NeuroImage 77 (2013) 44–51

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Technical Note

Measuring network's entropy in ADHD: A new approach to investigate neuropsychiatric disorders

João Ricardo Sato [a], Daniel Yasumasa Takahashi [b], Marcelo Queiroz Hoexter [c], Katlin Brauer Massirer [d], André Fujita [e,*]

[a] Center of Mathematics, Computation, and Cognition, Federal University of ABC, Rua Santa Adélia, 166, Santo André, SP, 09210-170, Brazil
[b] Department of Psychology and Neuroscience Institute, Green Hall, Princeton University, NJ 08540 Princeton, USA
[c] Laboratório Interdisciplinar de Neurociências Clínicas (LiNC), Department of Psychiatry, Federal University of São Paulo, São Paulo, Brazil
[d] Center for Molecular Biology and Genetics Engineering, University of Campinas, Av. Candido Rondon, 400, CP6010, Campinas, SP, 13083-875, Brazil
[e] Department of Computer Science, Institute of Mathematics and Statistics, University of São Paulo, Rua do Matão, 1010, Cidade Universitária, São Paulo, SP, 05508-090, Brazil

ARTICLE INFO

Article history:
Accepted 20 March 2013
Available online 6 April 2013

Keywords:
ADHD
Graph
Spectral analysis
Entropy
fMRI

ABSTRACT

The application of graph analysis methods to the topological organization of brain connectivity has been a useful tool in the characterization of brain related disorders. However, the availability of tools, which enable researchers to investigate functional brain networks, is still a major challenge. Most of the studies evaluating brain images are based on centrality and segregation measurements of complex networks. In this study, we applied the concept of graph spectral entropy (GSE) to quantify the complexity in the organization of brain networks. In addition, to enhance interpretability, we also combined graph spectral clustering to investigate the topological organization of sub-network's modules. We illustrate the usefulness of the proposed approach by comparing brain networks between attention deficit hyperactivity disorder (ADHD) patients and the brain networks of typical developing (TD) controls. The main findings highlighted that GSE involving sub-networks comprising the areas mostly bilateral pre and post central cortex, superior temporal gyrus, and inferior frontal gyri were statistically different (p-value = 0.002) between ADHD patients and TD controls. In the same conditions, the other conventional graph descriptors (betweenness centrality, clustering coefficient, and shortest path length) commonly used to identify connectivity abnormalities did not show statistical significant difference. We conclude that analysis of topological organization of brain sub-networks based on GSE can identify networks between brain regions previously unobserved to be in association with ADHD.

Introduction

Attention deficit hyperactivity disorder (ADHD) is a behavioral disorder that usually begins in childhood and often persists in adults. ADHD affects at least 3–5% of children globally (Nair et al., 2006). This condition is currently diagnosed by the combination of structured diagnostic interviews and assessments of cognitive, social, school, and family functioning (Pastor and Reuben, 2008). However, the available measurements are still considered subjective. In the search for objective biological markers and more quantitative analyses, we are investigating the application of computational and mathematical models to assist in the diagnosis of ADHD.

Functional brain mapping techniques, such as functional magnetic resonance imaging (fMRI), exploit recent advances of neuroimaging technology data acquisition. These techniques brought remarkable innovation in the search for investigation of neural modules associated with cognitive functions, also called functional brain mapping. On the other hand, specialization in neural modules is only one aspect of brain functioning. Given that the brain is composed of several anatomical interconnected modules, functional segregation in spatially distributed networks also plays an important role in neurodevelopment and cognitive function (Fair et al., 2007).

The investigation of large-scale brain networks has recently emerged as an exciting method to identify neural anatomical substrates of behavior, cognition and psychiatric disorders (Hagmann et al., 2012). Blood-oxygen-level-dependent (BOLD) signals from fMRI, for example, are used to extract information about neural inter-connectivity in the brain by correlating activity among distant brain regions. Graph modelling has become a very successful approach in the investigation of brain connectivity networks using fMRI data (Bullmore and Sporns, 2009). Rubinov and Sporns (2010, 2011) have suggested several descriptive measures of complex networks such as centrality (e.g.: betweenness) and modularity (e.g.: clustering) coefficients to explore topological features of brain networks. Two other approaches recently introduced by Zalesky et al. (2012a, 2012b) allow

* Corresponding author at: Rua do Matão, 1010, Cidade Universitária, São Paulo, SP, 05508-090, Brazil.
E-mail address: fujita@ime.usp.br (A. Fujita).

