MATHEMATICAL METHODS FOR BRAIN CONNECTIVITY

Held in Conjunction with the International Conference on Medical Image Computing and Computer Assisted Intervention

September 22, 2013
Nagoya, Japan

Archana Venkataraman	MIT CSAIL
Bertrand Thirion	Neurospin
Gaël Varoquaux	Neurospin
Maxime Descoteaux	Sherbrooke University
Rachid Deriche	Inria Sophia Antipolis-Méditerranée
Ragini Verma	University of Pennsylvania

Mathematical Methods for Brain Connectivity
MICCAI 2013 Workshop, Nagoya Japan
September 22, 2013

http://groups.csail.mit.edu/vision/mmbc2013/

PREFACE

Connectivity analysis is a burgeoning sub-field of medical imaging that focuses on pairwise relationships between brain regions. For example, anatomical connectivity informs us about neural pathways, which represent the internal wiring of the brain. In contrast, functional connectivity assesses neural synchrony, which relates to communication patterns within the brain. These interactions are crucial to building a comprehensive understanding of the brain that can guide the development of imaging biomarkers.

The 2013 workshop on Mathematical Methods for Brain Connectivity (MMBC) focuses on novel techniques that address the practical applications of brain connectivity. The five original research papers collected in this proceedings clearly demonstrate that the field remains as vibrant and diverse as ever. Topics range from empirical studies of structural connectivity to multimodal analyses that further our insight into the complex nature of the brain. This workshop, held in conjunction with the 16th International Conference on Medical Image Computing and Computer Assisted Intervention, provides a snapshot of the current state-of-the-art methods and highlights some open challenges in the field.

We would like to express our gratitude towards members of the Program Committee for ensuring the high quality of accepted papers. It is now our distinct pleasure to welcome participants to MMBC 2013 and to provide this record of the novel research presented at the workshop.

September 22, 2013   Archana Venkataraman
Nagoya, Japan   Bertrand Thirion
Gaël Varoquaux
Maxime Descoteaux
Rachid Deriche
Ragini Verma

TABLE OF CONTENTS

Workshop Organization

Archana Venkataraman	MIT CSAIL
Bertrand Thirion	Neurospin
Gaël Varoquaux	Neurospin
Maxime Descoteaux	Sherbrooke University
Rachid Deriche	Inria Sophia Antipolis-Méditerranée
Ragini Verma	University of Pennsylvania

Program Committee

Rafeef Abugharbieh	University of British Columbia
Alfred Anwander	Max Planck Institute
Alessandro Daducci	École Polytechnique Fédérale de Lausanne
Thomas Fletcher	University of Utah
Alexandre Gramfort	Telecom ParisTech
Ghassan Hamarneh	Simon Fraser University
Saad Jbabdi	University of Oxford
Georg Langs	Medical University of Vienna
Alexander Leemans	University Medical Center Utrecht
Christophe Lenglet	University of Minnesota
Gabriele Lohmann	Max Planck Institute
Bernard Ng	INRIA, Stanford University
Lauren ODonnell	Harvard Medical School
Jonas Richiardi	Stanford University, University of Geneva
Alard Roebroeck	Maastricht University
Daniel Rueckert	Imperial College London
Jean-Philippe Thiran	École Polytechnique Fédérale de Lausanne
Simon Warfield	Boston Childrens Hospital
Demian Wassermann	Harvard Medical School
Carl-Fredrik Westin	Harvard Medical School

ORAL SESSION I

Properties of Structural Connectivity

Disrupted Brain Connectivity in Alzheimer's Disease: Effects of Network Thresholding

Madelaine Daianu[1], Emily L. Dennis[1], Neda Jahanshad[1], Talia M. Nir[1], Arthur W. Toga[1],
Clifford R. Jack, Jr.[2], Michael W. Weiner[3,4], Paul M. Thompson[1*]
for the Alzheimer's Disease Neuroimaging Initiative[**]
[1]Imaging Genetics Center, Laboratory of Neuro Imaging,
UCLA School of Medicine, Los Angeles, CA, USA
[2]Department of Radiology, Mayo Clinic, Rochester, Minnesota, USA
[3]Department of Radiology, Medicine, and Psychiatry,
University of California San Francisco, CA, USA
[4]Department of Veterans Affairs Medical Center, San Francisco, CA, USA

Abstract. Diffusion imaging is accelerating our understanding of the human brain. As brain connectivity analyses become more popular, it is vital to develop reliable metrics of the brain's connections, and their network properties, to allow statistical study of factors that influence brain 'wiring'. Here we chart differences in brain structural networks between normal aging and Alzheimer's disease (AD) using 3-Tesla whole-brain diffusion-weighted images (DWI) from 66 subjects (22 AD/44 normal elderly). We performed whole-brain tractography based on the orientation distribution functions. Connectivity matrices were compiled, representing the proportion of detected fibers interconnecting 68 cortical regions. We found clear disease effects on anatomical network topology in the structural backbone – the so-called 'k-core' – of the anatomical network, defined by varying the nodal degree threshold, k. However, the thresholding of the structural networks – based on their nodal degree – affected the pattern and interpretation of network differences discovered between patients and controls.

Keywords. brain connectivity, k-core, threshold, DTI, tractography, graph theory

1. Introduction

Diffusion imaging has recently been added to several large-scale neuroimaging studies, including the Alzheimer's Disease Neuroimaging Initiative (ADNI), to monitor white matter deterioration using metrics not available with standard anatomical MRI. Diffusion MRI yields measures sensitive to fiber integrity and microstructure, such as the mean diffusivity and fractional anisotropy of local water diffusion [1]; in addition, tractography can be used to infer neural pathways and connectivity patterns, yielding additional, more complex mathematical metrics describing fiber networks.

Despite the enthusiasm for using diffusion imaging to map brain connectivity and how it changes with disease, there is a lack of serious groundwork validating these methods to see if the connections they map are correct and how acquisition and analysis

protocols affect them. Post-processed connectivity data is also affected by the level of thresholding applied to the brain connectivity matrices; thresholding is commonly applied to retain key information on the most crucial subnetworks, while eliminating false positive fibers or connections inaccurately inferred due to noise and imaging artifacts. There is no consensus about what might be the ideal level of thresholding to retain only the most relevant information in post-processed connectivity data. A common approach filters networks based on the nodal degree, leaving only the most highly connected nodes. As this loses information, some groups advocate defining metrics on the entire set of networks at all thresholds, using concepts such as the *Rips filtration* [2].

Here we studied anatomical fiber networks in 44 controls and 22 identically scanned people with Alzheimer's disease (AD) using novel mathematical network metrics derived from the 'structural backbone' – or k-core – of the human brain. Based on prior studies [3], we were interested in understanding how the different number of nodes, N, in filtered networks from healthy and diseased subjects affects graph theory measures computed from thresholded connectivity matrices. In the end, it would be unwise to infer that AD affects networks in a particular way, if networks filtered differently showed different disease effects. To explore this, we computed the network's structural core using a k-core decomposition [4] to find important sets of nodes that are highly and mutually interconnected. The level of the k-core, k, serves as a threshold to retain nodes in the connectivity matrix with degree k or higher. We systematically varied the values of k (k=1, …, 20) and analyzed the changes in the resulting network measures to understand how they are affected by thresholding the size or degree of the networks (N, k). We calculated global measures sensitive to anatomical network topology: the clustering coefficient (CC), characteristic path length (CPL), efficiency (EFF), and nodal degree (NOD) for all 66 subjects at each of the 20 k-core levels. All network measures showed group differences that depended heavily on the nodal degree and size of the threshold applied to the network. We aimed to find out which network measures are most and least sensitive to variation in the N and k levels, in terms of their ability to resolve differences between the healthy and diseased groups.

2. Methods

2.1. Subjects and Diffusion Imaging of the Brain

We analyzed diffusion-weighted images (DWI) from 66 subjects scanned as part of phase 2 of the Alzheimer's Disease Neuroimaging Initiative (ADNI2), a large multi-site longitudinal study to evaluate biomarkers to assist diagnosis and track disease progression. **Table 1** shows subject demographics and diagnostic information; data collection is ongoing. All 66 subjects underwent whole-brain MRI scanning on 3-Tesla GE Medical Systems scanners, at a variety of sites across North America, with the same protocol, which had been optimized for SNR. Standard anatomical T1-weighted SPGR (spoiled gradient echo) sequences were collected (256x256 matrix; voxel size =

1.2x1.0x1.0 mm^3; TI = 400 ms, TR = 6.984 ms; TE = 2.848 ms; flip angle = 11°) in the same session as the diffusion-weighted images (DWI; 256x256 matrix; voxel size: 2.7x2.7x2.7 mm^3; scan time = 9 min). 46 separate images were acquired for each DTI scan: 5 T2-weighted mages with no diffusion sensitization (b$_0$ images) and 41 diffusion-weighted images (b = 1000 s/mm^2).

Table 1. Demographic information for 44 controls and 22 AD patients scanned with diffusion MRI as part of ADNI. Their ages ranged from 55.7 to 90.4 years.

	Controls	AD	Total
N	44	22	66
Age	72.7 ± 5.9 SD	75.5 ± 10.0 SD	73.6 ± 7.5 SD
Sex	22M/22F	14M/8F	36M/30F

2.2   Image Analysis

Pre-processing and Co-registration

   Non-brain regions were automatically removed from each T1-weighted MRI scan, and from a T2-weighted image from the DWI set using the FSL tool "BET" (http://fsl.fmrib.ox.ac.uk/fsl/). Anatomical scans subsequently underwent intensity inhomogeneity normalization using the MNI "nu_correct" tool (www.bic.mni.mcgill.ca/software/). All T1-weighted images were linearly aligned using FSL (with 6 DOF) to a common space with 1mm isotropic voxels and a 220×220×220 voxel matrix. The DWI were corrected for eddy current distortions using the FSL tools (http://fsl.fmrib.ox.ac.uk/fsl/). For each subject, the 5 images with no diffusion sensitization were averaged, linearly aligned and resampled to a downsampled version of their T1-weighted image (110×110×110, 2×2×2mm). b_0 maps were elastically registered to the T1-weighted scan to compensate for susceptibility artifacts or EPI induced distortions.

Tractography and Cortical Extraction

   The transformation matrix from linearly aligning the mean b_0 image to the T1-weighted volume was applied to each of the 41 gradient directions to properly re-orient the orientation distribution functions (ODFs). We also performed whole-brain tractography as described in [5] on the sets of DWI volumes. We used a method based on the Hough transform to recover fibers, using a constant solid angle orientation density function to model the local diffusion propagator. The angular resolution of the ADNI data is deliberately limited to avoid long scan times that may increase patient attrition, but the ODF model makes best use of the limited available angular resolution.

   Elastic deformations obtained from the EPI distortion correction, mapping the average b_0 image to the T1-weighted image, were then applied to each recovered fiber's 3D

coordinates to more accurately align the anatomy. Each subject's dataset contained ~10,000 useable fibers (3D curves) in total. 34 cortical labels per hemisphere, as listed in the Desikan-Killiany atlas [6], were automatically extracted from all aligned T1-weighted structural MRI scans using FreeSurfer (http://surfer.nmr.mgh.harvard.edu/) [7].

NxN Matrices Representing Structural Connectivity

For each subject, a baseline 68x68 connectivity matrix was created, based on 34 right hemisphere ROIs and 34 left hemisphere ROIs. Each element described the estimated proportion of the total number of fibers, in that subject, that passes through each pair of ROIs. We note that various normalizations could be applied (e.g., using the volume or area of the target ROIs, or to turn these counts into densities), but for simplicity we here just used the fiber counts (normalized to the total number of fibers detected in the brain).

2.3   Brain Network Measures

Topological differences in the brain's networks may be analyzed using graph theory, which represents the brain network as a set of nodes and edges. The network's N nodes are typically defined as ROIs, usually on the cortex, segmented from anatomical MRI. These network nodes are linked by 'edges' whose weights denote some measure of connectivity between the two regions, such as the density or integrity of fiber tracts in DTI studies [8]. An NxN connection matrix may therefore be compiled to describe the network. A square matrix can represent any network of connections, and may also be displayed as a graph, i.e., a discrete set of nodes and edges [8], leading the way for analyses through the branch of mathematics known as graph theory. In our analysis, the matrix entries store the total proportion of fibers connecting each pair of regions (the nodes); these could also be considered as the "weights" of the edges that connect a pair of nodes [8].

From the connection matrices, we applied a threshold by computing the k-core for 20 levels of the nodal degree threshold, k, using a decomposition algorithm that identifies subsets of graphs (k-cores) by recursively removing nodes with degrees lower than k, such that k serves as a degree threshold for nodes [9]. For a graph $G = (N, E)$ with $|N| = n$ nodes and $|E| = e$ edges, a k-core is computed by assigning a subgraph, $H = (B, E|B)$ where set $B \subseteq N$ is a k-core of order k iff $\forall v \in B$: degree$_H \geq k$, and H is the maximal subgraph (most highly connected one) satisfying this property [9]. In other words, to compute the k-core of the connectivity matrix, we kept all nodes with a degree k or higher. These then become new 68x68 matrices, each being a somewhat thresholded version of the original; weights of nodes that did not satisfy the k-cutoff were replaced with zeroes.

We obtained the k-core matrices by varying k from 1 to 20 for both controls and AD subjects. The global graph theory measures (CC, CPL, EFF, and NOD) were derived from each k-core matrix for each subject, to yield four representative network measures at each k-level (i.e., each subject had 20 global metrics for CC, CPL, EFF and NOD). NOD was computed as a nodal measure first, and then averaged overall all 70 cortical regions for

each subject to output a global measure. These widely-used measures are detailed in [8], although their use in brain connectivity and AD research is yet to be extensively explored. CC and CPL measures were normalized based on 100 randomized networks of equal size and similar connectivity distribution. We tested for between-group differences using a linear regression, controlling for age and sex, with AD coded as 1 and controls as 0. We tested for differences between groups of controls and AD subjects for CC, CPL, EFF and NOD at each k-core value for the brain network. We also tested for within-group differences for network measures EFF and NOD, which were found to be "most significant" in the between-group comparison. For this, we compared every k-level across subjects within one diagnostic group with every other k-level in that group (i.e., EFF for controls at $k1$=1,2…19 was compared to EFF for controls at $k2$=($k1$+1)…20) using a 2-tailed paired t-test. We applied an FDR correction on all (20*20-20)/2 comparisons.

