal{U}}$ converging to 1 from below for large u_c.

The bias-adjusted result is

$$M|H \approx ac_{\mathcal{U}}2^{D/2+1}(D+2)^{-1}\,|\Lambda|^{-1/2}\,u_c^{-D/2}\,H^{D/2+1}\,\eta'^{-1}, \tag{13}$$

which is a scaled inverse χ^2 random variable with ν degrees of freedom and scale parameter

$$q_{\mathcal{U}}(H) = ac_{\mathcal{U}}2^{D/2+1}(D+2)^{-1}\,|\Lambda|^{-1/2}\,u_c^{-D/2}\,H^{D/2+1}$$

The marginal distribution of H is approximately exponential with mean $1/u_c$ [1], and thus the joint pdf of M and H is

$$f_{\mathcal{U}}(M,H) \approx \frac{(q_{\mathcal{U}}(H)\nu/2)^{\nu/2}}{\Gamma(\nu/2)}\frac{\exp\left[(q_{\mathcal{U}}(H)\nu/2)/M\right]}{M^{\nu/2+1}}\exp[-u_cH]u_c \tag{14}$$

for $M, H > 0$. The uncorrected P-value for cluster mass is then found with

$$P_{\mathcal{U}}(M > m) \approx \int_m^\infty \int_0^\infty f_{\mathcal{U}}(M, H)\, dH\, dM$$

using numerical integration over a fine grid.

B.0.4 The $\mathcal{Z}$ result

We repeat the preceding without the small excursion approximation. We call this the $\mathcal{Z}$ result, since $Z = H + u_c$ is left as is. Returning to (9) and marginalizing out H we get

$$\mathsf{E}_{\mathcal{Z}}(S) = a 2^{D/2} \left|\Lambda\right|^{-1/2} \mathsf{E}\{[H/(H + u_c)]^{D/2}\} \tag{15}$$

where the final term must be found numerically for a particular u_c. This provides the bias adjustment term

$$c_{\mathcal{Z}} = \mathsf{E}_{\mathrm{EC}}(S)/\mathsf{E}_{\mathcal{Z}}(S). \tag{16}$$

This provides an approximation for $M|H$ as a scaled inverse χ^2 random variable with ν degrees of freedom and scale parameter

$$q_{\mathcal{Z}}(H) = a c_{\mathcal{Z}} 2^{D/2+1} (D + 2)^{-1} \left|\Lambda\right|^{-1/2} (H + u_c)^{-D/2}\, H^{D/2+1}$$

and joint pdf of M and H of

$$f_{\mathcal{Z}}(M, H) \approx \frac{(q_{\mathcal{Z}}(H)\nu/2)^{\nu/2}}{\Gamma(\nu/2)} \frac{\exp\left[(q_{\mathcal{Z}}(H)\nu/2)/M\right]}{M^{\nu/2+1}} \exp[-u_c H] u_c. \tag{17}$$

As before, the uncorrected P-value for cluster mass is then found with

$$P_{\mathcal{Z}}(M > m) \approx \int_m^\infty \int_0^\infty f_{\mathcal{Z}}(M, H)\, dH\, dM$$

using numerical integration over a fine grid.

B.0.5 Corrected P-values

The uncorrected P-values can be transformed into family-wise error (FWE) corrected P-values with either a Bonferroni correction for the expected number clusters or the Poisson clumping heuristic [1, 3, **?**]. We opt for the later, as it provides a continuous transformation between uncorrected and corrected P-values.

A FWE corrected P-value accounts for the chance of the maximal statistic exceeding that actually observed. Assuming the clusters arise a Poisson process, this P-value is found as

$$\mathsf{P}(\max_\ell M_\ell > m) \approx 1 - \exp\{-\mathsf{E}(L) \cdot \mathsf{P}(M_\ell > m)\}, \tag{18}$$

where $E(L)$ is the expected number of clusters in the image. For moderate thresholds u_c Euler characteristic will count the number of clusters, and hence we approximate $E(L) \approx \mathsf{E}_{\mathrm{EC}}(L)$. The most accurate results for $\mathsf{E}_{\mathrm{EC}}(L)$ depends on the dimension and the topology of the search region [25]. For a 3D, approximately spherical search region