1053-8119/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2013.03.035

statistical inferences on graph differences between groups and correlation analysis with continuous data (e.g.: behavioral scores). During the development of their graph models, the authors were especially concerned about controlling the family-wise error of multiple comparisons. They introduced concepts as the Network Based Statistics (NBS) and Spatial Pairwise Clustering (SPC). As a direct application to analysis of resting state fMRI data, Zang et al. (2004) proposed an intra-regional functional connectivity measure named Regional Homogeneity (ReHo). Furthermore, Hagmann et al. (2012) and Power et al. (2010) applied complex network analysis to describe the brain networks during neurodevelopmental maturation stages.

The use of network modelling combining graph analysis and neuroimaging was also recently applied to the characterization of atypical neurodevelopment processes. Investigations performed by Lynall et al. (2010) and Fornito et al. (2012) demonstrated the suitability of graph-based approach when they evaluated networks of patients with schizophrenia. In addition, the application of graph theory has been very successful in the study of attention-deficit/hyperactivity disorder (ADHD). Wang et al. (2009) described differences in small-world measures in children with ADHD when compared to typical development (TD) controls. Fair et al. (2007) have identified the neural substrates associated to control networks that may contribute to the high heterogeneity of ADHD, using the community detection method. More recently, Tomasi and Volkow (2012) have used a data-driven graph theory approach to investigate functional connectivity between a large sample of ADHD children and TD controls. Higher connectivity was found in reward-motivation regions such as the ventral striatum and orbitofrontal cortex. Opposing, lower functional connectivity was found in regions of the dorsal attention such as the superior parietal cortex and unexpected functional attributes of the region precuneus were observed comparing neuroimaging from ADHD patients versus controls.

Certainly, complex network analysis techniques can be used to enhance the comprehension of connective topology; however, most available graph modelling studies focused only on centrality and segregation measurements. Therefore, other characteristics of the graph have been neglected. One graph feature that could be investigated more is the complexity of the topological organization of sub-networks. This feature cannot be described by centrality neither by clustering measures, which were proposed or applied in previous cited studies. From an information theory perspective, data complexity is quantified by entropy measures. Thus, the assessment of the graphs' entropy can be used as a complementary approach to the existing methods, specifically to characterize the organization of brain networks.

In this study, we applied the graph's entropy measure to identify differences in the complexities of sub-networks. We illustrate the usefulness and applicability of the proposed approach by comparing the complexity of brain networks between typical developing controls (TD) and ADHD patients based on resting state fMRI data. The rationale for this comparison is our hypothesis that these two groups differ not only in local centrality measures but also in the complexity of the network organization. Indeed, we found statistically different (p-value = 0.002) entropy between ADHD patients versus TD controls in sub-networks comprising mostly bilateral pre and post central cortex, superior temporal gyrus, and inferior frontal gyri. The comparison with other conventional graph descriptors (betweenness centrality, clustering coefficient, and shortest path length) commonly used to identify connectivity abnormalities was not significantly different between the two groups.

Material and methods

Functional magnetic resonance imaging (fMRI) dataset

The dataset composed of resting state fMRI data from 638 children were downloaded from the ADHD-200 Consortium website. This database is publicly available at ADHD-200 website (The ADHD-200 Consortium, 2012). The sample set used for our study was composed of 479 typically developing controls (TD, 253 males, mean age ± standard deviation of 12.23 ± 3.26 years) and 159 ADHD patients (combined: hyperactive/impulsive and inattentive, 130 males, 11.24 ± 3.05 years). The research performed by ADHD-200 contributing sites was conducted with local Internal Review Board approval, and was in accordance with local Internal Review Board protocols. All data distributed via the International Neuroimaging Data-sharing Initiative were fully anonymized in compliance with the HIPAA Privacy Rules. Further details about this dataset can be obtained at the ADHD-200 consortium website.