3. Results

The variation in the k-core levels (k=1, …, 20) affected the networks and, as expected, resulted in changing graph theory measures (CC, CPL, EFF and NOD) in each diagnostic group.
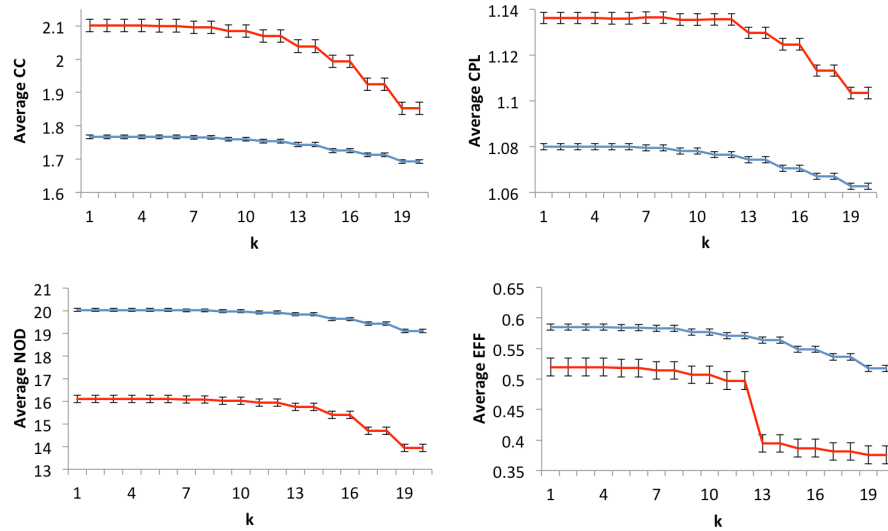
Figure 1. Average and global CC, CPL, EFF and NOD for the whole brain in 44 controls (*blue*) and 22 AD subjects (*red*), based on thresholding the network at k=1, …, 20. Error bars show the standard errors.

We performed *between* group comparisons to find out how effect sizes for group differences depended on the network degree threshold. Relative to controls, the AD group had a higher global CC (FDR critical p-value=6.26E-03) for the entire range of k-core values (k=1-20) and a higher global CPL (p-value=5.72E-3) for k-cores in the range k=1-

18. Obtaining a higher CC in AD, relative to controls, may not be entirely intuitive, but the CC can be disproportionately influenced by nodes with low degree [8]. NOD (FDR critical *p*-value=3.65E-05) and EFF (FDR critical *p*-value=6.21E-05) were lower in AD over the whole range of *k*-core values (*k*=1-20), relative to controls. Averaged network measures (**Figure 1**) and *p*-values (**Figure 2**) are plotted.

Furthermore, we tested for *within* group differences in all subjects for NOD and EFF, as these measures showed greatest effect sizes in the diagnostic group comparisons. The results are shown in a 20x20 matrix, where the EFF was calculated from matrices thresholded at each *k*-level. We compared the EFF network measure to the same network measure calculated from the other *k*-levels – always within the same diagnostic group, to avoid incorporating disease effects (**Figure 3**). EFF changed significantly as *k* varied in both controls and AD (FDR critical *p*-value=1.42E-02 for controls and 1.27E-02 for AD). Within-group measures for NOD were not significantly different across any *k*-levels in either group.
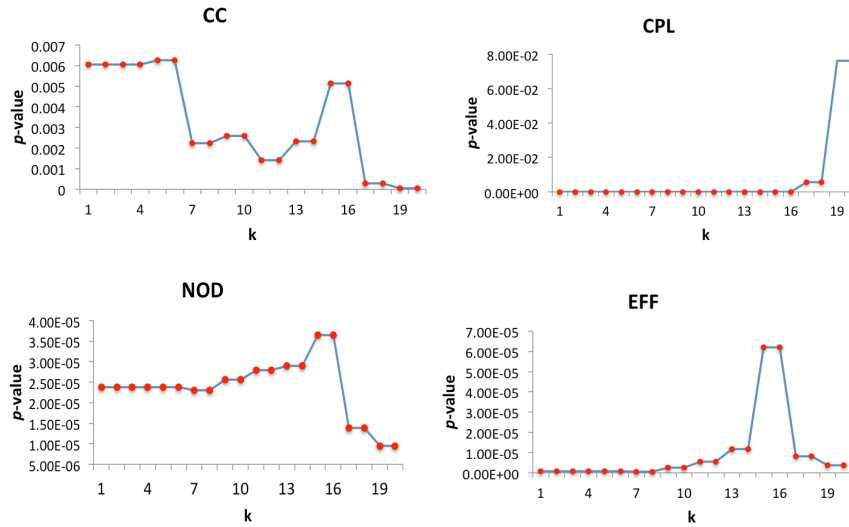
Figure 2. *P*-values from a regression controlling for age and sex, testing for significant differences between AD subjects and controls for whole-brain global CC, CPL, EFF and NOD in AD subjects versus controls. Red points highlight *p*-values that are less than the *p*-value threshold (CC *p*-value=6.26E-03, EFF *p*-value=6.21E-05, NOD *p*-value=3.65E-05 and CPL *p*-value=5.72E-03) that controls the FDR at 5%. This FDR correction allows us to state that the groups truly differ, even though multiple thresholds were tested.

Figure 3. Matrix (20x20) representing the *p*-values from the *within* group comparisons for EFF across all *k*-levels within each group (FDR critical *p*-value=1.42E-02 for controls and 1.27E-02 for AD). A given cell (*x,y*) in this matrix gives the *p*-value for the *t*-test comparing the value of EFF between *k*-cores where the minimum nodal degree is *x* and *y*, respectively. As expected, greatest differences in network measures were found between lowest and highest *k*-levels (red *p*-values).
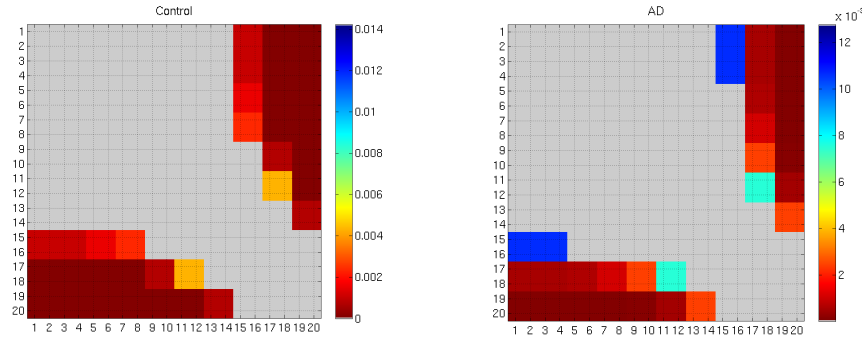
4. Discussion

Graph theory has been widely used to assess functional and anatomical networks in the brain, but not nearly so much attention has been paid to analyzing network variations due to choices made in analysis methods (i.e., network thresholding) and how they impact network topology comparisons. With the growing interest in connectivity analyses, it is important to understand how stable network measures are, and develop reliable guidelines when applying them to study disease. The interpretation of network breakdown in disease may be somewhat different depending on the criteria used to compare or filter networks.

Here we analyzed brain connectivity in cognitively impaired patients with AD and matched normal controls. We varied the nodal degree threshold applied to the connectivity matrices for both groups by using a wide range of *k*-core values (*k*=1, …, 20). Some network measures - CC, CPL, EFF and NOD - *declined* across all subjects as nodal degree threshold levels were increased. Network measures that showed the greatest differences between diagnostic groups over *k* levels ranging from 1 to 20 are in the following order (i.e., with the greatest size effect and smallest *p*-values): NOD, EFF, CPL, and CC. NOD and EFF were found to have greatest size effects among all measures (FDR critical *p*-value=3.65E-05 and 6.21E-05) (**Figures 1** and **2**). This led us to analyze within-group differences for NOD and EFF; we found that increasing levels of *k* significantly affects the apparent efficiency of the overall network in both controls and AD, while NOD was not affected by varying *k* levels (**Figure 3**).

The decline in all network measures with increasing *k* levels is expected in both diagnostic groups. This is because networks thresholded at higher *k* levels required a greater number of nodes to be connected (e.g., at *k*=20, approximately 30% of the nodes are connected). Similarly, AD is known to disrupt the overall network topology of the

brain [2,3] leading to fewer nodes when compared to controls. This is why NOD had the greatest effect size in the between-group comparisons.

An ideal network threshold for this data is in the range of k=15-18. This includes at least 22-26% of the nodes in each brain network, yielding the 'most significant' effects in both *between* and *within* group comparisons. Ideally, this threshold would tend to suppress noise and some imaging artifacts, removing weak connections while emphasizing stronger connections altered in disease. This range may vary with study-specific parameters.

We studied the effect sizes for the group differences here, to clarify how network filtering parameters influence the differentiation of diseased versus normal groups based on graph theory metrics. Although there is no universal method and no definitive answer as to how networks of different sizes and connectivity densities should be accurately compared and analyzed [10], maintaining these measures consistent across study groups is crucial for obtaining comparable results. Normalizing the network measures using randomized networks with the same number of nodes and connections may make graph metrics more stable with respect to differences in N and k [10]. In the end, methods based on network filtrations may supersede those applied to thresholded networks, if they better detect disease effects on brain connectivity.

References

1. Clerx, P., et al., *New MRI markers for Alzheimer's disease: a meta-analysis of diffusion tensor imaging and a comparison with medial temporal lobe measurements*. **J Alzheimer's Dis.** 2012; 29(2):405-29.
2. Lee H., et al., *Persistent Brain Network Homology from the Perspective of Dendrogram*. **IEEE Transactions on Medical Imaging** 31(12):2267-77.
3. Daianu M., et al., *Breakdown of Brain Connectivity between Normal Aging and Alzheimer's Disease: A Structural k-core Network Analysis*. **Brain Connect.** 2013 [Epub ahead of print].
4. Hagmann, P., et al., *Mapping the structural core of the human cerebral cortex*. **PLoS Biology** 2008;6(7):1479-1493
5. Aganj, I., et al., *A Hough transform global probabilistic approach to multiple-subject diffusion MRI tractography*. **Med Image Analysis** 2011;15(4):414-25.
6. Desikan, R. S., et al., *An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest*. **NeuroImage** 2006; 31(3):968-80.
7. Fischl B., et al., *Automatically parcellating the human cerebral cortex.* **Cerebral Cortex** 2004;14:11-22.
8. Sporns, O., *Networks of the Brain*. Cambridge, MA, 2011;5-31.
9. Alvarez-Hamelin, J. I., et al., *Large scale networks fingerprinting and visualization using k-core decomposition*. Weiss Y, Scholkopf B, Platt J. editors. Cambridge (Massachusetts) 2006; MIT Press 44-55.
10. Wijk, B. C. M., et al., *Comparing Brain Networks of Different Size and Connectivity Density Using Graph Theory*. **PLoS One** 2010; 5(10):e13701.

Rich Club Analysis of Structural Brain Connectivity at 7 Tesla versus 3 Tesla

Emily L. Dennis[1], Liang Zhan[1], Neda Jahanshad[1], Bryon A. Mueller[2], Yan Jin[1], Christophe Lenglet[3], Essa Yacoub[3], Guillermo Sapiro[4], Kamil Ugurbil[3], Noam Harel, Arthur W. Toga[1], Kelvin O. Lim[2], Paul M. Thompson[1]

[1]Imaging Genetics Center, Laboratory of Neuro Imaging, UCLA, Los Angeles, CA
[2]Department of Psychiatry, University of Minn., Minneapolis, MN
[3]Center for Magnetic Resonance Research, University of Minn., Minneapolis, MN
[4]Department of Electrical and Computer Engineering, Duke University, Durham, NC

Abstract

The 'rich club' is a relatively new concept in brain connectivity analysis, which identifies a core of densely interconnected high-degree nodes. Establishing normative measures for rich club organization is vital, as is understanding how scanning parameters affect it. We compared the rich club organization in 23 subjects scanned at both 7 and 3 Tesla, with 128-gradient high angular resolution diffusion imaging (HARDI). The rich club coefficient (RCC) did not differ significantly between low and high field scans, but the field strength did affect which nodes were included in the rich club. We also examined 3 subjects with Alzheimer's disease and 3 healthy elderly controls to see how field strength affected the statistical comparison. RCC did not differ with field strength, but again, which nodes differed between groups did. These results illustrate how one key parameter, scanner field strength, impacts rich club organization - a promising concept in brain connectomics research.

1   Introduction

The 'rich club' is an emerging concept in the graph theoretical analysis of brain networks. Initially described in [1], it was first applied to brain networks in [2]. In graph-based analyses of brain connectivity, brain regions are represented as nodes and a set of edges represent the connections between them. These connections may be defined based on fiber tracts extracted from diffusion MRI, or based on time-course correlations between different brain regions in functional MRI data. For brain networks, the rich club is defined as a central core of high-degree nodes that are more highly interconnected than would be expected simply from their high degree. Some authors argue that the rich club is crucial for understanding global network efficiency; an attack on rich club nodes disproportionately affects global efficiency [2]. Rich club connectivity organization changes with age as brain connectivity strengthens [3]. Brain networks are complex, and the rich club concept offers a principled approach for dimension reduction: it identifies a key set of crucial nodes that contribute maximally to network

efficiency. The *rich club coefficient*, $\phi(k)$, represents the density of connections between the rich club nodes according to the following equation:

$$\text{Eq. 1} \quad \phi(k) = \frac{E_{>k}}{N_{>k}(N_{>k}-1)}$$

Here k is the degree of the nodes, $E_{>k}$ is the number of links between nodes with degree k or greater, and $N_{>k}$ is the total number of possible connections if those nodes of degree k were fully connected. Van den Heuvel and Sporns also introduced $\phi_{norm}(k)$, which is $\phi(k)$ divided by the rich club coefficient calculated in a series of random networks ($\phi_{random}(k)$) of the same size with a similar distribution of edges [2]. A $\phi_{norm}(k)$ value > 1 indicates rich-club organization in the network.

The rich club, as it pertains to brain networks, is a new topic of interest, and little is known about its stability across MRI scanning parameters. If network parameters depend on the scanner field strength or other imaging parameters, researchers need to be aware of this, to ease pooling of multi-site data and resolve apparent discrepancies among studies. We hypothesized that the brain's fiber network could be imaged reproducibly at 7 and 3 Tesla to yield the same rich club content and coefficient, despite differences in scanning protocols (high-field versus standard magnetic field strength). We set out to examine how MR field strength affects rich club properties in 23 young adults scanned at both 7 and 3 T. In pilot work, we also examined 3 subjects with Alzheimer's disease (AD) and 3 age- and sex-matched healthy controls (HC) to see how statistical comparisons might be affected by field strength. These data come from prior work by our group examining how MR field strength affects connectivity, including more basic tractography measures [4].