$$\mathsf{E}_{\mathrm{EC}}(L) = \lambda(\Omega)|\Lambda|^{1/2}(2\pi)^{-2}(u_c^2 - 1)\exp[-u_c^2/2]; \tag{19}$$

where $\lambda(\Omega)$ is the volume of the search region. In addition, for a high threshold u_c, the number of clusters above the threshold will be approximated by [1, 23]

$$E(L) = \lambda(\Omega)|\Lambda|^{1/2}(2\pi)^{-2}u_c^2\exp[-u_c^2/2]$$

B.0.6 Smoothness Estimation & Λ

The preceding results depend on the roughness of the component random fields, as parameterized by $|\Lambda|$. Worsley et al. [24] proposed re-expressing this as the FWHM Gaussian kernel required smooth an independent random field into one with roughness Λ. Assuming the smoothing is aligned with the major axes of the image, this relationship is

$$|\Lambda|^{1/2} = \frac{(4\ln 2)^{D/2}}{\prod_d \mathrm{FWHM}_d}$$

where FWHM$_d$ is the smoothness in the d-th dimension. If the smoothness is not known, $|\Lambda|^{1/2}$ can be estimated from the residual images of a general linear model [13]

B.1 Student's *t-image*

A Student's *t*-statistic image with small degrees of freedom is not well approximated by Gaussian random field [25], however our results assume the statistic image is Gaussian. Gaussianization of t statistic images will match the univariate distribution at each voxel, but the converted image will have greater roughness than the component fields. Worsley et al. and Holmes showed that if the roughness of a Gaussian image is Λ, the roughness for a Student's *t*-statistic image can be estimated by $\Lambda_T = \lambda_n \Lambda$, where $n > 4$ is the number of scans used to generate the t image and λ_n is the correction factor [11, 24]. When applying our method to Gaussianized data we adjust Λ accordingly.