Pre-processing of images

The fMRI dataset was pre-processed and functional networks of 351 brain's regions-of-interest (ROI) for each subject were estimated by Spearman's correlation (which is robust against outliers and monotonic non-linear relationships). The estimated functional network is clustered by spectral clustering algorithm (Luxburg, 2007; van den Heuvel et al., 2008) in order to identify partitions of the brain that are co-activated. Then, several features (betweenness centrality, clustering coefficient, shortest path length) for each cluster were quantified. These features were compared between the TD and ADHD groups in order to identify networks potentially related to the disorder.

The pre-processing of imaging data was performed using the Athena pipeline. The pre-processed data is available at the Neurobureau website. The pipeline focused on providing systematic processing of fMRI data, including the following steps: exclusion of the first four scans; slice timing correction; deoblique dataset; correction for head movements; masking the volumes to exclude non-brain regions; co-registration of mean image to the respective anatomic image of the subject; spatial normalization to MNI space (4 mm × 4 mm × 4 mm resolution); extraction of BOLD (Ogawa et al., 1990) time series from white matter (WM) and cerebrospinal-fluid (CSF); removing effects of WM, CSF, motion and trend using linear multiple regression; temporal band-pass filter ($0.009 < f < 0.08$ Hz); spatial smoothing the filtered data using a Gaussian filter (FWHM = 6 mm). The 351 regions-of-interest (ROI) considered as the graph nodes were defined by using the functional parcellation defined by the CC400 atlas (Craddock et al., 2012), obtained in a data-driven fashion by using the spectral clustering of the BOLD signals at distinct voxels. The average time series within the ROI was considered as the region representative.

Graphs

A *graph* is a pair of sets $G = (P,E)$, where P is a set of n vertices (or nodes) $v_1, \ldots, v_n$, *and* E is a set of m edges that connect two elements of P. Any undirected graph G with n nodes can be represented by its adjacency matrix A with $n \times n$ elements A_{ij}, whose value is $A_{ij} = A_{ji} = 1$ if nodes i and j are connected, and 0 otherwise.

In the current study, nodes and edges represent the 351 ROIs and the functional connectivity for each pair of ROIs, respectively. The functional connectivity was obtained by calculating the p-value corresponding to the Spearman's correlation for each pair of ROIs, with p-values corrected for False Discovery Rate (FDR) (Benjamini and Hochberg, 1995). The graph's adjacency matrices were constructed by adopting a p-value threshold of 0.05 after FDR correction, i.e., $A_{ij} = A_{ji} = 1$ if q-value < 0.05, and $A_{ij} = A_{ji} = 0$, otherwise. Thus, the graphs are undirected and the adjacency matrices A are symmetric.

Graph Spectral clustering

The problem of clustering ROIs (modules) consists of finding a partition of the network (graph) where different sub-networks (clusters) are lowly connected whereas the ROIs within a group are highly

connected. One well-known solution for this problem is the spectral clustering (Luxburg, 2007; van den Heuvel et al., 2008), described as following:

Input: Let W be the dissimilarity matrix of a graph G and k the number of desired clusters.

1. Compute the Laplacian matrix $L = D = W$, where D is the degree matrix, i.e., the diagonal matrix with the degrees $d_1, \ldots, d_n$ of the vertices $v_1, \ldots, v_n$, respectively, on the diagonal.
2. Compute the first k eigenvectors $u_1, \ldots, u_k$ of L.
3. Let $U \in \mathbb{R}^{n \times n}$ be the matrix containing the vectors $u_1, \ldots, u_k$ as columns.
4. For $i = 1, \ldots, n$, let $y_i \in \mathbb{R}^k$ be the vector corresponding to the i-th row of U.
5. Cluster the points $(y_i) i = 1, \ldots, n$ in $\mathbb{R}^k$ with the k-means algorithm into clusters $C_1, \ldots, C_k$.

Output: Clusters $C_1, \ldots, C_k$.

The number of clusters k was estimated by the silhouette method (Rousseeuw, 1987).