2 Methods

2.1 Subject demographic and image acquisition

Whole brain anatomical and DW-MRIs at both 7 and 3 T were collected from 23 young adults (11 female, mean age=23.8, SD=2.6) and 6 elderly subjects (3 AD: 2 female, mean age=76.1, SD=3.2; 3 HC: 2 female, mean age=78.3, SD=2.4); the protocols are detailed in [5,6]. Standard head coils were used on both systems: the 12-channel receive-only array on the 3T, and a Nova 24 channel transmit/receive coil on the 7T. The reconstruction method for the 3T scanner was adaptive recombine (AC), while the default multi-channel reconstruction method for the DWI data on the 7T scanner was sum-of-squares (SOS). 3T DW-images were acquired with the following acquisition parameters: GRAPPA mode; acceleration factor PE=2; TR/TE=7800/82 ms; FOV=192x192 mm, isotropic voxel size=2 mm. 143 images were collected per subject: 15 b_0 and 128 diffusion-weighted (b=1000 s/mm^2). 7T DW-images were acquired with the following acquisition parameters: GRAPPA mode; acceleration factor, PE=2; TR/TE=5700/57 ms; FOV=256x256 mm, isotropic voxel size=2 mm. 143 images were collected per subject: 15 b_0 and 128 diffusion-weighted (b=1000 s/mm^2). T1-weighted anatomical images were acquired at 3 Tesla with the following acquisition parameters: GRAPPA mode; acceleration factor PE=2; T1/TR/TE=1100/2530/3.65 ms; echo spacing = 8.5 ms; flip angle = 7°; slice thickness = 1.0 mm, with an acquisition matrix of 256x256. All subjects gave informed consent

after study protocols were explained.

2.2   Image preprocessing and registration

All DWI data were visually inspected by an experienced rater for evidence of the known Siemens vibration dropout artifact [7]. No dropout artifact was found in the DW data. All raw DWI images were corrected for distortions due to eddy currents and motion using the *eddy_correct* function from the FSL toolbox (http://fsl.fmrib.ox.ac.uk/fsl) [8,9]. Geometric distortions due to magnetic susceptibility were then corrected using a field map collected just before the DTI, using the FSL *prelude and fugue* functions. Non-brain regions were removed from a T2-weighted image (b_0) in the corrected DWI dataset using the *bet* function in FSL. A trained neuroanatomical expert manually edited the T2-weighted scans to refine the brain extraction and to ensure the same brain coverage among different protocols. This step was important to avoid bias, as different connectivity patterns might be recovered if brain coverage varies. All analyses below are based on this preprocessed dataset.

2.3   Brain connectivity computation

Tractography and cortical networks were computed in the native space of the data. The Diffusion Toolkit (http://trackvis.org/dtk/, [10]) uses these parameters to generate 3D fiber tracts, using the Orientation Distribution Function model, computed using the 2nd order Runge-Kutta method [11]. We used all voxels (with FA≥ 0.2) as seed voxels to generate the fibers. Paths were stopped when they reached a region with FA < 0.2; they were also stopped if the fiber direction encountered a sharp turn (critical angle threshold $\geq 30°$). After tractography, a spline filter was applied to each generated fiber, with units expressed in terms of the minimum voxel size of the dataset (2 mm). Each subject's dataset contained 25,000-40,000 useable fibers (3D curves). Duplicate fibers and very short fibers (< 10mm) were removed. Although we did not do this here, some researchers normalize fiber count by ROI volume at this point [12].

Cortical and subcortical ROIs were defined using the Harvard Oxford Cortical and Subcortical probabilistic atlases [13]. Midline cortical masks were bisected into left and right components, to define separate hemispheric ROIs for each cortical region. Since this is a probabilistic atlas, the masks were set to a liberal threshold of 10% to include tissue along the gray-white matter interface, where fiber orientation mapping and tractography are most reliable [14]. To register these ROIs to each subject's DTI space, we used FSL's *flirt* function to determine the optimal affine transformation between the MNI152 T1 average brain (in which the Harvard Oxford probabilistic atlases are based) and each subject's unique FA image. We used a 12 degree-of-freedom registration with a mutual information cost function. We applied the resulting transformation to register the 110 ROIs to each subject's DTI space using nearest neighbor interpolation. To ensure that ROI masks did not overlap with each other after registration, each voxel was uniquely assigned to the mask for which it had the highest probability of membership. For a list of ROIs, see [4]. We did not include the brainstem and cerebellum ROIs, giving us a total of 110 ROIs.

For each pair of ROIs, the number of detected fibers connecting them was determined from the tractography results. A fiber was considered to connect two ROIs if it intersected both ROIs. This process was repeated for all pairs, resulting in an 110x110

matrix. This matrix is symmetric and has a zero diagonal (no self-connections).

2.4 Rich club analyses

On these 110x110 matrices, we used the Brain Connectivity Toolbox ([15]; https://sites.google.com/a/brain-connectivity-toolbox.net/bct/ Home) to compute the rich club coefficient ($\phi(k)$). We calculated the $\phi(k)$ over values of k (the degree of the nodes) ranging from 0-110 to capture all possible values. To generate the normalized rich club coefficient ($\phi_{norm}(k)$), we simulated 50 random networks. These matrices were first binarized so the actual weights of the edges were not factored in, simply the number of connections. Analyses comparing $\phi(k)$ and $\phi_{norm}(k)$ were performed across all subjects. In order to compare rich club organization, we constructed average graphs for the 3T and 7T datasets. For the young cohort, these were averaged across all 23 subjects, and the group-averaged matrices were thresholded to include only connections found in at least 75% of subjects; this step is helpful to suppress false positive fibers arising from tractography errors. For the elderly cohort, these were averaged separately for the AD and HC subjects. Given the small sample size, we did not threshold the group-averaged matrices. To determine the k cut-off for rich club membership, we used the same criteria as [2]: we included nodes having a degree at least one standard deviation above the average degree. For the young cohort, the average degree for the 3T group-averaged network was 57.4, while for the 7T network it was 54.7. This was not a significant difference, however. This resulted in a k cutoff of 69 at 7T and 71 at 3T, and we used these thresholds for **Figure 1**. We will call these 'analogous k-levels' from here on in the paper. For the elderly cohort, the average degree for the 3T AD group-averaged network was 93.5, and in HC it was 91.2; this group difference was significant ($p=0.025$). For the 7T group-averaged networks, the average degree for AD was 73.0, and in HC it was 79.9, a difference that was also significant ($p=2.1\times10^{-16}$). These averages are higher than for the young cohort because we could not filter these networks in the same way, given the small sample size.

3 Results

3.1 Rich club coefficient ($\phi(k)$ and $\phi_{norm}(k)$)

We ran a paired-sample t-test at each k-level (the nodal degree threshold) to look for protocol effects on our subjects' connectomes, by studying both $\phi(k)$ and $\phi_{norm}(k)$ across subjects. We did not detect any significant differences in rich club coefficient, either $\phi(k)$ or $\phi_{norm}(k)$, between protocols in the young cohort. We also did not detect any differences in $\phi(k)$ or $\phi_{norm}(k)$ between AD and HC at 3T or 7T, although admittedly we were underpowered to pick up group differences.

3.2 Rich club organization – Young cohort results

When $k=69$ for 7T and $k=71$ for 3T, as justified above, there were differences in the rich club organization of the group-averaged 3T and 7T matrices. These mostly resulted from the fact that the 3T rich club at $k=71$ included 21 nodes, while the 7T rich club at $k=69$ included only 19 nodes. This was due to a slight difference in average degree, as mentioned above. These results are shown in **Figure 1**.
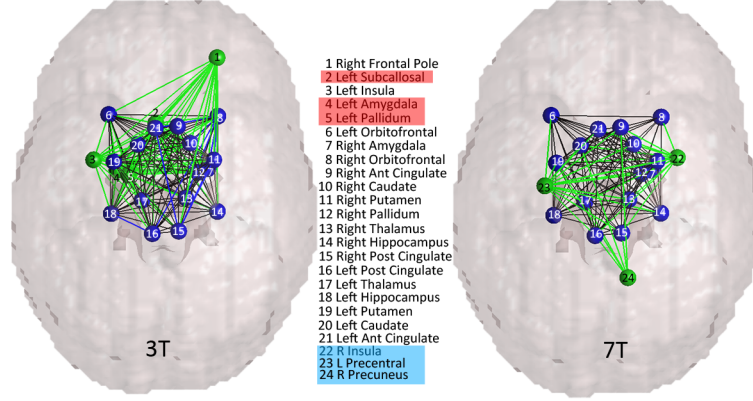
Figure 1. **Differences in rich club organization between 3T and 7T scanning protocols.** Green nodes are unique to one rich club, blue are common to both. Black edges are common to both rich clubs, blue edges are unique to one rich club but among common nodes, green edges are unique to one rich club due to the unique nodes. 3T *k*-level=71, 7T *k*-level=69. The red highlighted nodes are unique to the 3T rich club when N_{node} was kept constant at 19, instead of *k*-level, blue highlighted nodes are unique to the 7T rich club in the N_{node} analysis**.**

Knowing that a difference in degree impacted rich club membership, we also compared rich club organization when the *number* of nodes was held constant, rather than looking analogous *k*-levels, as before. When the node-number (symbolized here by N_{node}) was held constant at 19 nodes, there were still differences in rich club membership. When held constant at N_{node}=19, the common nodes between the two connectomes were nodes 6-21, as listed in **Figure 1**. Those unique to the 7T rich club were nodes 22-24, highlighted in blue, while those unique to the 3T rich club were nodes 2, 4, and 5, highlighted in red. Interestingly, those unique to the 3T rich club are all subcortical, while those unique to the 7T rich club, when N_{node}=19, are all cortical. Following up on this N_{node} analysis, we compared $\phi(k)$ when node number was held constant. **Table 1** shows these results, across a range of N_{node} tested. To do this, we looked at the number of nodes present at *k*-levels 69-75 for 7T, and then found where this N_{node} boundary was in the 3T network. The N_{node} comparisons are not exact between 3T and 7T, as that would require arbitrarily cutting off nodes that had the same *k*-level. As we were comparing $\phi(k)$, which is still calculated based on level, arbitrarily cutting off nodes would make comparing the $\phi(k)$ values invalid. This kind of analysis is intended to give a clearer idea of the relationship between the $\phi(k)$ of the 3T and 7T connectomes when the number of nodes is held constant. As seen in **Table 1**, $\phi(k)$ is significantly greater in the 3T connectome across most of this range. We chose to look at a range beginning with 7T *k*-level=69, as that was the 'high degree' threshold for 7T mentioned above, and ending with the last *k*-threshold at which rich club organization was detectable (75 for 7T, 77 for 3T). We ran a paired-sample *t*-test on the distribution of $\phi(k)$ at a given *k*-level between protocols, and used the false discovery rate method (FDR) to correct for multiple comparisons (q<0.05, [16]).

Table 1. **Comparison of high-field and standard field protocols maintaining** N_{node}**.** Across a range, $\phi(k)$ was averaged across all subjects. N_{node} indicates the number of nodes at a given k-level. P denotes the p value for a two-tailed t-test comparing the means of the $\phi(k)$ for the 3T and 7T protocols at a given k-level. All results are corrected for multiple comparisons, with the FDR method ($q<0.05$).

7T k level	69	70	71	72	73	74	75
N_{node}	19	14	14	11	10	9	8
Avg. $\phi(k)$	0.9217	0.9332	0.9434	0.9529	0.9603	0.9674	0.9732
3T k level	72	74	74	75	75	76	77
N_{node}	19	15	15	11	11	9	9
Avg. $\phi(k)$	0.947	0.9577	0.9577	0.9652	0.9652	0.9712	0.9758
p	0.0014	0.0039	0.066	0.073	0.47	0.55	0.64

3.3   Rich club organization – AD/HC comparison

We compared which nodes were included at the statistically determined k cut-offs for rich club membership (3T AD=104, 3T HC=105; 7T AD=90, 7T HC=96). We expected to find differences between the AD and HC subjects, but were most interested in how the differences between groups varied with field strength. Given our small sample size (3 AD, 3 HC), these results are preliminary. These are summarized in **Figure 2**. There were many differences that were only detectable at 3T. Both groups had larger rich clubs at 3T, like the young cohort, again due to lower degree at 7T.
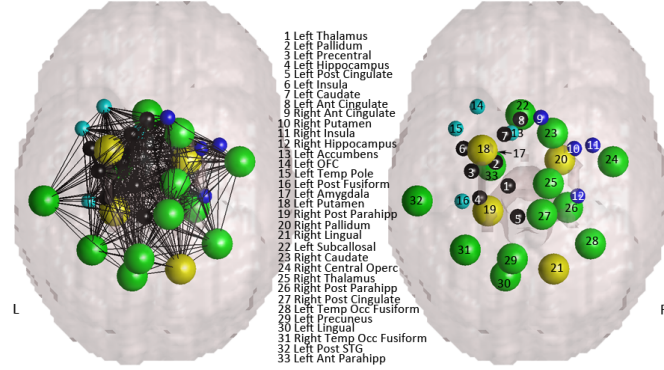

1 Left Thalamus
2 Left Pallidum
3 Left Precentral
4 Left Hippocampus
5 Left Post Cingulate
6 Left Insula
7 Left Caudate
8 Left Ant Cingulate
9 Right Ant Cingulate
10 Right Putamen
11 Right Insula
12 Right Hippocampus
13 Left Accumbens
14 Left OFC
15 Left Temp Pole
16 Left Post Fusiform
17 Left Amygdala
18 Left Putamen
19 Right Post Parahipp
20 Right Pallidum
21 Right Lingual
22 Left Subcallosal
23 Right Caudate
24 Right Central Operc
25 Right Thalamus
26 Right Post Parahipp
27 Right Post Cingulate
28 Left Temp Occ Fusiform
29 Left Precuneus
30 Left Lingual
31 Right Temp Occ Fusiform
32 Left Post STG
33 Left Ant Parahipp

Figure 2. Differences in the comparison of AD versus HC subjects between 3T and 7T scanning protocols. Left image shows rich club nodes with connections, as averaged across all subjects (thresholded to include connections found in at least 66% of subjects), right image shows just the nodes, for clarity. Small nodes showed no effect of field strength, large nodes showed an effect of field strength; black nodes both groups had; blue nodes present in AD only; cyan nodes present in HC only; green nodes had group diff. only in 3T; yellow nodes had group diff. only in 7T. The right side of the image denotes the left side of the brain.