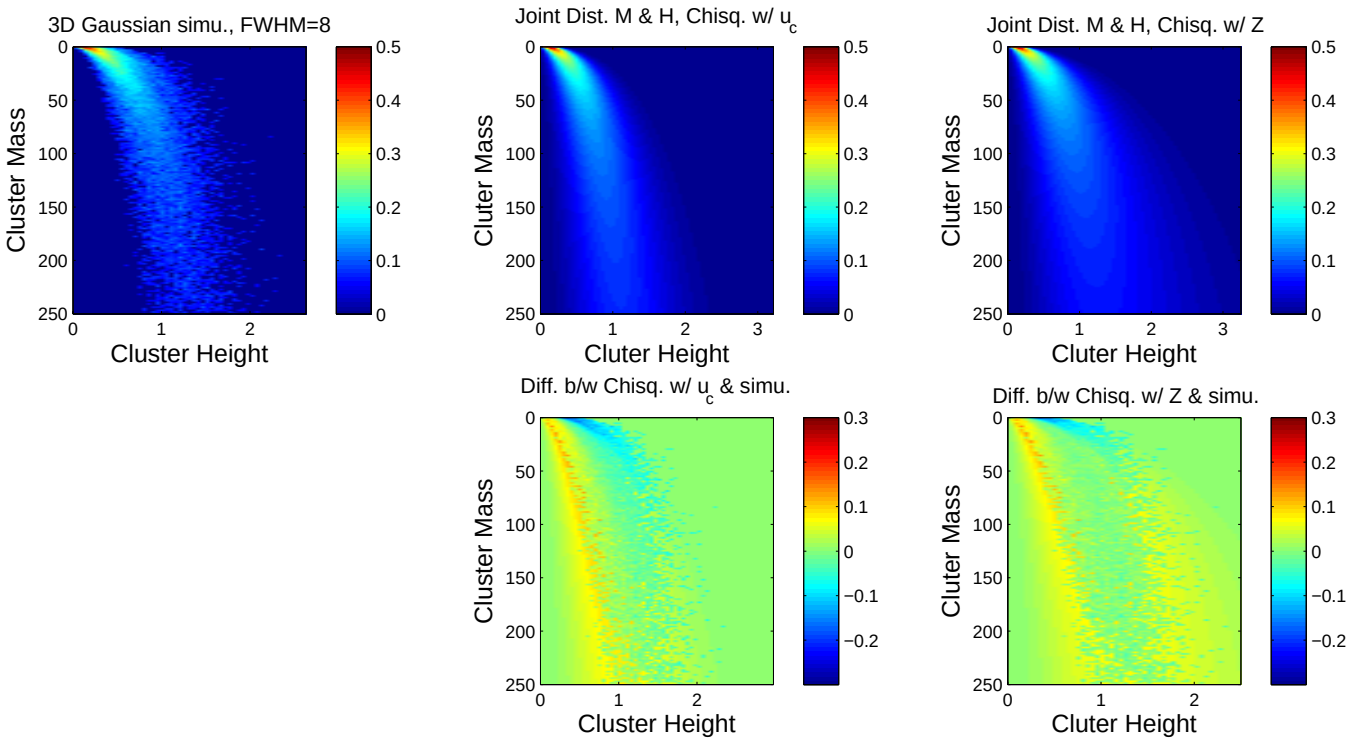

Figure 1: Comparison of true and theoretical joint distributions of cluster mass and peak height intensity, for Gaussian images. On top left is the true distribution obtained from simulation, on the top middle is the $\mathcal{U}$ result and on the top right is the $\mathcal{Z}$ result. Below each of the theoretical results is the true minus estimated distributions. While only an intermediate result, the agreement is reasonable, with better performance obtained with the $\mathcal{Z}$ result. All distributions are transformed by the fourth root to improve visualization. Unless otherwise noted, simulation settings used in the figures are: u_c = 2.3263 (p=0.01), $64 \times 64 \times 30$ image at FWHM 8 voxels.