Descriptive measures

The comparison of functional connectivity patterns between classes of subjects requires descriptive measures that contain relevant or discriminative neurobiological information. Several network measures were recently developed and used (Rubinov and Sporn, 2010). Some of them are described as follows:

- *Average betweenness centrality*: The *betweenness centrality* (relative importance of a vertex within the graph) is a measure of a node's centrality in a network equal to the number of shortest paths from all vertices to all others that pass through that node (Freeman, 1978). The average betweenness centrality is the sum of the betweenness centralities divided by the number of vertices n.
- *Average clustering coefficient:* The *local clustering coefficient* for a vertex is given by the proportion of edges between the vertices within its neighborhood divided by the number of edges that could exist among them. The average clustering coefficient is the sum of the local clustering coefficients divided by the number of vertices n (Wasserman and Faust, 1994).
- *Average shortest path length:* The *shortest path length* between nodes v_i and v_j is the number of edges in a shortest path connecting them. The average shortest path length is the average of all the shortest path lengths for all pair of nodes v_i and v_j with $i \neq j$.

Graph spectrum and entropy

Another descriptive measure used to analyze networks is the concept of graph's spectrum entropy, introduced by Takahashi et al. (2012). The spectral density of the adjacency matrix of a graph has a tight relationship with the graph's structure. This idea was proposed by Takahashi et al. (2012) as a measure of the "uncertainty" of the graph.

The *spectrum* of an undirected graph G is the set of eigenvalues of its adjacency matrix A. A graph with n nodes has n real eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Given a family of random graphs g generated by some probability law, the eigenvalues are random vectors for which we can take the expectation with respect to the law of the graph's family. Takahashi et al. (2012) proposed that brain functional networks can be modeled as random graphs and they defined the *spectral density* of a general graph family g as

$$\rho_g(\lambda) = \lim_{n \to \infty} \left\langle \frac{1}{n} \sum_{j=1}^{n} \delta\left(\lambda - \lambda_j / \sqrt{n}\right) \right\rangle,$$

where δ is the Dirac delta function and the brackets $\langle\rangle$ indicate the expectation with respect to the law of the random graph. In order to estimate the spectral densities, first, the eigenvalues of the adjacency matrix are computed. Then, the eigenvalues' distribution is constructed by using a Gaussian kernel regression with the Nadaraya–Watson estimator (Nadaraya, 1964; Watson, 1964). The Nadaraya–Watson estimator is used to obtain a smoother histogram that improves the estimation. Then, the density is normalized to obtain the integral below the curve equal to one. The bandwidth of the kernel was chosen by $\left(\frac{\lambda_1 - \lambda_n}{\#\text{bins}}\right)$, where the number of bins (#bins) was selected by using the Sturges' criterion (Sturges, 1926).

Now, let ρ_g be the spectra of the adjacency matrix of a random graph g. The *spectral entropy* $H(\rho_g)$ is defined as (Takahashi et al., 2012)

$$H\left(\rho_g\right) = -\int_{-\infty}^{+\infty} \rho_g(\lambda) \log \rho_g(\lambda) d\lambda$$

where we assume $0 \log 0 = 0$. It is necessary to mention that the entropy defined above (differential entropy (Cover and Thomas, 2006) can assume negative values and without loss of generality, it can be applied to sub-graphs. The higher the entropy, the higher is the uncertainty, and the lower the entropy, the less uncertainty is the graph.

In order to provide an intuitive idea of graph entropy, we compute the approximate spectral entropy for the Erdös–Rényi random graph with parameter p as follows. For large n, the spectral density can be written as

$$\rho(\lambda) \sim \frac{\sqrt{4p(1-p) - \lambda^2}}{2\pi p(1-p)}$$

for $0 < |\lambda| < 2\sqrt{p(1-p)}$ and 0 otherwise. Using the above spectral density, we obtain the entropy of an Erdös–Rényi random graph as