4   Discussion

In this paper we used a unique dataset, comparing rich clubs recovered from the same groups of subjects scanned at both 3T and 7T. As the rich club is a relatively new metric in brain connectivity analyses, it is important to know how much measures

vary across key scan parameters. In a prior paper studying this same dataset [4], the 7T protocol had higher SNR than the 3T protocol, as expected from MR theory, but this did not affect the estimates of FA (fractional anisotropy). We ran binary graph theory analyses (not weighted), so some of the more subtle differences between protocols may be washed out. We did not have the space here to consider weighted rich club analyses, but will do so in the future.

One might expect the rich club to contain more nodes at higher field, based on the presumably more accurate and complete recovery of connections. The more extensive rich club seen in the 3T connectome may be due to the higher noise level in this dataset compared to the 7T connectome, which increases the likelihood of false positive fibers. In our group-averaged networks, we did threshold the connectivity matrices to include only those connections found in at least 75% of the subject pool, which should decrease the number of false positives. If specific areas of the brain, such as subcortical structures, are particularly vulnerable to false positives with the 3T protocol, consistent tractography errors may even be made across subjects that are able to survive this thresholding. A weighted analysis might be more sensitive to this, if the false positive connections are weak. We intend to pursue these analyses.

Comparing networks of analogous k-levels is intuitive, as the rich club is defined by statistically high degree nodes. Even so, an analysis of networks thresholded to contain the same *number* of nodes (retaining those with highest degree) is complementary, as it can reveal the true direction of associations masked by differences in degree. When we compared $\phi(k)$ between connectomes, keeping the N_{node} constant (or closer than it would be in a k-level analysis), we found significant differences in $\phi(k)$, across a range of nodal degree thresholds, k, and N_{nodes}. Across these significant ranges, $\phi(k)$ was higher in the 3T connectome than the 7T connectome. Higher $\phi(k)$ indicates a greater density of connections between rich club nodes. This could also be due to differences in signal to noise as discussed above. The 7T protocol revealed a more 'trimmed down' rich club network. We started out with a binarized analysis, but a weighted analysis may reveal a very different rich club, as significant increases have been found in the density of subcortical connections in the 7T protocol [4]. Other parcellations will obviously yield different results.

In our elderly dataset, we compared the rich clubs of AD and HC subjects at both 3T and 7T to see how field strength might affect group comparisons. Again, the 3T rich clubs included more nodes than the 7T rich clubs, which led to 12 nodes showing group differences in the 3T matrices that no longer showed group differences at 7T. This could be due to increased noise in the 3T data, or decreased resolution for subcortical structures, or increased susceptibility artifacts in the 7T data. These are only preliminary data, but have important implications for future work using rich club measures to investigate the effects of neurological disorders, especially if rich club is ever to be used as a biomarker of disease.

5   Conclusion

Here we compared the rich club coefficient ($\phi(k)$) and anatomical network organization at 3T and 7T in a group of 23 subjects scanned with both protocols. $\phi(k)$ did

not depend on field strength when compared at analogous k-levels, but it did differ when the N_{node} was kept constant. The 3T connectivity matrices had a higher average degree than did the 7T matrices, leading to our comparing $\phi(k)$ on networks with the same N_{node}, as well as on a k-level. When comparing rich clubs at an analogous k-level, we found a number of differences in which nodes were included in the rich club between protocols. The 3T connectome had a far more extensive rich club than the 7T connectome. When we examined our elderly AD and HC subjects, we similarly found differences only in which nodes were included in the rich club. These preliminary results need further analysis, however, for rich club measures to be reliable bi-omarkers. As the rich club coefficient is a new metric intended to represent a crucial contributor to network efficiency, we believe these results are important for understanding some of the fundamental factors that may affect rich club calculations.

References

1. Colizza, V. et al. (2006). Detecting rich-club ordering in complex networks. **Nat Phys**, 2, 110-115.

2. Van den Heuvel, M. & Sporns, O. (2011). Rich-club organization of the human connectome. **J Neuroscience**, 31(44), 15775-15786.

3. Dennis, E. L. et al. (2013). Development of the "Rich Club" in Brain Networks from 438 Adolescents and Adults Aged 12 To 30. **IEEE ISBI**, 620-623.

4. Zhan, L. et al. (2013). Magnetic resonance field strength effects on diffusion measures and brain connectivity networks. **Brain Connectivity**, 3(1), 72-86.

5. Stanisz, G. J. et al. (2005). T1, T2 relaxation and magnetization transfer in tissue at 3T. **Magn Reson Med**, 54, 507–512.

6. Yacoub, E. et al. (2003). Spin-echo fMRI in humans using high spatial resolutions and high magnetic fields. **Magn Reson Med,** 49, 655–664.

7. Gallichan, D. et al. (2010). Addressing a systematic vibration artifact in diffusion-weighted MRI. **Hum Brain Mapp**, 31, 193–202.

8. Smith, S. M. et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. **NeuroImage**, 23, 208–219.

9. Woolrich, M. W. et al. (2009). Bayesian analysis of neuroimaging data in FSL. **NeuroImage**, 45, S173–S186.

10. Wang, R. et al. (2007). Diffusion Toolkit: A Software Package for Diffusion Imaging Data Processing and Tractography. **Proc Intl Soc Magn Reson Med**, 15.

11. Basser, P. J. et al. (2000). In vivo fiber tractography using DT-MRI data. **Magn Reson Med**, 44, 625–632.

12. Duarte-Carvajalino, J. M. et al. (2012) Hierarchical topological network analysis of anatomical human brain connectivity and differences related to sex and kinship. **NeuroImage**, 59 (4), 3784-3804.

13. Desikan, R. S. et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. **NeuroImage**, 31, 968–980.

14. Morgan, V. L. et al. (2009). Integrating functional and diffusion magnetic resonance imaging for analysis of structure-function relationship in the human language network. **PLoS One**, 4, e6660.

15. Rubinov, M. & Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. **NeuroImage**, 52, 1059-1069.

16. Benjamini, Y & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. **J Roy Stat Soc B**, 57(1), 289-300.

ORAL SESSION II

Multimodal & Population Analysis of Connectivity

Coupled Intrinsic Connectivity: A Principled Method for Exploratory Analysis of Paired Data

Dustin Scheinost, Xilin Shen, Emily Finn, Rajita Sinha,
R. Todd Constable, and Xenophon Papademetris

Departments of Biomedical Engineering, Diagnostic Radiology,
Interdepartmental Neurosceince Program, Psychiatry, and Nuerosurgery
Yale University, New Haven, CT

Abstract. We present a novel voxel-based connectivity approach for paired functional magnetic resonance imaging (fMRI) data collected under two different conditions labeled the Coupled Intrinsic Connectivity Distribution (coupled-ICD). Our proposed method jointly models both conditions to incorporate additional spatial information into the connectivity metric. When presented with paired data, conventional voxel-based methods analyze each condition separately. However, nonlinearities introduced during processing can cause this approach to underestimate differences between conditions. We show that commonly used methods can underestimate functional changes and evaluate our coupled-ICD solution using a study comparing cocaine-dependent subjects and healthy controls. Our approach detected differences between paired conditions in similar brain regions as the conventional approaches while revealing additional changes. Follow-up seed-based analysis confirmed, via cross validation, connectivity differences between conditions in regions detected by coupled-ICD that were undetected using conventional methods. This approach of jointly analyzing paired connectivity data provides a new and important tool with many clinically relevant applications.

Keywords: Functional Connectivity, Resting-state, fMRI, Cocaine

1 Introduction

Functional connectivity holds promise as a clinical tool to detect abnormal brain organization in clinical populations. The most common approaches rely on regions of interests (ROIs) or "seeds" to characterize connectivity; however, seed-based approaches can only examine connectivity in reference to the seed region. Choosing which seeds to examine is often a difficult question as the wrong choice in seed regions could occlude important patterns of connectivity.

Voxel-based metrics can be used as a data-driven way to define seed regions for further analysis [1–3]. However, voxel-based approaches and seed approaches can often produce seemingly conflicting results. For example, voxel-based results may suggest an increase in connectivity for a region while follow-up seed analysis with the region may show decreases in connectivity to the region. This discrepancy arises because each approach is fundamentally different. Voxel-based

metrics essentially work as compression mechanisms, reducing all information about the connections to a voxel into a few summary parameters used for group comparisons. This compression of information is nessessary as connectivity matrix genernated from voxel-based approaches (typically 20,000x20,000 matrix at fMRI resolutions) are difficult to interpret and are problematic for statistical inferences. In contrast, seed-based approaches directly compare correlations between regions at the group level rather then these summary parameters.

For the special case of paired data such as pre- and post-treatment, the standard approach with voxel-based metrics is to compure the summary parameters separately for each condition and then perform statistical analysis to compare the two. We observe that this approach is suboptimal as the compression into a summary parameters is performed twice (once for each condition). Thus, with this approach, comparisons are made on how these summary parameters change rather then how correlations between regions change. Further, non-linearities introduced during processing – such as only examining the positive correlations [1, 4] – guarantee that the difference in the summary parameter is not the same as the summary parameter of a difference. Hence, information about how each correlation changes due to the treatment is also lost with current approaches.

In this work, we propose a method where within-subject differences across conditions are first computed and then a single summary measure can be calculated for these differences. We label our approach the coupled Intrinsic Connectivity Distribution (coupled-ICD) as it extends the recently developed Intrinsic Connectivity Distribution (ICD) method [9]. Unlike other voxel-based metrics, coupled-ICD mimics seed-based approaches by directly comparing correlations between each condition and, then, summarizing these changes into summary parameters for group comparisons. Thus, coupled-ICD should produce regions more suitable for seed-based connectivity. To assess our coupled-ICD measure, we used a data set of cocaine-dependent subjects and healthy controls scanned while presented with relaxing and drug-related imagery cues. We show that our coupled-ICD has higher sensitivity than conventional approaches for detecting differences between conditions. Finally, using cross-validation on separate, independent sub-sample of our data, we show regions detected by coupled-ICD are predictive of seed-based difference in connectivity.

2   Theory

Voxel-based measures of functional connectivity [4, 1, 9] aim to reduce large amounts of information to a voxel into a much smaller set of summary parameters. Typically, this compression is formulated based on graph theory [8] where the brain is treated as a graph or network and each voxel represents a node in this graph. These nodes (or voxels) are connected to each other by edges based on the similarity of their timecourses.

Measures of node centrality such as the measure degree are the primary metrics used for compression. For any voxel x, these measures can be calculated from the distribution of connection strength, $f(x, r)$, where x is the current

voxel, and r is a correlation or any other measure of timecourse similarity. First, $f(x, r)$ is estimated by computing the histogram of the correlations r for the timecourse at voxel x to the timecourse at every other voxel in the brain. Degree can be estimated as the integral of this distribution from any threshold τ to 1, or $\int_{\tau}^{1} f(x, r) dr$. Weighted degree measures such as weighted Global Brain Connectivity (wGBC) [1] can be estimated as the mean of this distribution. In contrast, ICD models the corresponding survival function to $f(x, r)$. Each point on the survival function is simply degree evaluated at that particular threshold τ and, thus, ICD parameterizes how degree for a voxel changes as the threshold used to determine if two voxels are connected is increased. Previously, it was shown that a stretch exponential decay with unknown variance parameter α and shape parameter β was sufficient to model this survival function. Modeling the survival function with a stretch exponential is equivalent to modeling the underlying distribution as a Weibull distribution: $f(x, r, \alpha, \beta) = \frac{\beta}{\alpha}(\frac{r}{\alpha})^{\beta-1} \exp(-(\frac{r}{\alpha})^{\beta})$.

The presented approach, coupled-ICD, extends conventional voxel-based connectivity in a critical way as the graph summarized by coupled-ICD is a graph defined by differences in correlations and not simply the correlations. As such, coupled-ICD takes advantage of the paired nature of the data by explicitly comparing the same edge under two conditions. Typically, graphs at the voxel level become difficult to analyze due to memory constraints and multiple comparison issues with groups of 10 or more subjects. However, for the special case of paired scans, only two graphs need to be analyzed simultaneously. Paired scans reduce the complexity of this problem.

Directly comparing each correlation (like seed-based appoarches) allows additional spatial information about changes of corresponding edges to be incorporated into a summary parameter of connectivity for a single subject. Specifically, computing the differences between the weights of corresponding edges of a graph and then summarizing the differences takes into account the topological (spatial) structure of the graphs. With this approach, information about how each edge has changed due to condition or time can be incorporated into the summary parameter for the subject. This information is lost with current approaches. Due to non-linearities in the calculation of these summary parameters, the difference between the summary parameters (degree, wGBC, or ICD) of two graphs is not the same as the summary parameter of the difference between the graphs. For example, given the ambiguity of negative correlations, many current approaches examine only the positive correlations [4, 9, 1].

Given a set of paired data, coupled-ICD can be computed by repeatedly calculating conventional seed connectivity maps treating each voxel as a seed, and summarizing the difference between the seed maps for each condition (Fig. 1). First, for any voxel x, the correlation between the timecourse at voxel x to the timecourse at every other voxel in the gray matter is calculated for each condition in the paired data. These correlation maps are then subtracted from one another. Coupled-ICD then summarizes this map of differences in the same way that ICD (or degree) summarizes a map of connections to a voxel. First, for each voxel, a distribution of these differences is estimated with a histogram. Second, this dis-
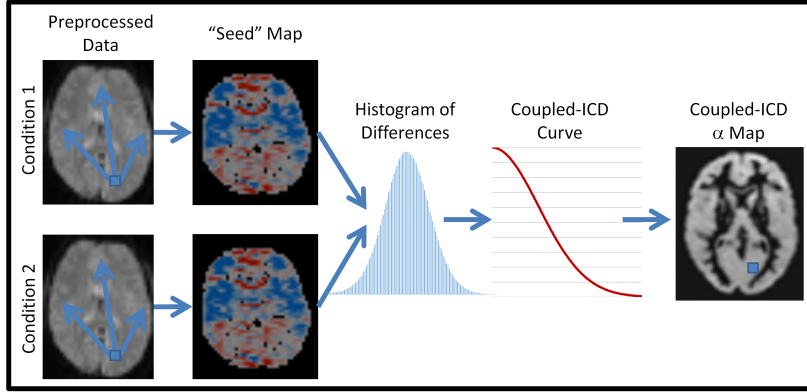
Fig. 1. Flow chart describing coupled-ICD. For paired data, coupled-ICD jointly analyzes both conditions and then creates a summary of the difference in connectivity between conditions for each voxel. First, a seed connectivity map is created for a voxel (shown as the blue square through the flow chart) in each condition. The resulting survival function of the distribution of the difference (labeled coupled-ICD curve) is calculated and modeled with a stretched exponential. This process is repeated for each voxel in the gray matter. The final output is an image where each voxel represents a summary of the difference between two seed maps using that voxel as the seed region.

tribution is modeled as a Wiebull distribution which corresponds to modeling the survival function of the histogram as a stretch exponential. Group comparisons can be performed by comparing the parameters with standard methods.