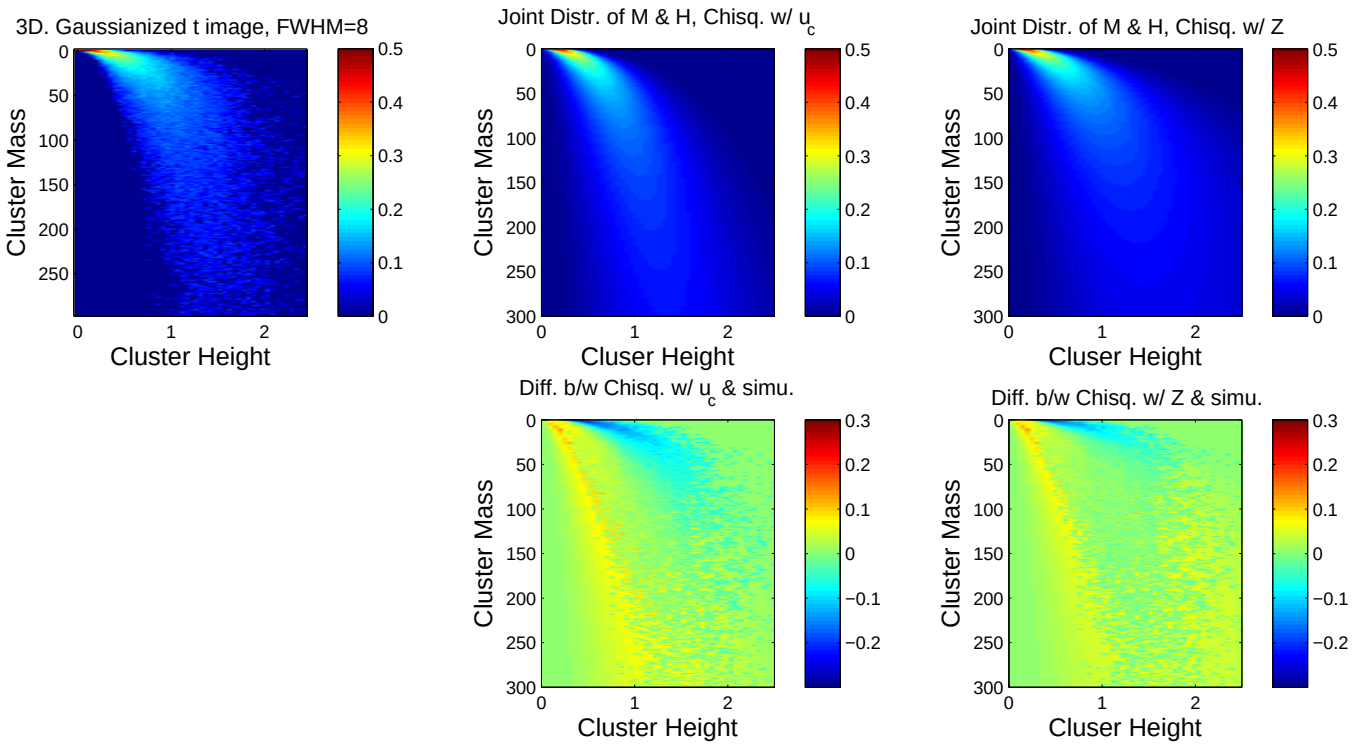

Figure 2: Comparison of true and theoretical joint distributions of cluster mass and peak height intensity, for Gaussianized t_{14} images. Same format as in Figure 1. Again the agreement between simulated truth and derived theoretical result is good, with a closer match seen with the $\mathcal{Z}$ result.

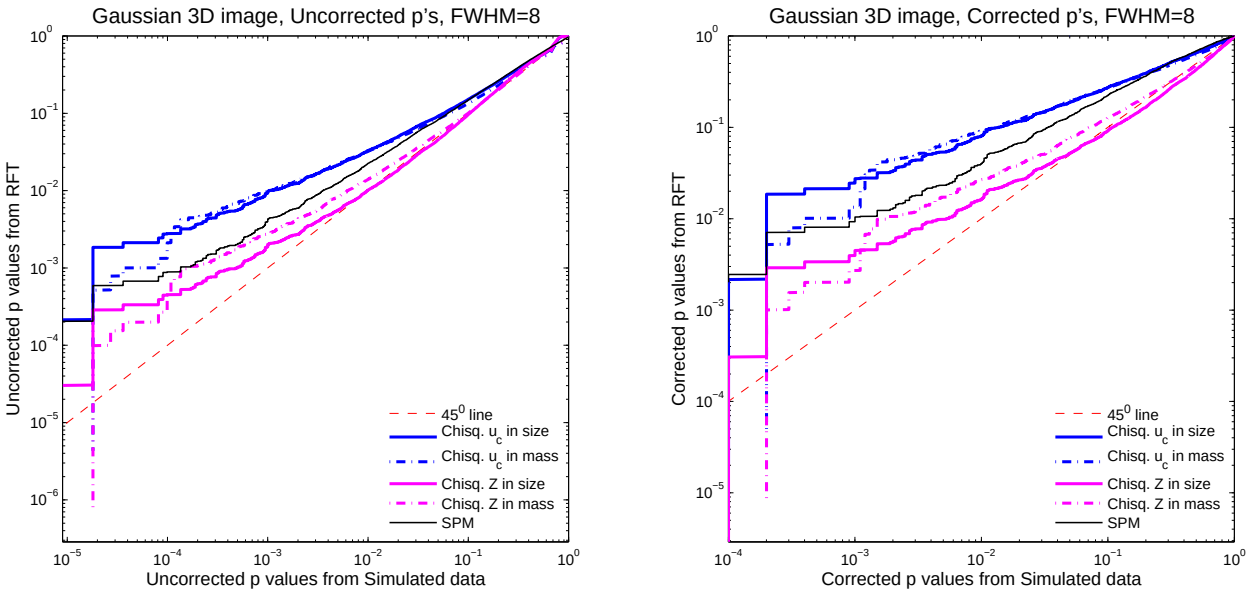

Figure 3: Monte Carlo simulation P-values versus theoretical P-values for uncorrected and corrected P-values with Gaussian images. Values in the plot above the identity indicate conservative performance, below the identity invalid performance. Our $\mathcal{Z}$ cluster mass method exhibits slight conservative performance, but much less conservative than the other methods.