$$H(\rho) \sim \frac{1}{2} 1n\left(4\pi^2 p(1-p)\right) - \frac{1}{2}.$$

By analyzing this formula, we verify that the maximum spectral entropy for the Erdös–Rényi graph is achieved for $p = 0.5$. This result is in accordance to our intuition that the Erdös–Rényi random graph with $p = 0.5$ is the one with the largest uncertainty. Moreover, notice that the entropy function is symmetric because the spectral density of the Erdös–Rényi graph generated with parameter p is equal to the spectral density of the Erdös–Rényi graph generated with parameter $1 - p$. The entropy achieves its minimum values for $p = 0$ and $p = 1$, which is the situation of the empty (without edges) and complete ($m = n^2$ edges) graphs, respectively. The empty graph is the one with the lowest entropy among graphs, followed by the complete graph. For instance, an empty graph has all its eigenvalues zero, i.e., 0^n (we will denote the multiplicities as exponents). For this case, equation 2 does not define directly the entropy, but interpreting its entropy as the limit of entropy of probability distributions increasingly concentrating on zero. The entropy value of an empty graph is $-\infty$. The spectrum of the complete graph is composed of $(n-1)^1$ and $(-1)^{n-1}$ eigenvalues. By taking $n \to \infty$, its entropy is also $-\infty$.

Another example is the $G_{1,n-1}$ complete bipartite graph, also called "star" due to its topology (graph with n nodes and $n-1$ edges where one node is connected to all other $n-1$ nodes). The spectra of $G_{1,n-1}$ is 0^{n-2}, $\pm\sqrt{n-1}$. By considering large networks, i.e., $n \to \infty$, its entropy is also $-\infty$. Therefore, by excluding trivial graphs such as empty and complete, the star graph is the one with the lowest entropy. Interestingly, scale-free networks (Barabási and Albert, 1999) with high scaling exponent have low entropy (the eigenvalue density has a peak in zero, i.e., most part of the eigenvalues

are concentrated in zero) and also present topologies similar to a star graph, where one vertex is the hub (highly connected) and a high number of vertices are connected to this hub. For more details, refer to (Takahashi et al., 2012).

Brain networks analysis

The fMRI data set of TD controls and ADHD was analyzed as described in the schema of Fig. 1.

For each subject, matrices of Spearman's correlation coefficients were calculated between the 351 ROIs. One hundred subjects with TD were randomly selected (training data) and the average correlation matrix was calculated, obtaining a representative connectivity matrix of the group with TD. These subjects are not included in further groups' comparison analysis (testing data). The dissimilarity matrix was constructed by calculating one minus the respective p-values of correlation coefficients. Notice that the higher is the absolute values of correlation coefficient, the lower is the p-value (for a fixed sample size), and consequently, the higher is the dissimilarity. The choice of the proposed dissimilarity measure instead of the standard one minus the correlation coefficient is due to the fact that we are interested in ROIs that are highly correlated, independent whether they present positive or negative correlation. Moreover, the use of p-value normalizes correlation coefficient by data variance.

The identification of modules of ROIs was carried out by applying the spectral clustering algorithm on this dissimilarity matrix. The number of clusters was determined by using the silhouette method (Rousseeuw, 1987).

Then, the labels of clusters identified in the set of 100 subjects (representative connectivity matrix) were applied to the remaining data set (testing data) composed of 538 subjects (379 subjects with TD + 159 with ADHD). The graph's adjacency matrices were constructed for data of each of the 538 children, using a p-value threshold of 0.05 after FDR (False Discovery Rate) correction (Benjamini and Hochberg, 1995), i.e., $A_{ij} = A_{ji} = 1$ if q-value < 0.05, and $A_{ij} = A_{ji} = 0$, otherwise. Notice that the descriptive features analyzed here (average betweenness centrality, average clustering coefficient, average shortest path length, and graph's entropy) are only applicable for unweighted graphs, i.e., graphs with adjacency matrices composed of 0s (absence of an edge) or 1s (presence of an edge). Therefore, it was necessary to discretize the dissimilarity matrices in order to obtain the adjacency matrices.

In the following, the averages of different descriptive measures of each sub-graph (average betweenness, clustering coefficient, average shortest path length, and entropy) were statistically compared between TD and ADHD by using the General Linear Model. Gender, age and site of image acquisition were included as covariates in order to remove their effects.