Coupled-ICD can be used to model increases in connectivity, decreases in connectivity, or the magnitude of the changes in connectivity. For simplicity, we focus only on modeling magnitude of the changes in connectivity noting that the modeling and the interpretation of the parameters is similar for the other cases. When modeling the magnitude of changes, a larger α parameter indicates a larger variance in the distribution and that a larger number of connections exhibit a strong change in correlations between the two conditions.

3 Functional Connectivity Estimation

Subjects: The data set consisted of 28 cocaine-dependent (CD) subjects and 38 healthy control (HC) subjects aimed at examining influence of cue state and diagnostic group on brain activity. Subjects performed four fMRI scans while listening to imagery scripts of either neutral relaxing cues or drug related cues (two scans of each). Complete details can be found elsewhere [7, 10].
Preprocessing: Images were slice-time corrected using sinc interpolation and motion corrected using SPM5. All further analysis was performed using in-house software. Several covariates of no interest were regressed from the data including linear and quadratic drift, six rigid-body motion parameters, mean cerebral-spinal fluid (CSF) signal, mean white matter signal and mean global signal.

Finally, the data were temporally smoothed with a zero mean unit variance Gaussian filter (cutoff frequency=0.12Hz). A gray matter mask was applied to the data so that only voxels in the gray matter were used in the calculation.

ICD and Degree: The timecourse for voxel x was correlated with the timecourse for every other voxel in the gray matter. As the removal of the global mean makes the signs of the correlation ambiguous [4], only the positive correlation was used in analysis. For each voxel, a distribution of connection strength was estimated for the positive correlation coefficients using a 100 bin histogram. ICD was used to model this distribution. First, the histogram was converted to the corresponding survival function and this survival function was modeled with a stretched exponential. This results in two summary parameters for each voxel reflecting that voxels connectivity to the rest of the brain; the α parameter was used in the group comparisons. As this survival function describes how the network theory measure degree changes with connection threshold, degree at any threshold can be estimated as a single point on the survival function. Degree was estimated with a connection threshold of $r = 0.25$. ICD or degree maps for each paired condition were then subtracted from each other resulting in a single map per subject describing the difference in connectivity between conditions.

Coupled-ICD: Similar to the ICD and degree estimations, the timecourse for voxel x was correlated with the timecourse for every other voxel in the gray matter. As coupled-ICD operates on paired data, this process was performed on both conditions resulting in two seed connectivity maps with voxel x as the seed. These maps are then subtracted. A distribution of the differences in connection strength was estimated for the absolute value of the differences using a 200 bin histogram. A larger number of bins was used to keep the bin width the same as the ICD analysis while accommodating the wider range of possible values (difference in correlations has a range of $[-2, 2]$ while correlation has a range of $[-1, 1]$). We chose to model the absolute value of the differences to highlight regions of the brain that show large differences between two conditions. A similar procedure can be used to analyze the increases or decreases between conditions. As described above, this histogram was converted to the corresponding survival function and this survival curve was modeled with a stretched exponential.

Seed Connectivity: A follow-up seed-based analysis (similar to $[4, 2, 3]$) was performed on a sample region showing large differences as detected using coupled-ICD. The voxel-based analysis and follow-up seed-based analysis were run on independant data by splitting the data into two groups by subjects. Fourteen CD subjects and 19 HC subjects were randomly chosen for voxel-based analysis. The remaining subjects were used for seed-based analysis. Splitting the data into two groups allows the seed connectivity results to act as a cross-validation of the coupled-ICD results. A seed was placed in putamen based on voxels showing significant differences ($p < 0.05$, corrected) between HC and CD subjects. The timecourse of the reference region in a given subject was then computed as the average timecourse across all voxels in the reference region. This timecourse was correlated with the timecourse for every other voxel in the gray matter to cre-
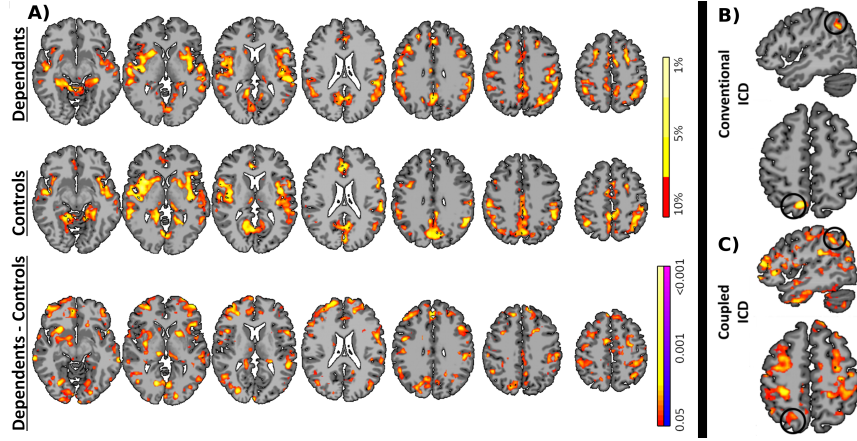
Fig. 2. Evaluation of coupled-ICD. A) Coupled-ICD detects widespread differences in connectivity while subjects are experiencing either relaxing or drug-related imagery for both (*top*) the cocaine-dependent (CD) subjects. and (*middle*) healthy controls (HC). (*Bottom*) Group level comparisons (CD vs HC) revealed that the groups significantly ($p < 0.05$, corrected) differ in response to the imagery conditions. **B)** Conventional ICD analysis detected significant group differences ($p < 0.05$ corrected) in the parietal and occipital lobes. **C)** Coupled-ICD detected significant ($p < 0.05$, corrected) difference in these areas as well as several other areas. This result highlights the additional information that can be captured by jointly analyzing paired conditions.

ate a map of r-values, reflecting Seed-to-whole-brain connectivity. These r-values were transformed to z-values using Fisher's transform.

Group Analysis: To facilitate group statistics, all single subject results were spatially smoothed with a $6mm$ Gaussian filter and non-linearly warped to common space using Bioimage Suite [5]. Between group differences were calculated using two-sample t-test with significance assessed at $p < 0.05$. AFNI's Alphasim was used for multiple comparison correction.

4 Results

Coupled-ICD detected widespread significant differences due to condition between the two groups. The coupled-ICD maps for each group and the between-group comparisons are shown in Fig. 2. Both groups showed large differences between the drug-related and relaxing imagery in the posterior cingulate cortex, bilateral angular gyrus, bilateral insular cortex, bilateral putamen, medial prefrontal/anterior cingulate cortex, and visual processing areas. The CD group showed significantly greater differences between conditions in many of these regions and additionally in the prefrontal lobe.

The between-group differences detected by coupled-ICD were compared with between-group differences detected by the conventional ICD and degree approaches. Conventional ICD detected two clusters that satisfied our criteria for
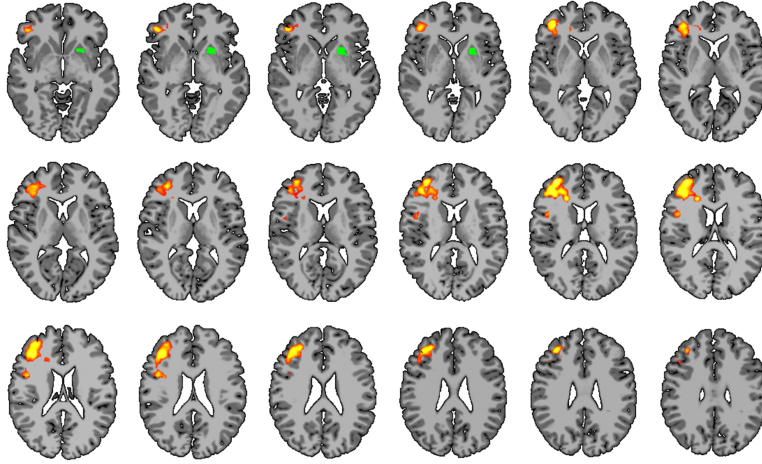
Fig. 3. Seed validation. A follow-up, seed-based connectivity analysis was performed on independant data using a region in the left insula detected by coupled-ICD but not by conventional ICD and degree analysis. Several areas of significant differences were detected ($p < 0.05$, corrected). As the seed analysis was performed on independant data, the seed-based results provide evidence via cross-validation that coupled-ICD is detecting an effect not detected by conventional analysis

significant differences-in the parietal and occipital lobes-while the degree approach did not detect any significant differences. Group differences for both coupled-ICD and ICD are shown in Fig. 2. Coupled-ICD identifies both clusters detected by conventional ICD along with additional widespread changes.

To explore the differences detected by coupled-ICD, a follow-up seed-based analysis was performed using a seed defined in the left putamen where significant between-group differences were found using coupled-ICD. The left putamen was chosen as a seed due to the substantial literature (see [7] for example) implicating this region in addiction. Significant ($p < 0.05$, corrected) interactions between group and condition were observed in right frontal lobe (Fig. 3).

5 Discussion

We present a principled method for exploratory analysis of paired conditions to detect regions that differ significantly in their connectivity patterns between conditions. We show that our coupled-ICD approach is superior in detecting group differences in connecitivity due to paired conditions. We show that coupled-ICD is a viable solution as a data-driven way to pick seeds for further analysis. Standard seed-based analysis, performed on data independent from the coupled-ICD results, showed that the regions detected by coupled-ICD exhibit significant differences in seed connectivity. While similar to conventional voxel-based metrics of connectivity, the present approach extends connectivity in a critical

way, coupled-ICD summaries differences in correlations, rather compute difference in summaries of correlation. This extension allows coupled-ICD to mimic seed-based approaches and gain additional information for group comparisons.

Numerous clinical applications could benefit from measuring changes in the functional organization of the brain at the voxel level for paired data, yet the translational technology for detecting changes in connectivity remains elusive. Coupled-ICD represents a principled method for exploratory analysis of paired conditions to detect regions that differ significantly in their connectivity patterns between conditions. Thus, coupled-ICD could potentially fill this important void not currently covered by conventional approaches.

References

1. Cole, M.W., Yarkoni, T., Repovs, G., Anticevic, A., Braver, T.S. (2012) Global connectivity of prefrontal cortex predicts cognitive control and intelligence. The Journal of Neuroscience, 32:8988-99.
2. Constable, R.T., Vohr, B.R., Scheinost, D., Benjamin, J.R., Fulbright, R.K., Lacadie, C., Schneider, K.C., Katz, K.H., Zhang, H., Papademetris, X., Ment, L.R. (2012) A left cerebellar pathway mediates language in prematurely-born young adults. Neuroimage, 64C:371-378.
3. Hampson, M., Tokoglu, F., Shen, X., Scheinost, D., Papademetris, X., Constable, R.T. (2012) Intrinsic brain connectivity related to age in young and middle aged adults. PloS one, 7:e44067.
4. Buckner, R.L., Sepulcre, J., Talukdar, T., Krienen, F.M., Liu, H., Hedden, T., Andrews-Hanna, J.R., Sperling, R.A., Johnson, K.A. (2009) Cortical Hubs Revealed by Intrinsic Functional Connectivity: Mapping, Assessment of Stability, and Relation to Alzheimer's Disease. The Journal of Neuroscience, 29:1860-1873.
5. Joshi, A., Scheinost, D., Okuda, H., Belhachemi, D., Murphy, I., Staib, L.H., Papademetris, X. (2011) Unified framework for development, deployment and robust testing of neuroimaging algorithms. Neuroinformatics, 9:69-84.
6. Li, C.S., Sinha, R. (2008) Inhibitory control and emotional stress regulation: neuroimaging evidence for frontal-limbic dysfunction in psycho-stimulant addiction. Neurosci Biobehav Rev, 32:581-97.
7. Potenza, M.N., Hong, K.I., Lacadie, C.M., Fulbright, R.K., Tuit, K.L., Sinha, R. (2012) Neural correlates of stress-induced and cue-induced drug craving: influences of sex and cocaine dependence. Am J Psychiatry, 169:406-14.
8. Rubinov, M., Sporns, O. (2010) Complex network measures of brain connectivity: Uses and interpretations. NeuroImage, 52:1059-1069.
9. Scheinost, D., Benjamin, J., Lacadie, C.M., Vohr, B., Schneider, K.C., Ment, L.R., Papademetris, X., Constable, R.T. (2012) The Intrinsic Connectivity Distribution: A Novel Contrast Measure Reflecting Voxel Level Functional Connectivity. Neuroimage.
10. Seo, D., Jia, Z., Lacadie, C.M., Tsou, K.A., Bergquist, K., Sinha, R. (2011) Sex differences in neural responses to stress and alcohol context cues. Hum Brain Mapp, 32:1998-2013.