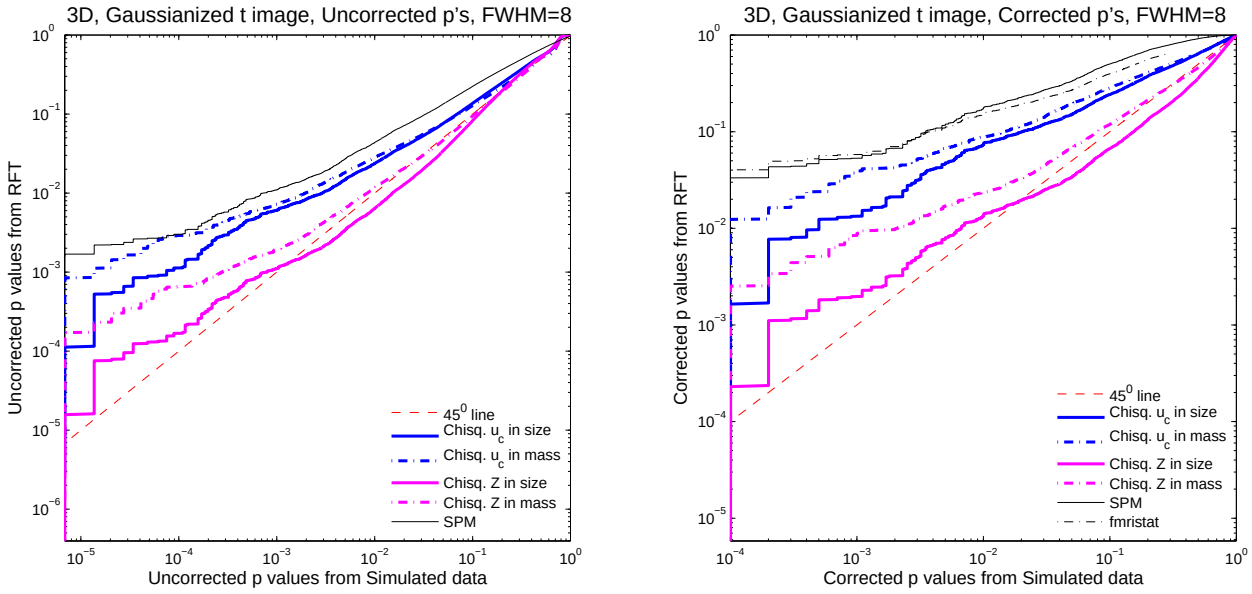

Figure 4: Monte Carlo simulation P-values versus theoretical P-values for uncorrected and corrected P-values with Gaussianized t_{14} images. Despite Gaussianization, our $\mathcal{Z}$ cluster mass method provides close to exact performance, and less conservative performance than other methods.

