A NOVEL THRESHOLD-FREE NETWORK-BASED STATISTICAL METHOD: DEMONSTRATION AND PARAMETER OPTIMISATION USING IN VIVO SIMULATED PATHOLOGY

Lea Vinokur[1,2], Andrew Zalesky[3,4], David Raffelt[1], Robert Smith[1], and Alan Connelly[1,2]

[1]The Florey Institute of Neuroscience and Mental Health, Heidelberg, Victoria, Australia, [2]Department of Florey Neurosciences, University of Melbourne, Melbourne, Victoria, Australia, [3]Melbourne School of Engineering, University of Melbourne, Melbourne, Victoria, Australia, [4]Melbourne Neuropsychiatry Centre, University of Melbourne, Melbourne, Victoria, Australia

Target Audience: Researchers interested in direct inference of group differences in structural or functional connectivity matrices.

Introduction: The structural and functional connectivity of the brain is increasingly being interrogated using the framework of the *connectome*, where connections between pre-defined regions of grey matter are represented in matrix form. In multi-subject analysis, statistical inference of individual connections (edges) is difficult due to the large number of multiple comparisons. Classic FWER corrections such as Bonferroni are too conservative since connectivity values are not necessarily independent. Network-Based Statistics (NBS)[1] was introduced to address this issue, and improves sensitivity when connections with common nodes have correlated test-statistics. In a similar manner to cluster-based image statistics, NBS joins supra-threshold connections to form connected network components. FWE-corrected p-values can then be assigned to each network component based on network size. As in cluster-based imaging statistics, the main limitation of NBS is the sensitivity to the chosen component-forming threshold. In this work we aim to address this limitation, and propose a modified NBS (called NBS-TFCE) that adapts the concept of Threshold-Free Cluster Enhancement (TFCE)[2] to the network domain. TFCE removes the dependence on an arbitrary threshold by introducing a two other input parameters E and H (see Eq. 1). Recommended values for E and H are justified in the original TFCE method[2], however due to conceptual differences between imaging and network-based statistics, it's not known if these recommended values are also appropriate for NBS-TFCE. Identifying optimal parameters for NBS-TFCE is important to ensure robust comparisons with NBS. In this work we describe the NBS-TFCE method, with particular emphasis on testing the sensitivity of E and H parameters using simulated pathology introduced into the in-vivo structural connectome of a group of healthy control subjects.

Methods: The NBS-TFCE method inputs a test-statistic matrix, and outputs an enhanced matrix where the NBS-TFCE value at connection c is equal to the sum of the component extents, e, as the input test-statistic matrix is thresholded at various heights, h, up to the height of h_c (Eq. 1). As in NBS, the component extent e at a given threshold h is defined as the number of supra-threshold connections that are joined by common nodes (computed using a breadth first search). For statistical inference, we compute FWE corrected p-values for each connection by comparison with a null distribution of maximal test-statistics (via non-parametric permutation testing).

$$NBS\text{-}TFCE(c) = \int_{h=h_0}^{h_c} e(h)^E h^H dh \quad (1)$$

An example enhancement of a t-statistic connectome is illustrated in Fig. 1. Shown in Fig.1-a are edges where simulated pathology was introduced. Fig.1-b illustrates the t-statistic matrix computed with realistic noise (see below). As seen in Fig.1-c, following enhancement with the proposed NBS-TFCE method the signal-to-noise is improved compared to the raw t-statistic matrix.

In [2] the recommended default values of E and H are 0.5 and 2.0 respectively. By setting E to less than 1, the contribution to the enhancement grows less than linear with cluster size. In contrast, with H larger than 1 more weight is given to clusters at higher thresholds. While these defaults seem sensible in practice, the NBS-TFCE method differs in kind since each edge in the matrix is highly connected to many other edges. The low degree of separation between any two edges causes large components sizes at low h thresholds (a limitation also of NBS). To explore the effect of E and H, we assessed the performance of NBS-TFCE enhancement using simulated pathology within three different networks.