Power Estimates for Voxel-Based Genetic Association Studies using Diffusion Imaging

Neda Jahanshad[1], Peter Kochunov[2], David C. Glahn[3],
John Blangero[4], Thomas E. Nichols[5,6], Katie L. McMahon[7],
Greig I. de Zubicaray[8], Nicholas G. Martin[9], Margaret J. Wright[9],
Clifford R. Jack, Jr.[10], Matt A. Bernstein[10], Michael W. Weiner[11,12],
Arthur W. Toga[1], and Paul M. Thompson[1] for ADNI

[1] Imaging Genetics Center, Lab of Neuro Imaging, UCLA, Los Angeles, CA USA
[2] Maryland Psychiatric Research Center, U of Maryland, Baltimore, MD USA
[3] Olin Neuropsychiatry Research Center, Yale School of Medicine, New Haven, CT
[4] Department of Genetics, Texas Biomedical Research Institute, San Antonio, TX
[5] Department of Statistics & Warwick Manufacturing Group, U of Warwick, Coventry, UK
[6] Oxford Centre for Functional MRI of the Brain (FMRIB), Oxford University, UK
[7] University of Queensland, Centre for Advanced Imaging, Brisbane, Australia
[8] University of Queensland, School of Psychology, Brisbane, Australia
[9] Queensland Institute of Medical Research, Brisbane, Australia
[10] Department of Radiology, Mayo Clinic, Rochester, Minnesota, USA
[11] Department of Radiology, Medicine, and Psychiatry, UC San Francisco, CA, USA
[12] Department of Veterans Affairs Medical Center, San Francisco, CA, USA

Abstract: The quest to discover genetic variants that affect the human brain will be accelerated by screening brain images from large populations. Even so, the wealth of information in medical images is often reduced to a single numeric summary, such as a regional volume or an average signal, which is then analyzed in a genome wide association study (GWAS). The high cost and penalty for multiple comparisons often constrains us from searching over the entire image space. Here, we developed a method to compute and boost power to detect genetic associations in brain images. We computed voxel-wise heritability estimates for fractional anisotropy in over 1100 DTI scans, and used the results to threshold FA images from new studies. We describe voxel selection criteria to optimally boost power, as a function of the sample size and allele frequency cut-off. We illustrate our methods by analyzing publicly-available data from the ADNI2 project.

Keywords: Neuroimaging genetics, heritability, GWAS, DTI, multiple comparisons correction.

1 Introduction

Imaging genetics is an emerging field in which variations in the human genome are related to brain differences. Genome-wide association studies (GWAS), for example, test for statistical associations between brain measures and up to a million single

nucleotide polymorphisms (SNPs[1]) in the genome. To simplify the screening effort, studies often focus on a single measure extracted from brain scans, such as the overall volume of the hippocampus [1]. Despite successful findings with simple summary measures, the image space contains many more features, e.g., at individual voxels, which can be used for genetic testing. Testing all these features, however, imposes a high cost of multiple comparisons. GWAS tests at each voxel are computationally feasible but underpowered, due to the large number of tests performed [2]. Efforts to boost voxelwise power include set-based multivariate tests (multilocus genetic tests), ICA [3], and sparse regression methods in the image and/or genome [4-6]. If we are able to analytically reduce the image space *a priori* to only those regions where we expect sufficient power for association, we could maximize our search space while minimizing the penalty for multiple comparisons.

Power in GWAS is a topic of great interest, as any algorithms to boost power would make genetic studies more efficient, faster, or more cost-effective. Power calculations can inform data collection, i.e., how many subjects to recruit. Power can be computed for a GWAS analysis of a single measure, or (conversely) to find how many subjects are required to detect a pre-defined effect size on a brain measure, for one single nucleotide polymorphism (SNP) with a given minor allele frequency (MAF). One commonly used tool for calculating power is the Genetic Power Calculator [7]. Power and required sample sizes can be computed for association tests with binary outcomes (e.g., diagnosis in a case-control study) [8], or quantitative traits such as brain measures. Another approach to optimize power is to select subjects non-randomly, to optimize power to detect effects of rare variants [9]. However, in large neuroimaging studies, often data have already been collected, sometimes without genetic discovery in mind, and the challenge is to compute image features to maximize power to detect genetic associations. A key goal is to prioritize the list of >100,000 possible voxelwise phenotypes[2] to only those that meet a desired expected power threshold, of say 80%, for a given sample size, N.

Power calculators depend on distributional assumptions, and may require certain parameters to be known, such as the means and variances of the trait in each of the genotype groups. Often, we may not know these means and standard deviations [10]. One approach [11] suggests the use of heritability to avoid the need to specify the mean and variance required for the non-centrality parameter. Other common assumptions are that allele frequencies and genotype counts are known. However, these are often not known beforehand, and ignoring uncertainties can lead to overly optimistic power estimates [12]. Even with a fixed population sample with allele frequencies in Hardy-Weinberg equilibrium (HWE), the genotype distribution of the actual sample can vary. It becomes essential to compute the expected power by taking into account the expected genotype distribution, as has been shown for single alleles [12, 13]. While these problems have been addressed in the context of single variants, with known or unknown allele frequencies, p, power analysis for GWAS studies is still challenging. In GWAS, over one million alleles may be tested, all with different

---

[1] At each location on the genome, a person has a specific "letter" or nucleotide; SNPs are common variants in the genetic code, carried by at least 1% of the population.

[2] A *phenotype* is a biological measure that is subjected to genetic analysis, such as the size of a brain region, or a diffusion imaging measure in a specific region.

minor allele frequencies (MAFs). For gene discovery, including GWAS studies of brain imaging phenotypes, all these alleles will have unknown and varying genotype counts, although a limit can be set as to how much the SNP deviates from HWE, where HWE represents the expected genotype frequencies for each allele given the MAF in the population of interest. If, for example, the allele MAF=0.3 for a given population, then with a sample size of N=1000, $N*p^2$=90 people in the cohort would be expected to have 2 copies of the minor allele (n_2), $2*N*p*(1-p)$=420 people should have 1 copy, (n_1), and $N*(1-p)^2 = 490$ people should have 0 (n_0). Deviations from this count are assessed using a χ^2 test with one degree of freedom (dof); samples with genotype counts that differ significantly, are not in HWE.

Studies using DTI have reported moderate correlations between DTI-based measures of fractional anisotropy (FA) and common variants in specific candidate genes (e.g., *CLU*, *HFE*, *NTRK1*, and *BDNF*) known to associate with neurological diseases and disorders. This motivates the use of FA as phenotypes for GWAS, as they may implicate genes that affect disease risk. Recently, we [14] computed the first large-scale heritability study on voxelwise FA maps, using meta-analysis methods on two cohorts comprising a total of over 1100 subjects. Here, we use these existing voxelwise measures of heritability to estimate the power of GWAS at each voxel of a new dataset. We limit our multiple comparisons correction problem by filtering SNPs based on a user-defined threshold for the MAF. We compute the expected power at each voxel, with appropriate limits on the number of subjects in the genotype groups, based on the HWE. This allows us to focus on voxels of the FA map where we would expect a certain level of power for detecting a genetic effect, given a specific N.

2 Methods

2.1 Heritability and Power Estimates

We can use linear regression to test the effect of a genetic variant on a quantitative trait, such as a brain imaging measure. In this way, the expected value for the trait is modeled as $\mu_i = \mu_0 + \beta*i + \beta_x*x + \varepsilon$, where i =0,1,2 represents the number of copies of the minor allele, and β_X and x represent the nuisance variables such as age and sex. Here the F-ratio can be used to test the equality of the means (H_0: β=0, H_A: $\beta\neq0$). The F-ratio follows a non-central F-distribution with 1 and N-2 dof, with a non-centrality parameter λ: $\lambda = \dfrac{\left(\sum_i n_i(i-\bar{x})(\mu_i-\mu)\right)^2}{\sum_i n_i(i-\bar{x})^2 \sigma^2}$, where $\bar{x} = \dfrac{\sum_i n_i*i}{\sum_i n_i}, \mu = \dfrac{\sum_i n_i\mu_i}{\sum_i n_i}$ and σ is the within-group SD. The power is $\pi = \Pr\left(F_{1,N-2,\lambda} > F_c\right)$, where F_c is the (1-α)*100 percentile of the central F-distribution, and α is the significance threshold (e.g., 0.05). Note that if we add d covariates to the model, our distribution would have 1 and N-2-d dof. However, the means and variances within the groups may be unknown - or impractical to calculate - for each of the possible GWAS SNPs.

Alternatively, the heritability of a trait (h^2) is defined as the proportion of the observed variance in a trait (here a brain measure) that is attributable to genetic

variation; $h^2 = V_g/V_T$ – or, the model sum-of-squares, divided by the total sum-of-squares (SSM/SST= 1-(SSE/SST)). As suggested in [11], the *F*-statistic can be directly calculated as a function of h^2 and *N*. This is because according to an ANOVA table for a regression (**Table 1**), the *F*-statistic=(SSM*DFE)/(SSE*DFM); as SSE= $(1-h^2)$*SST and SSM= h^2*SST, then $F = \frac{h^2 * DFE}{(1-h^2) * DFM}$, and in the case of one explanatory variable, $d=1$, $F = \frac{h^2 * (N-2)}{1-h^2}$. We can therefore use the trait's h^2 to calculate power, π.

Table 1 ANOVA table for a multiple regression model, with *d* explanatory variables.

ANOVA	dof	Sum of squares	Mean squares	F
Model	DFM=1	SSM= $\sum_j (\bar{y}_j - \bar{y})^2$	MSM=SSM/DFM	MSM/MSE
Error	DFE=*N*-*d*-1	SSE= $\sum_j (y_j - \bar{y}_j)^2$	MSE=SSE/DFE	
Total	DFT=*N*-1	SST=SSM+SSE	MST=SST/DFT	

Recently, we [14] meta-analyzed voxelwise measures of heritability from FA maps in two family-based cohorts totaling over 1100 subjects. One cohort was a sample of Caucasians imaged at 4T with 94 directional gradients and voxels of size 1.7x1.7x2mm, and the other was a Mexican-American sample imaged at 3T with 64 gradients and 1.8x1.8x3mm voxels. Given differences in cohort demographics and scanning protocols, and the overall large sample size (1100+), we treat these meta-analyzed values as the best available heritability estimates for the voxelwise phenotypes in any studies of FA mapped to the same space. We carry these heritability measures forward, to estimate power in a new GWAS study of unrelated individuals. We express the power as a function of *N*, calculating the probability under the *F*-distribution (1 and *N-2* dof) for the heritability at every voxel.

2.2 HWE, MAF, and multiple comparisons correction

Tests of HWE assess whether the sample genotype frequency is aligned with the frequency expected based on reference data compiled from human populations (such as *HapMap*). Including SNPs with significant deviations from HWE can be problematic [16], but the threshold for determining significant deviations is often different between studies [17]. Thresholding this deviation at more stringent values eliminates group formations that could artificially appear to enhance power. A direct filter on the SNPs involves thresholding the MAF, to remove rare variants and SNPs with allele frequencies too low to be adequately sampled. Given the relatively lower *N* in imaging studies compared to other genetic studies, filtering out low frequency SNPs can reduce the number of tests, alleviating the multiple comparisons penalties, without compromising power. We filter SNPs based on MAF, and adjust possible outcomes with HWE deviations. Once SNPs have been filtered, the significance threshold can be based on the number of SNPs (N_{SNPs}). The false discovery rate (FDR) *q*-value which ensures that the false positive rate across all *voxels,* is controlled at q_{FDR}= 0.05/N_{SNPs}. We count the number of voxels (N_{VOX}) that pass the threshold. Now power may be calculated as $\pi = \Pr(F_h > F_c)$, with F_c determined by α=0.05/ N_{SNPs}/N_{VOX}.

2.3 Accounting for uncertainties in genotype frequency

As mentioned in [12, 13], for a given allele frequency p, the expected power is the weighting of the estimate π (here a function of h^2 and N) by the probabilities of the multinomial counts, $\mathbf{n}=(n_0, n_1, n_2)$. By contrast with prior papers, we limit $\mathbf{n}$ to only those combinations that fall within HWE_c.

Additionally in this work, as opposed to those mentioned above, we are conducting a GWAS of all SNPs with $p > MAF_c$ (a lower bound on the MAF of the SNPs) rather than a single SNP. Therefore, rather than setting a beta-prior on p, we sum over all possible frequencies, each one with respect to only the possible $\mathbf{n}$ in HWE, and all possible $p > MAF_c$. Note that, by definition, the MAF ≤ 0.5.

$$E(v)_{GWAS}\left[\pi(h^2,N)\,|\,\mathbf{n}_{HWE}\right] = \sum_{p=MAF_c,\Delta p}^{0.5-\Delta} \frac{\Pr(p,p+\Delta)}{\Pr(MAF_c \leq p \leq 0.5)} \sum_{\sum_i n_i = N; n_i \geq 0\,\forall i} \pi(h^2,N)\frac{N!}{n_0!n_1!n_2!}2^{n_1}p^{2n_0+n_1}(1-p)^{n_1+2n_2}$$

This estimates power for GWAS at the individual voxel, v – and we only include those voxels that pass the FDR correction as mentioned above. The estimated power is a weighted function of: (1) the probability that the p falls within a given interval, Δ, and (2) the probability of the genotype group distribution within HWE. π can now be taken out of both summations, and a single weight, $wt(MAF_c,HWE_c)$, is given to each voxel. $E(v)$ can be written as $\pi(h^2(v),N)*wt$. As an example, we use data from the second phase of the Alzheimer's Disease Neuroimaging Initiative study (ADNI2; publicly available at http://adni.loni.ucla.edu), and show the proportion of SNPs that fall across a range of allele frequencies (**Figure 1**). Using the formulation above, we compute the N_{VOX} that surpass a power threshold while varying the N_{SNPs} examined by varying MAF_c and HWE_c, as well as how this estimate is modified by sample size, N.

2.4 Voxelwise GWAS of the ADNI2 dataset

To date, ADNI2 has genotyped 78 Caucasian subjects scanned with DTI. Scans were processed and aligned to the heritability map detailed in [14]. Limiting our search space to only voxels that pass the 0.8 power threshold with $MAF_c=0.1$ and $HWE_c=1\text{x}10^{-5}$, we ran a GWAS at all remaining voxels using an additive model; this is modeled with a linear regression for each SNP.

3 Results

3.1 Voxels with power > 0.8 as functions of N, MAF_c, HWE_c

The whole-brain white matter skeleton of the DTI FA-template consists of 97158 voxels total. In **Figure 2**, we show how the expected N_{VOX} changes as we threshold the MAF from 0.01 to 0.2 in step sizes of 0.01, and as we become more stringent with the HWE_c. We do not eliminate SNPs due to divergence, but rather filter out the possible combinations of subjects that could result in the given p. Therefore, as we are more

strict with HWE$_c$ ($1x10^{-6} < \dots < 0.01 < 0.05$), more possibilities are deemed divergent and possibly biased, so power is lower, and fewer voxels are selected.

3.2 Voxelwise GWAS in the ADNI2 dataset

In our real data analysis, we first find the number of voxels that will have an expected power $E_{GWAS}(v) > 0.8$ with MAF$_c$ = 0.1 (N_{SNPs} = 522,077 out of 661,478), in the ADNI2 dataset of 78. As this is a dataset including several patients with Alzheimer's disease, we would not necessarily expect the HWE to be upheld, as disease could impact the frequency of some alleles – especially AD risk alleles. We were therefore less strict with HWE$_c$, selecting $p<10^{-5}$ as our cutoff. We obtained a mask consisting of 1664 voxels (**Figure 3A**), and ran a GWAS at all voxels.