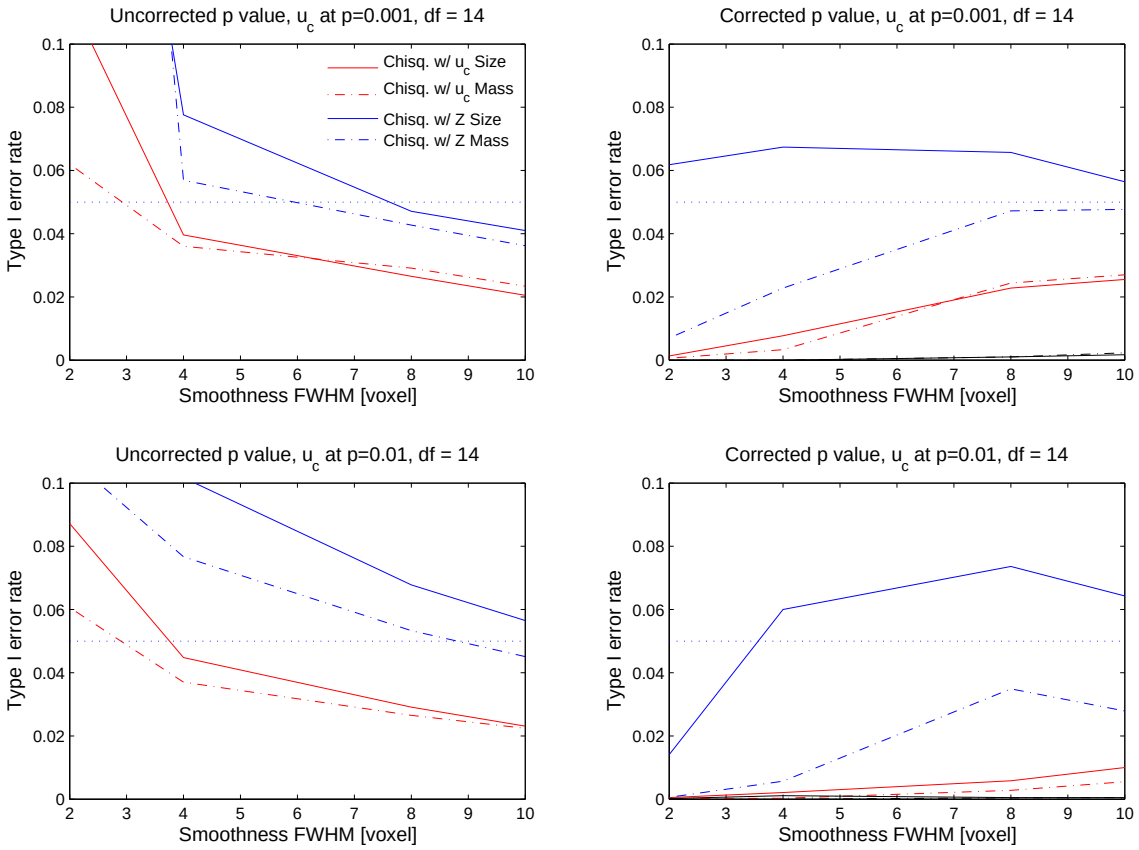

Figure 5: Type I error rate for Gaussianized t images, for both $P = 0.01$ and $P = 0.001$ cluster-forming thresholds, with different smoothness. While uncorrected P-values perform poorly under low smoothness, our $\mathcal{Z}$ cluster mass method has the corrected P-values are closest to the nominal $\alpha = 0.05$ level without being invalid.

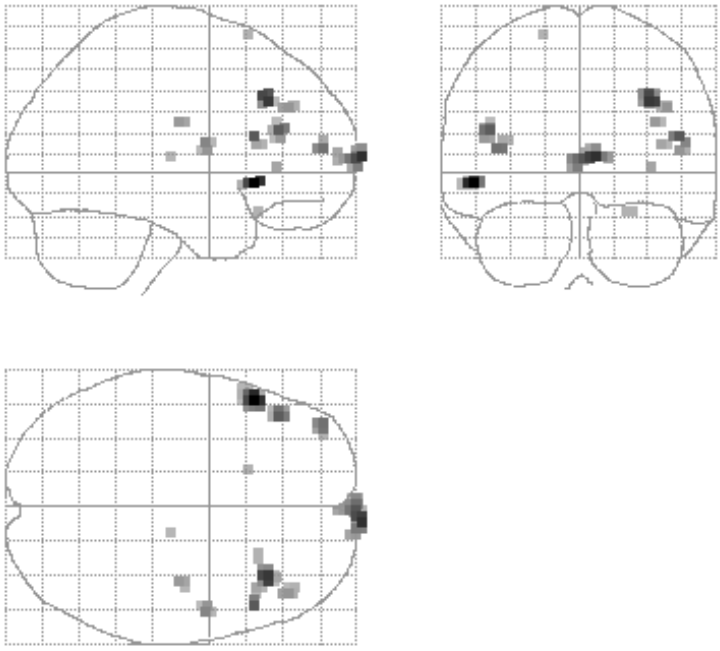

Figure 6: Results for “sentence” effect in FIAC single subject data.