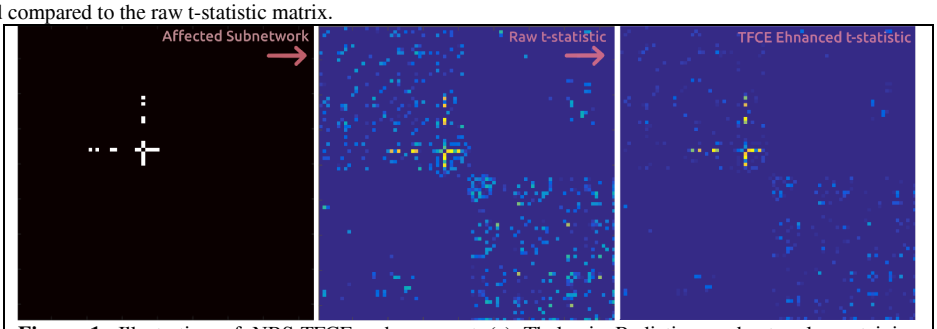

Figure 1: Illustration of NBS-TFCE enhancement (a) Thalamic Radiations subnetwork containing simulated pathology with introduced connectivity reduction (b) Resulting t-statistic image following edge-wise t-test (SNR ~7.5) (c) TFCE-enhanced t-statistic image (SNR ~34)

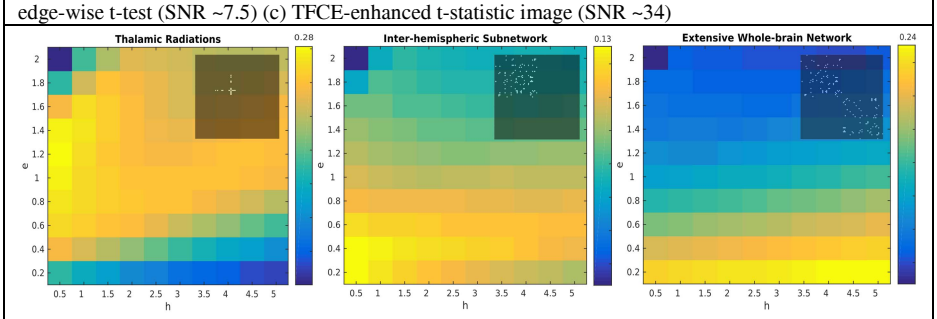

Figure 2: E vs. H plots for the AUC calculated on ROC plots of TPR vs FWE

In vivo simulated pathology: Data were acquired from 22 controls on a 3T Siemens Trio (60 Directions, b=3000 s/mm^2, 2.3mm). DWIs were preprocessed[3]. FODs were computed by constrained spherical deconvolution[4]. The iFOD2 probabilistic tractography algorithm, incorporating the Anatomically-Constrained Tractography (ACT) framework[5], was used to generate 100 million streamlines, with an appropriate subset of 10 million streamlines selected using the SIFT method[6]. Cortical parcellation was performed using the FreeSurfer software package, resulting in 84 nodes according to the Desikan-Killiany atlas. Connectivity matrices (using the number of streamlines connecting each node pair as the connectivity metric) were thresholded to reduce the network density to 20%.

For each of the 2000 permutations, a different subset of 11 subjects was assigned to the pathology group. Simulated pathology was introduced by synthetically reducing the connectivity of a pre-defined interconnected component by a set amount (30% in this case). In this study, three distinct pathologies were tested: (a) a small star-shaped topology of the thalamus and prominent radiations (11 nodes and 10 edges); b) a larger intra-hemispheric network (30 nodes, 48 edges); (c) an extensive network comprising of two clusters of 30 nodes, one in each hemisphere, with 146 edges in total. ROC curves were plotted as by computing the TPR as the number of TP connections within the pathological network, while the FWE-FPR was computed as the proportion of permutations with at least one FP in unaffected areas.

Results: Figure 2 shows the ROC area-under the curve (AUC) values as a function of parameters E and H for each of the three simulated pathologies, along with corresponding network topology illustrated in the corner of each plot. Note that AUC is computed for FWE-FPR < 0.05 and multiplied by 20.

Discussion: In this work, we introduce a potentially important modification to address the main limitation of the NBS approach to network analysis, and investigate the sensitivity of the choice of the parameters E and H on the type of pathology being investigated. The images in Figure 2 suggest that there is no single optimal combination of E and H for all of the data investigated; rather, that the ideal parameter set depends on the topology and/or size of the underlying affected network. Further investigation is necessary to establish the effects of other influences such as non-stationarity, SNR, and network sparsity, and to determine whether a parameter set can be found that performs adequately over a range of pathologies The latter would enable the method to be applied without requiring *a priori* information regarding the nature of network differences, which would be greatly beneficial for neuroscientific investigation.

References: [1] Zalesky A, Neuroimage 2010 ;53(4):1197-207 [2] Smith SM, Neuroimage 2009 ;44(1):83-98 [3] Raffelt D, Neuroimage 2012;59(4):3976-94. [4] Tournier J-D, Neuroimage 2007;35(4):1459-72. [5] Smith RE, Neuroimage 2012;62(3):1924-38. [6] Smith RE, Neuroimage 2013;67:298-312.