When correcting for all the voxels and the SNPs tested, no SNP had a significance level less than $0.05/(1664*522,077) = 5.75x10^{-11}$ in the ADNI2 sample. However, as these voxels were chosen to have the highest power to detect a genetic association, we show that when plotting the minimum p-values at each voxel, against what would be expected from a null distribution (here, a Beta distribution with parameters $(1, N_{SNPs})$ [18]) in a Q-Q plot (**Figure 3B**), we show our p-values tended to be lower than what would be expected from a null distribution.

4 Discussion

We have presented a method to estimate power for imaging genetics studies that apply genome-wide scanning to multiple phenotypes; in fact, the same method could also be adapted to prioritize targets for genetic analysis. Most power estimates for GWAS studies are computed to plan future data collection, but here we assume imaging data has already been collected (or there is an upper limit on N). We then show how to optimize the study to focus the GWAS on the most powerful regions of the high-dimensional image space. Our approach is also flexible - users can select a MAF or HWE cutoff to estimate the best thresholds to optimize the search space.

Limitations of our study include the following: (1) the heritability presented here may differ from that of the population a user is studying. However, we used the best available estimates of heritability of FA to date, given the large sample size, and the use of meta-analysis to mitigate differences in demographics and imaging protocols; (2) Studies have also shown that choosing the inappropriate model for the SNP (e.g., additive when in fact it has a dominant effect) can bias power calculations [13]. Even so, most GWAS studies assume an additive model for each SNP, so we follow this convention. (3) Here, we do not account for the correlation between voxels, or the linkage disequilibrium (LD) structure of the SNPs. If LD is considered, tests are not all independent and we can reduce the number of effective tests, reducing the multiple comparisons penalty. Our method can incorporate this correlation, so our power estimates (assuming SNP independence) are somewhat conservative. Regardless, our methods can estimate the power of voxelwise association tests, providing a starting ground for GWAS of spatially extended phenotypes.

Additionally, it is possible that limiting rarer variants by removing the SNPs with low MAF, may result in removal of SNPs that may have the greatest effect sizes, however with a limited number of scans, it is unlikely that the effect of these rare variants would be accurately obtained.

Here, we show that meta-analyzed heritability estimates are useful for genetic studies on the entire image space. We focus on voxelwise analyses, but so long as heritability estimates have been made on multiple imaging phenotypes, this method can be extended beyond maps of FA values, to mean volume or anisotropy values within regions of interest, or to all of the individual network elements that make up the human connectome.

References

1.   Stein, JL, et al., *Identification of common variants associated with human hippocampal and intracranial volumes*. Nature Genetics, 2012. **44**(5): p. 552-61.
2.   Hibar, DP, et al., *Voxelwise gene-wide association study (vGeneWAS): multivariate gene-based association testing in 731 elderly subjects*. NIMG 2011. **56**(4) 1875-91.
3.   Calhoun, VD et al., *A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data*. NIMG 2009. **45** S163-72.
4.   Ge, T., et al., *Increasing power for voxel-wise genome-wide association studies: the random field theory, least square kernel machines and fast permutation procedures*. NIMG 2012. **63**(2) 858-73.
5.   Vounou, M et al., *Sparse reduced-rank regression detects genetic associations with voxel-wise longitudinal phenotypes in Alzheimer's disease*. NIMG 2012 **60**(1) 700-16.
6.   Silver, M., et al., *Identification of gene pathways implicated in Alzheimer's disease using longitudinal imaging phenotypes with sparse regression*. NIMG 2012. **63**(3)1681-94.
7.   Purcell, S et al., *Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits*. Bioinformatics, 2003. **19**(1): p. 149-50.
8.   Menashe, I, et al., *PGA: power calculator for case-control genetic association analyses*. BMC Genetics, 2008. **9**: p. 36.
9.   Guey, LT, et al., *Power in the phenotypic extremes: a simulation study of power in discovery and replication of rare variants*. Genet Epid, 2011.
10. Minikel, EV *Power for GWAS and extreme phenotype studies*. 2012 [cited 2013 2/8/2013]:http://www.cureffi.org/2012/12/05/power-for-gwas-and-extreme-phenotype-studies/.
11. Feng, S, et al., *GWAPower: a statistical power calculation software for genome-wide association studies with quantitative traits*. BMC Genet, 2011. **12**: p. 12.
12. Ambrosius, WT, et al, *Power for genetic association studies with random allele frequencies and genotype distributions*. Am J of Hum Genet, 2004. **74**(4): p. 683-93.
13. Kozlitina, J, et al., *Power of genetic association studies with fixed and random genotype frequencies*. Annals of Human Genetics, 2010. **74**(5): p. 429-38.
14. Jahanshad N, et al., *Multi-site genetic analysis of diffusion images and voxelwise heritability analysis: A pilot project of the ENIGMA-DTI working group*. Neuroimage. In Press, 2013
15. Yang, J, et al., *Common SNPs explain a large proportion of the heritability for human height*. Nat Genet, 2010. **42**(7): p. 565-9.
16. Trikalinos, TA et al., *Impact of violations and deviations in Hardy-Weinberg equilibrium on postulated gene-disease associations*. Am J Epid 2006 **163**(4)300-9

17.   Chan, EK et al., *The combined effect of SNP-marker and phenotype attributes in genome-wide association studies*. Animal Genet, 2009. **40**(2) 149-56.

18.   Stein, JL et al., *Voxelwise genome-wide association study (vGWAS)*. NIMG 2010. **53**(3) 1160-74.
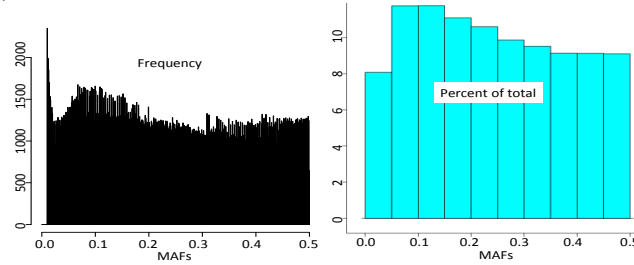
Figure 1 Histograms of frequency and percent ($\Delta=0.001,0.05$) of SNP MAFs in ADNI2.
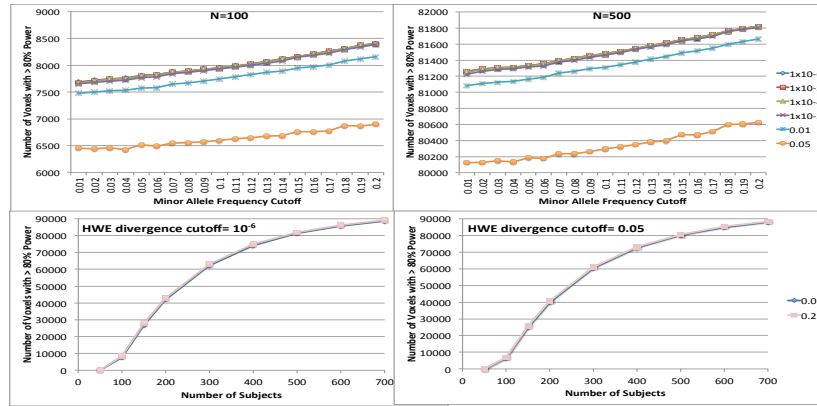

Figure 2: The *top row* shows the number of voxels with sufficient power to detect a genetic effect (>80% power) with respect to the MAF cutoff (MAF_c) and the divergence threshold for HWE (HWE_c), for samples with N=100 and N=500 subjects. As the MAF threshold is increased, fewer SNPs are tested, and this reduces the multiple comparisons correction and increases power for those tested SNPs, if there is a true associated variant in the group. As in prior GWAS power studies, the *bottom row* shows the number of subjects, regardless of MAF_c or HWE_c, and has a far larger influence than either one. The number of voxels with sufficient power levels off here, as N > 500.
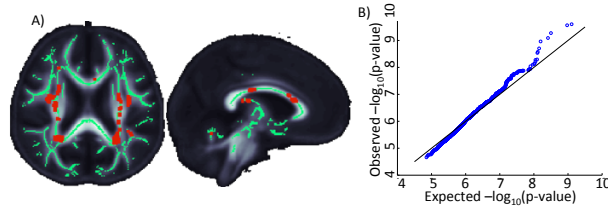

Figure 3: The 1664 voxels found to have sufficient power (> 80%) in the ADNI2 sample (N=78) were used for GWAS, and are *highlighted in red*. Plotting the minimum p-value at each voxel against what would be expected from the null hypothesis shows trends towards significance in this subsample of voxels.

Global changes in the connectome in autism spectrum disorders

Caspar J. Goch[a][⋆], Basak Oztan[b][⋆], Bram Stieltjes[c], Romy Henze[c,d], Jan Hering[a], Luise Poustka[e], Hans-Peter Meinzer[a], Bülent Yener[b] and Klaus H. Maier-Hein[a,c]

[a]German Cancer Research Center, Medical and Biological Informatics, Heidelberg, Germany;
[b]Rensselar Polytechnic Institute, Computer Science Department, Troy, New York, United States of America;
[c]German Cancer Research Center, Quantitative Imaging-based Disease Characterization, Heidelberg, Germany;
[d]Heidelberg University Hospital, Child and Adolescent Psychiatry, Section Disorders of Personality Development, Heidelberg, Germany;
[e]Department of Child and Adolescent Psychiatry and Psychotherapy, Central Institute of Mental Health, Mannheim, Germany
[⋆] These authors contributed equally to this work.

Abstract. There is an increasing interest in connectomics as means to characterize the brain both in healthy controls and in disease. Connectomics strongly relies on graph theory to derive quantitative network related parameters from data. So far only a limited range of possible parameters have been explored in the literature. In this work, we utilize a broad range of global statistic measures combined with supervised machine learning and apply it to a group of 16 children with autism spectrum disorders (ASD) and 16 typically developed (TD) children, which have been matched for age, gender and IQ. We demonstrate that 86.7% accuracy is achieved in distinguishing between ASD patients and the TD control using highly discriminative graph features in a supervised machine learning setting.

Keywords: connectomics, network analysis, diffusion imaging, autism, classification

1 Introduction

The past decades have seen an increasing interest in using diffusion weighted imaging to examine the way the human brain is connected [1]. Differences in these connections have been found for many mental illnesses, e.g. autism spectrum disorders (ASD) [2]. These techniques have mainly been used to look at the integrity of single tracts [3] or a few global measures of the connectome, especially small-worldness and the clustering coefficient [4]. A few recent studies also looked at local changes in different brain areas, especially those related to speech [5, 6]. A disadvantage of these approaches is the need for anatomical knowledge about

the areas of interest and the lack of global information. Therefore, changes not localized in these specific areas are typically overlooked.

Recent studies focused on a limited number of measures to characterize the changes induced by ASD in connectome. However, the understanding and diagnosis of ASD can be improved upon a comprehensive evaluation of the connectome topology with a large number of network features at global scale. Global scale graph features are shown to successfully characterize structure-function relationships in various biological systems. Specifically, in histopathological image analysis and tissue modeling applications, *cell-graphs* are utilized for the computer-aided diagnosis of brain, breast, and bone cancers [7–10] and also for the modeling of stem cells [11], cell-mediated collagen remodeling [12], and salivary gland branching morphogenesis [13]. In this paper, we extend quantitative connectomics by investigating the roles of global graph features in capturing ASD induced changes. We demonstrate that support vector machines based supervised learning achieves 86.7% accuracy in classifying the ASD and TD connectomes.

2 Materials and Methods

Data Acquisition: Evaluation was performed on a group of 18 right-handed children (16 male and 2 female) with a mean (standard deviation) chronological age of 9.7 (2.1) with a diagnosis of Asperger Syndrome or High Functioning Autism. The control group of 18 typically developed children of age 9.7 (1.9) was matched for age, sex and IQ. Data acquisition was done using a 1.5 T scanner (Siemens Avanto). T1 images for parcellation were taken with the following settings: MPRAGE TR/TE/TI/α = 1.9 s/4 ms/1.1 s/8°, FOV = 256 × 256 mm^2, matrix = 256 × 256, scan time 6 min). Diffusion weighted imaging was performed using single shot EPI with a dual bipolar diffusion gradient and a double spin echo for reduction of eddy currents with the following parameters: TR/TE 4700/78, FOV 192 mm, data matrix of 96 × 96 yielding an in-plane resolution of 2.0 mm, 50 axial slices with a thickness of 2.0 mm and no gap, with six gradient directions (b=1000 s/mm^2) and a b=0 image. This scheme was repeated 15 times.

Preprocessing and Fiber Tracking: The entire image processing pipeline is depicted in Fig. 1. The T1 weighted image was used to create a parcellation of the brain using freesurfer [14] as well as a binary mask of the brain. DWI data was motion and eddy-current corrected using FSL [15]. Q-ball images were then generated using solid angle reconstruction as provided by MITK [16]. Fiber tractography was performed on the q-ball images using the global tractography approach as presented by Neher et al. [17] using the brain mask to restrict the search space for possible fibers. To evaluate the robustness of our chosen tractography algorithm and the influence of the probabilistic tracking on our results we did four independent trackings for each patient. The same settings were used for each tracking: 10^8 iterations, particle length of 3.7 mm, particle width of 0.1 mm, particle weight of 0.0015, start temperature of 0.1, end temperature of

0.001, energy balance of 0, minimal fiber length of 20 mm and curvature threshold of 45°. Two patients and two controls were excluded due to heavy image artifacts and a resulting failure of the processing pipeline.
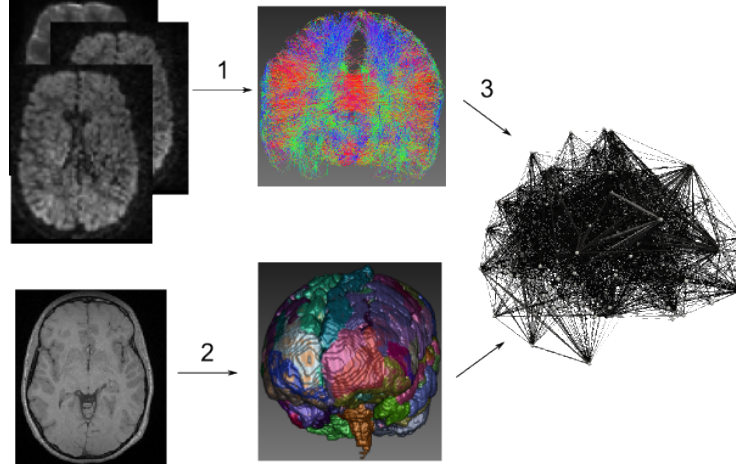

Fig. 1. Preprocessing pipeline. 1. Diffusion images are used to create a fiber image. 2. The anatomical MR image is used to create a parcellation of the brain. 3. The parcellation and the fiber image are used to create a network.