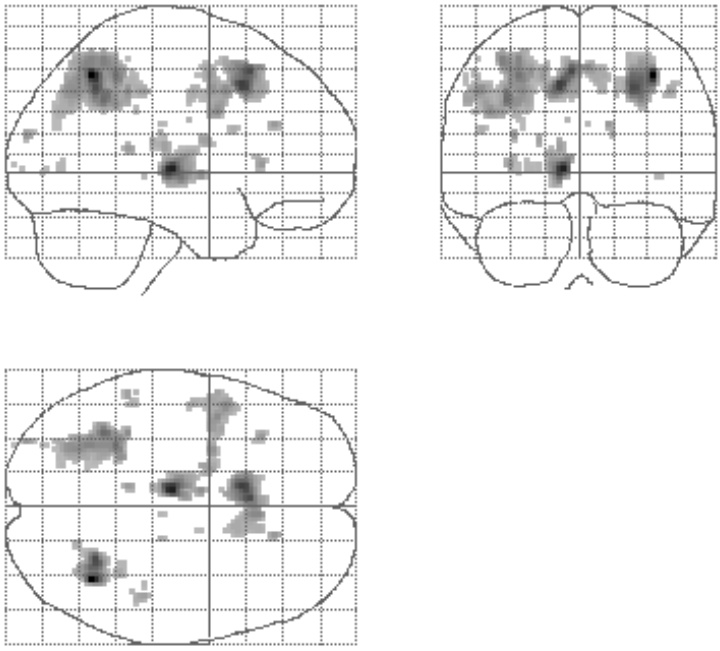

Figure 7: Results from item recognition effect in the working memory data.

FWHM	Intensity	cluster extent via SPM					Proposed cluster mass (Z) method				
		Radius					Radius				
		1	3	5	7	10	1	3	5	7	10
2	0.5	0.0122	0.0123	0.0195	0.0425	0.1211	0.0006	0.0006	0.0006	0.0008	0.0044
	1.0	0.0122	0.0249	0.1991	0.5681	1.0000	0.0006	0.0012	0.0282	0.1601	0.9999
	1.5	0.0122	0.1368	0.8113	0.9961	1.0000	0.0006	0.0212	0.4738	0.9345	1.0000
	2.0	0.0122	0.5482	0.9972	1.0000	1.0000	0.0006	0.1976	0.9773	1.0000	1.0000
4	0.5	0.0156	0.0159	0.0166	0.0243	0.0447	0.0035	0.0036	0.0037	0.0051	0.0129
	1.0	0.0156	0.0160	0.0287	0.1071	0.3821	0.0035	0.0037	0.0088	0.0468	0.2027
	1.5	0.0156	0.0174	0.0902	0.4732	0.9162	0.0035	0.0051	0.0575	0.3099	0.8146
	2.0	0.0156	0.0189	0.3125	0.8912	0.9993	0.0035	0.0110	0.2560	0.8008	0.9976
8	0.5	0.0131	0.0131	0.0138	0.0164	0.0222	0.0227	0.0231	0.0243	0.0264	0.0356
	1.0	0.0131	0.0132	0.0157	0.0219	0.0667	0.0227	0.0243	0.0272	0.0405	0.0941
	1.5	0.0131	0.0134	0.0174	0.0365	0.2309	0.0227	0.0244	0.0360	0.0858	0.2864
	2.0	0.0131	0.0141	0.0191	0.0675	0.5780	0.0227	0.0254	0.0590	0.2206	0.6418
10	0.5	0.0091	0.0091	0.0095	0.0105	0.0154	0.0314	0.0316	0.0331	0.0353	0.0432
	1.0	0.0091	0.0095	0.0102	0.0135	0.0357	0.0314	0.0322	0.0350	0.0459	0.0885
	1.5	0.0091	0.0095	0.0109	0.0191	0.1029	0.0314	0.0333	0.0413	0.0820	0.2237
	2.0	0.0091	0.0095	0.0115	0.0272	0.2836	0.0314	0.0344	0.0567	0.1751	0.4891
12	0.5	0.0084	0.0086	0.0087	0.0095	0.0131	0.0495	0.0501	0.0508	0.0530	0.0631
	1.0	0.0084	0.0086	0.0090	0.0114	0.0254	0.0495	0.0505	0.0540	0.0641	0.1015
	1.5	0.0084	0.0086	0.0091	0.0143	0.0521	0.0496	0.0513	0.0585	0.0930	0.2130
	2.0	0.0084	0.0086	0.0097	0.0159	0.1230	0.0496	0.0519	0.0712	0.1680	0.4199

Table 1: Power of cluster extent inference method via SPM and the propose cluster mass inference method via RFT for Gaussian image, the cluster defined threshold is 2.3263 (p=0.01).

