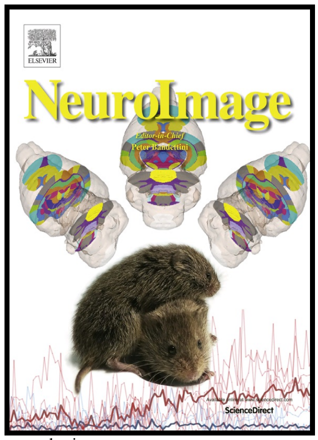
# Author's Accepted Manuscript




www.elsevier.com

Topological false discovery rates for brain mapping based on signal height

Junning Li, Jin Kyu Gahm, Yonggang Shi, Arthur W. Toga

PII: S1053-8119(16)30522-5
DOI: http://dx.doi.org/10.1016/j.neuroimage.2016.09.045
Reference: YNIMG13472

To appear in: *NeuroImage*

Received date: 30 June 2016
Accepted date: 18 September 2016

Cite this article as: Junning Li, Jin Kyu Gahm, Yonggang Shi and Arthur W. Toga, Topological false discovery rates for brain mapping based on signal height *NeuroImage,* http://dx.doi.org/10.1016/j.neuroimage.2016.09.045

# Topological False Discovery Rates for Brain Mapping Based on Signal Height

Junning Li, Jin Kyu Gahm, Yonggang Shi, Arthur W. Toga*

*Laboratory of Neuro Imaging (LONI), Stevens Neuroimaging and Informatics Institute, Keck School of Medicine, University of Southern California, Los Angeles, USA*

**Abstract**

Correcting the effect of multiple testing is important in statistical parametric mapping. If the threshold is too liberal, then spurious claims may flood in; if it is too conservative, then true hints may be overlooked. It is highly desirable to combine random field theory and the false discovery rate (FDR) to achieve more powerful detection under gauged topological errors. However, the current FDR method based on peak height does not fully meet this expectation, and sometimes is more conservative than the traditional family-wise error rate method, for unexplained reasons. In this paper, we introduce a new topological FDR method based on signal height. As analyzed in theory and validated with extensive experiments, it controls error rates much more accurately than the peak FDR method does, and substantially gains detection power. In addition, we discover reasons behind the peak FDR method's under-performance, and formulate equations to predict the two methods' behavior.

*Keywords:* statistical parametric mapping, false discovery rate

## 1. Introduction

Statistical parametric mapping is widely used to locate functional activation or genetic influence in the brain. To tell true signals from random noises, a crucial step is to set a threshold on a map. An experiment with a dead salmon [1] has demonstrated the importance of careful error control. With an arbitrary threshold, correlation can be "detected" between brain activities of the dead fish and human photos shown in front it. This kind of spurious results essentially come from multiple testing. When statistical tests are conducted on numerous points in a volume or on a surface, the chance of making wrong decisions becomes much higher than a single test.

A good threshold should control a certain error rate. Two widely used criteria are (1) the family wise error rate (FWER) [2] which is the probability for any false positive to occur, and (2) the false discovery rate (FDR) [3, 4] which is the expected portion of false positives among reported ones. Rather than prohibiting even a single error, the FDR provides a trade-off between errors and discoveries. Since its debut in 1995 [3], the FDR has been actively adopted.

When a user specifies a target error rate, he/she expects to accurately achieve it. Overly stringent control might miss true hints and reduce detection power. On the other hand, overly liberal control might let spurious claims flood in and later mislead the research community to spend resources on hypotheses which did not hold in the first place.

It is a challenging task to accurately control errors in statistical parametric mapping. First, the test domain is a continuous manifold composed of an infinite and uncountable number of points, even though in practice we usually discretize it as a regular grid or a triangular mesh. Second, geometric and topological properties, such as spatial transformation and connected components, should be considered. Moreover, spatial dependence among signals further complicates the problem.

Random field theory (RFT) is an elegant tool to tackle the aforementioned challenges. With RFT, we can study probabilistic distributions of topological features after segmentation at a threshold, for instance, Euler characteristics or the number of clusters. It takes adjacency between points into account, instead of treating them as unrelated ones. It measures spatial correlation with signals' intrinsic volumes, invariant under spatial transformation. These features meet researchers' need, so RFT has become the statistical core of popular software packages such as Statistical Parametric Mapping (SPM)

☆This work is supported by grants P41EB015922, U54EB020406, R01MH094343, and K01EB013633 from the National Institutes of Health (NIH).

☆☆We acknowledge data obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data, but did not participate in analysis or writing of this paper. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

*Corresponding author

*Email addresses:* Junning.Li@loni.usc.edu (Junning Li), JinKyu.Gahm@loni.usc.edu (Jin Kyu Gahm), yonggang.shi@loni.usc.edu (Yonggang Shi), toga@loni.usc.edu (Arthur W. Toga)