Network Generation: Connectivity matrices were created from the tractography result and the parcellation. DWI data and T1 images were registered using ANTs[1] for affine registration. Each label of the freesurfer segmentation was represented by one node if at least one fiber originated or ended within it. Two nodes were linked by an edge if at least one fiber connected the corresponding volumes. If a fiber could not be assigned two different non-white-matter labels it was disregarded. After network creation edges between nodes were eliminated if they represented less than N fibers to remove connections induced by noise. For this purpose, we performed a parameter search where we varied N between 14 and 30 with steps equal to 1 and select the value that yield the highest classification accuracy. Our analysis indicated $N = 26$ corresponded to the best classification performance.

Extraction of Connectome Features: We extracted 32 features for each patient's connectome. These features quantify the compactness, clustering, and spatial uniformity of the hypothesized connections within the brain. Graph features and their explanations are given in Table 1.

---

[1] http://www.picsl.upenn.edu/ANTS/

Table 1. Extracted graph features and their descriptions.

Feature Name	Description
Number of Nodes	Number of regions in brain
Number of Edges	Number of hypothesized communications
Average Degree	Number of edges per node
Clustering Coefficient C	Ratio of total number of edges among the neighbors of the node to the total number of edges that can exist among the neighbors of the node per node
Clustering Coefficient D	The average of the ratio of the links a node's neighbors have in between to the total number that can possibly exist
Clustering Coefficient E	Ratio of total number of edges among the neighbors of the node to the total number of edges that can exist among the neighbors of the node per node excluding the isolated nodes
Average Eccentricity	Average of node eccentricities, where the eccentricity of a node is the maximum shortest path length from the node to any other node in the graph
Diameter	Maximum of node eccentricities
Radius	Minimum of node eccentricities
Average Path Length	Average distance between the nodes of a graph, where the distance between two nodes is the number of edges in the shortest path that connects them
Average Betweenness	Average of node betweenness, where the betweenness of a node is the number of shortest paths from all nodes to all others that pass through that node
Giant Connected Component Ratio	Ratio between the number of nodes in the largest connected component in the graph and total the number of nodes
Number of Connected Components	Number of clusters in the graph excluding the isolated nodes
Average Connected Component Size	Number of nodes per connected component
Percentage of Isolated Points	Percentage of the isolated nodes in the graph, where an isolated node has a degree of 0
Percentage of End Points	Percentage of the end nodes in the graph, where an end node has a degree of 1
Number of Central Points	Number of nodes within the graph whose eccentricity is equal to the graph radius
Percentage of Central Points	Percentage of nodes within the graph whose eccentricity is equal to the graph radius
Spectral Radius	Largest valued eigenvalue of adjacency matrix
Second Largest	Second largest values eigenvalue of adjacency matrix
Adjacency Trace	Sum of the eigenvalues of adjacency matrix
Adjacency Energy	Sum of the squares of eigenvalues of adjacency matrix
Spectral Gap	Number of 0 valued eigenvalues of adjacency matrix
Laplacian Trace	Sum of the eigenvalues of laplacian matrix
Laplacian Energy	Sum of the squares of eigenvalues of laplacian matrix
Number of 0s	Number of eigenvalues that are equal to 0 in normalized laplacian matrix
Number of 1s	Number of eigenvalues that are equal to 1 in normalized laplacian matrix
Number of 2s	Number of eigenvalues that are equal to 2 in normalized laplacian matrix
Lower slope	The slope of the line fitted for the eigenvalues of the normalized laplacian matrix that are between 0 and 1 when sorted
Upper slope	The slope of the line fitted for the eigenvalues of the normalized laplacian matrix that are between 1 and 2 when sorted
Normalized Laplacian Trace	Sum of the eigenvalues of normalized laplacian matrix
Normalized Laplacian Energy	Sum of the squares of eigenvalues of normalized laplacian matrix

Classification and Validation: Support vector machine (SVM) classification was employed for the classification of the two groups. Though alternate supervised learning techniques may also be utilized, as we shall see in the next section, SVM classifier yielded the highest classification accuracy among the other well known candidates. We used radial basis function, also referred to as Gaussian kernel, in the form of $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\frac{|\mathbf{x}_i - \mathbf{x}_j|^2}{2\sigma^2})$ to transform the increase in the dimensionality of the data for better separability. We performed a parameter search to identify σ that achieves the highest classification accuracy. We sought σ in the set of candidate values that varied from 1.0 to 6.0 with 0.1 steps and determined that σ equaling 3.6 achieved the best performance in the identification of the patient's state.

The data is normalized so that the features have zero mean and unit variance to reduce the scale differences within different features. In order to obtain unbiased performance estimates, patient-based leave-one-out cross-validation was performed. The feature set was first divided into 32 disjoint partitions for each patients data. For each patient, a classifier was trained with the remaining 31 patients data and then tested on the retained data. The results for each patient were then combined to find the overall classification accuracy.

3 Results

We generated brain connectome networks as described previously for 32 patients each of which with four independent trackings. We then characterized the graphs using the 32 features described in Table 1 and using SVM classifier with RBF kernel we discriminated ASD patients from TD control with leave-one-patient-out cross-validation.

Given the large number of features, we performed feature selection based on *t-statistic* to identify the most discriminative features. For a given feature i, the *t*-statistic to test whether the population means are different is calculated as

$$t(i) = \frac{|\mu_1(i) - \mu_2(i)|}{\sqrt{\frac{\sigma_1^2(i)}{N_1} + \frac{\sigma_2^2(i)}{N_2}}} \tag{1}$$

where $\mu_k(i)$, $\sigma_k(i)$, and N_k are the sample mean, standard deviation, and size of the kth class ($k \in \{1, 2\}$) for ith feature, respectively. The features with high discriminative power get higher score. We tested the grading accuracy of the feature sets constituted by the first M most discriminative features. We varied M from 1 to 32, and report the grading accuracy in Fig. 2. It is seen that a classification accuracy of 86.72% can be achieved using the top four or five features. When we investigated the results of this case, it is seen that 9 out of 64 ASD trackings were identified as TD control and eight out of 64 TD control trackings were classified as ASD, and the rest of the trackings were classified accurately.

In order to compare our result to our earlier study that only considered the betweenness centrality of speech related locations in the brain [6], we also performed classification using the average betweenness centrality alone. Our result
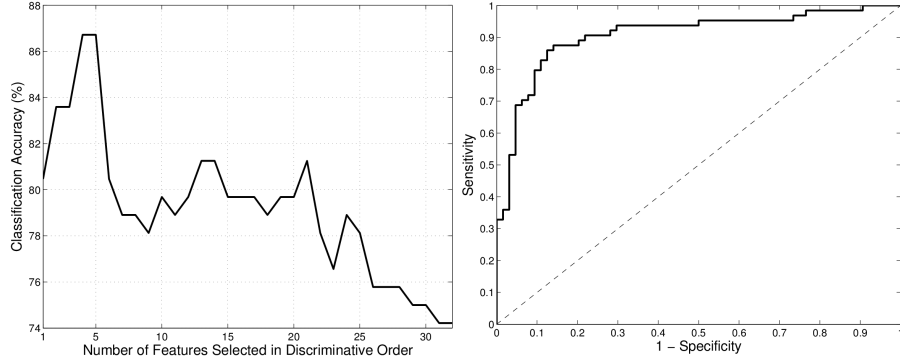
Fig. 2. Left: Influence of the number of discriminative features selected for classification on the classification accuracy. Highest grading accuracy achieved either the top four or five features selected. **Right:** Receiver operating characteristics for the SVM classifier with RBF kernel. The area under the curve is 0.9067.

showed 78.9% classification accuracy can be achieved using this feature alone. It is clear that considering additional features improved the classification accuracy significantly.

Independent of the learning method, we could achieve a consistent classification accuracy over 80%. Table 2 compares the classification accuracies of different classification methods. It is clear that SVM classifier achieves the highest overall accuracy in identifying the patient's neurological state. This is not unexpected as SVM classifiers are known to be highly successful in biomedical applications [18].

Table 2. Classification accuracy for different learning methods. SVM with RBF kernel yields the highest classification accuracy.

Learning Method	Classification Accuracy (%)
Support Vector Machines (RBF Kernel)	**86.72**
Support Vector Machines (Linear Kernel)	85.16
Linear Discriminant Analysis	84.38
Naïve Bayes Classifier	78.13
AdaBoost (Decision Stumps)	81.25

We then investigated how often a feature was in the top five of features for classification for a range of thresholds where the discriminative influence of each feature was given by t-statistic. Table 3 shows the frequency of discriminative features that appear in the top five feature for different thresholds. The Giant Connected Component Ratio was consistently a discriminative feature for every threshold in the range. For threshold $N = 26$, with the highest classification accuracy the top five features with the highest t-statistics were Clustering Coef-

ficient D, Giant Connected Component Ratio, Average Connected Component Size, Normalized Laplacian Trace, and Normalized Laplacian Energy.

Table 3. Histogram of highest discriminative features where the frequency shows the number of times the feature was in the top five discriminative features according to t-statistic for a link threshold (N) ranging from 14 to 30.

Feature	Frequency
Giant Connected Component Ratio	17
Clustering Coefficient D	16
Normalized Laplacian Trace	15
Average Connected Component Size	11
Normalized Laplacian Energy	10
Second Largest Eigenvalue Adjacency	7
Clustering Coefficient C	4
Average Betweenness Centrality	1

Finally, we give the receiver operating characteristics (ROC) to evaluate the performance of the classification. ROC curve plots the *sensitivity* against the $1-specificity$ at different threshold settings. For the SVM classifier, we used the distance from the maximum-margin hyperplane as the decision threshold. Figure 2 shows the ROC curve for our classifier. The area under the curve (AUC) is 0.9067, which is considered as a well-discriminating classifier.

4   Discussion

We show that global connectome features are useful to divide a group into patients suffering from ASD and healthy controls with good accuracy. A range of features, which have been neglected in the literature so far can be a valuable tool in identifying changes in the structure of the connectome.

Our patients have been matched for IQ and as such provide a sample of ASD that is closest to a normal population and as such presents the most prominent challenge considering classification in the context of ASD. In this light, our classification results are surprisingly good. Thus, quantitative connectomics may provide a powerful tool to further the understanding of the functioning of the human brain, both under normal conditions as well as in disease.

Identification of the features of the connectome which are consistently and significantly affected in disease using the full power of network graph analysis is an important step in this direction.

References

1. Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci. 2009 March;10:186–198.

2. Travers BG, Adluru N, Ennis C, Tromp DPM, Destiche D, Doran S, et al. Diffusion Tensor Imaging in Autism Spectrum Disorder: A Review. Autism Research. 2012;.

3. Lewis WW, Sahin M, Scherrer B, Peters JM, Suarez RO, Vogel-Farley VK, et al. Impaired Language Pathways in Tuberous Sclerosis Complex Patients with Autism Spectrum Disorders. Cerebral Cortex. 2012;.

4. Griffa A, Baumann PS, Thiran J-P, Hagmann P. Structural connectomics in brain diseases. NeuroImage. 2013;.

5. Li H, Xue Z, Ellmore TM, Frye RE, Wong STC. Network-based analysis reveals stronger local diffusion-based connectivity and different correlations with oral language skills in brains of children with high functioning autism spectrum disorders. Human Brain Mapping. 2012;.

6. Goch CJ, Stieltjes B, Henze R, Hering J, Meinzer HP, Fritzsche KH. Reduced centrality of Wernicke's area in autism. Proceedings of SPIE Medical Imaging 2013: Biomedical Applications in Molecular, Structural, and Functional Imaging. 2013; p. 867223–867223–6.

7. Demir C, Gultekin S, Yener B. Learning the topological properties of brain tumors. IEEE/ACM Transactions on Computational Biology and Bioinformatics **2**(3), 262–270 (2005);.

8. Bilgin C, Demir C, Nagi C, Yener B. Cell-graph mining for breast tissue modeling and classification. Proc. of 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 5311–5314, IEEE (2007);.

9. Bilgin C, Bullough P, Plopper G, Yener B. Ecm-aware cell-graph mining for bone tissue modeling and classification. Data mining and knowledge discovery **20**(3), 416–438 (2010);.

10. Oztan B, Shubert K, Bjornsson C, Plopper G, Yener B. Biologically-Driven Cell-Graphs for Breast Tissue Grading. Proc. of IEEE Symp. on Biomed. Imaging, 141–144 (2013);.

11. Lund AW, Bilgin CC, Hasan MA, McKeen LM, Stegemann JP, Yener B, Zaki MJ, Plopper GE. Quantification of spatial parameters in 3D cellular constructs using graph theory. Journal of Biomedicine and Biotechnology. (2009);.

12. Bilgin CC, Lund AW, Can A, Plopper GE, Yener B. Quantification of three-dimensional cell-mediated collagen remodeling using graph theory. PloS one **5**(9), (2010);.

13. Dhulekar N, Bange L, Baskaran A, Yuan D, Oztan B, Yener B, Ray S, Larsen M. A novel dynamic graph-based computational model for predicting salivary gland branching morphogenesis. Bioinformatics and Biomedicine (BIBM), 2012 IEEE International Conference on, 1–8 (2012);.

14. Fischl B, van der Kouwe A, Destrieux C, Halgren E, Ségonne F, Salat DH, et al. Automatically Parcellating the Human Cerebral Cortex. Cerebral Cortex. 2004;14(1):11–22.

15. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. FSL. NeuroImage. 2012;62(2):782 – 790.

16. Fritzsche KH, Neher PF, Reicht I, van Bruggen T, Goch C, Reisert M, et al. MITK Diffusion Imaging. Methods of Information in Medicine. 2012;.

17. Neher PF, Stieltjes B, Reisert M, Reicht I, Meinzer HP, Fritzsche KH. MITK Global Tractography. SPIE Medical Imaging 2012: Image Processing. 2012;.

18. Yang ZR. Biological applications of support vector machines. Briefings in bioinformatics **5**(4), 328–338 (2004);.