Cluster	Cluster			Uncorrected P values			Corrected P values		
No	Extent	Height	Mass	Extent	Height	Mass	Extent	Height	Mass
1	13	5.09	9.35	0.0069	0.0008	0.0011	0.1606	0.0192	0.0279
2	24	4.52	12.54	0.0009	0.0092	0.0004	0.0238	0.2096	0.0106
3	13	4.45	7.97	0.0069	0.0122	0.0018	0.1606	0.2665	0.0451
4	5	4.10	2.09	0.0633	0.0463	0.0404	0.7999	0.6920	0.6425
5	10	4.08	3.60	0.0140	0.0508	0.0138	0.2992	0.7251	0.2959
6	6	3.87	2.60	0.0446	0.1056	0.0269	0.6782	0.9319	0.4960
7	5	3.65	1.22	0.0633	0.2134	0.0967	0.7999	0.9956	0.9145
8	5	3.48	0.98	0.0633	0.3492	0.1334	0.7999	0.9999	0.9664
9	3	3.43	0.64	0.1447	0.4013	0.2324	0.9764	1.0000	0.9973
10	1	3.34	0.25	1.0000	0.5261	0.6816	1.0000	1.0000	1.0000
11	2	3.21	0.22	0.2433	0.7304	0.7648	0.9979	1.0000	1.0000
12	1	3.18	0.09	1.0000	0.7924	1.0000	1.0000	1.0000	1.0000
13	1	3.16	0.07	1.0000	0.8429	1.0000	1.0000	1.0000	1.0000

Table 2: Real data results for FIAC single subject data analysis, comparing extent, peak height and mass statistics for cluster inference. The cluster mass has good sensitivity, and, in particular, when any of the three inference methods are significant, cluster mass is usually significant,

Cluster No	Random Field Theory Cluster Mass Inference								
	Cluster			Uncorrected p-values			Corrected p-values		
	Extent	Height	Mass	Extent	Height	Mass	Extent	Height	Mass
1	347	5.47	182.19	0.0005	0.0001	0.0002	0.0043	0.0011	0.0018
2	540	4.99	262.29	0.0001	0.0012	0.0001	0.0007	0.0111	0.0004
3	620	4.82	272.05	0.0000	0.0026	0.0001	0.0004	0.0231	0.0004
4	1150	4.34	448.15	0.0000	0.0192	0.0000	0.0000	0.1602	0.0000
5	481	4.02	119.41	0.0001	0.0621	0.0008	0.0012	0.4313	0.0076
6	40	3.43	5.26	0.1012	0.4110	0.1684	0.6014	0.9761	0.7836
Cluster No ‡	Permutation-based Cluster Mass Inference								
	Cluster			Uncorrected P values			Corrected P values		
	Extent	Height	Mass	Extent	Height	Mass	Extent	Height	Mass
1	347	5.47	182.19	0.0018	0.0000	0.0007	0.0098	0.0002	0.0034
2	540	4.99	262.29	0.0008	0.0008	0.0003	0.0039	0.0051	0.0015
3	620	4.82	272.05	0.0006	0.0018	0.0002	0.0037	0.0117	0.0012
4	1150	4.34	448.15	0.0000	0.0132	0.0000	0.0002	0.0803	0.0002
5	481	4.02	119.41	0.0010	0.0461	0.0018	0.0049	0.2305	0.0093
6	40	3.43	5.26	0.0658	0.3327	0.1202	0.2759	0.7515	0.4312

Table 3: Real data results for the small group fMRI data, comparing RFT parametric and permutation nonparametric inferences. Note the similarity between the RFT P-values and permutation P-values, even though the RFT method depends on many assumptions and approximations.