Results

Number of clusters

The identification of sub-networks associated to ADHD subjects was based on clustering of the ROIs (see Material and Methods) using the spectral clustering algorithm. A preliminary step to every clustering algorithm including the spectral clustering method is the need to define the number of clusters k to be used in further analysis. Here, we estimated the number of clusters by the silhouette method (Rousseeuw, 1987). The silhouette measures the quality of clustering. A sudden change in the silhouette value occurring with an increase in the number of clusters indicates a breakdown of a cluster.

The number of clusters was plotted against the silhouette value as illustrated in Fig. 2. The sudden change in the silhouette value is represented by the largest gap between values associated to the number of clusters four and five. This gap defines the number of clusters in our dataset as being equal to four. Therefore, in our data-driven analysis, the average functional brain network is basically composed of four modules (sub-networks). The number of clusters was very robust with respect to the choice of the initial 100 individuals used to compute the clusters. The same procedure was repeated 10 times with samples represented by different sets of 100 subjects and in all the cases the number of chosen clusters was four.

Clustering the brain regions

After defining the number of clusters as being four, the ROIs were clustered into sub-networks as can be visualized in Fig. 3. The clusters

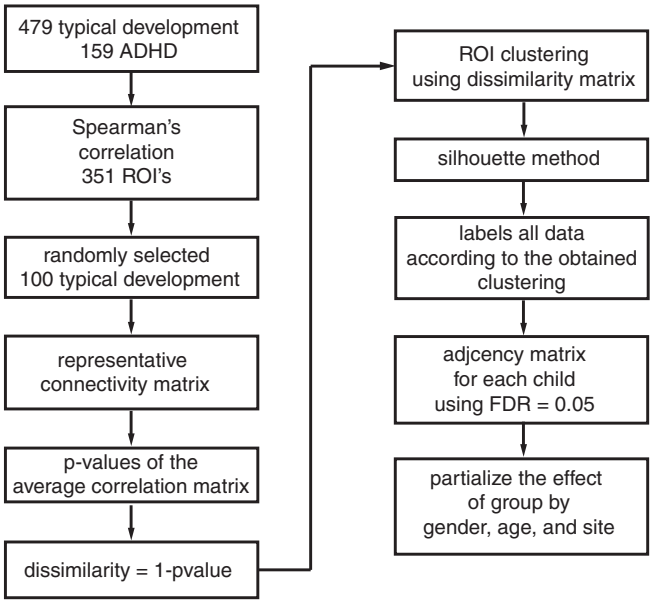

Fig. 1. Schema describing the procedures for statistical/computational analysis applied on children with TD and ADHD.

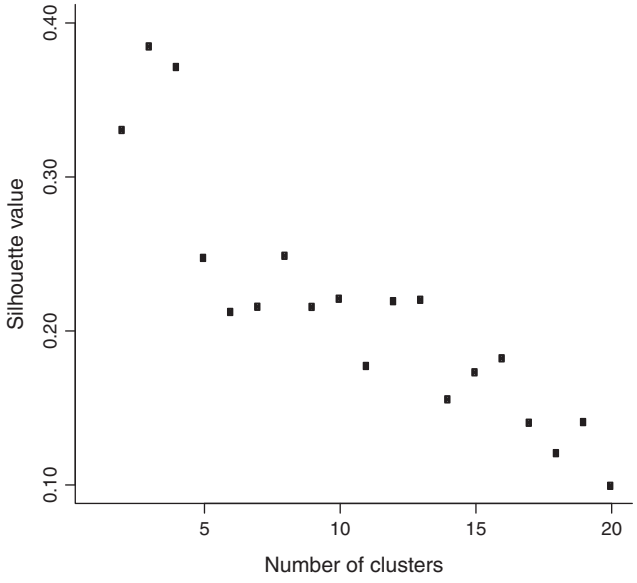

Fig. 2. Selection of the number of clusters. The number of clusters versus the silhouette value. Notice that the silhouette value for one cluster is not defined. The gap between the number of clusters four and five suggests that the number of clusters in this dataset is four.

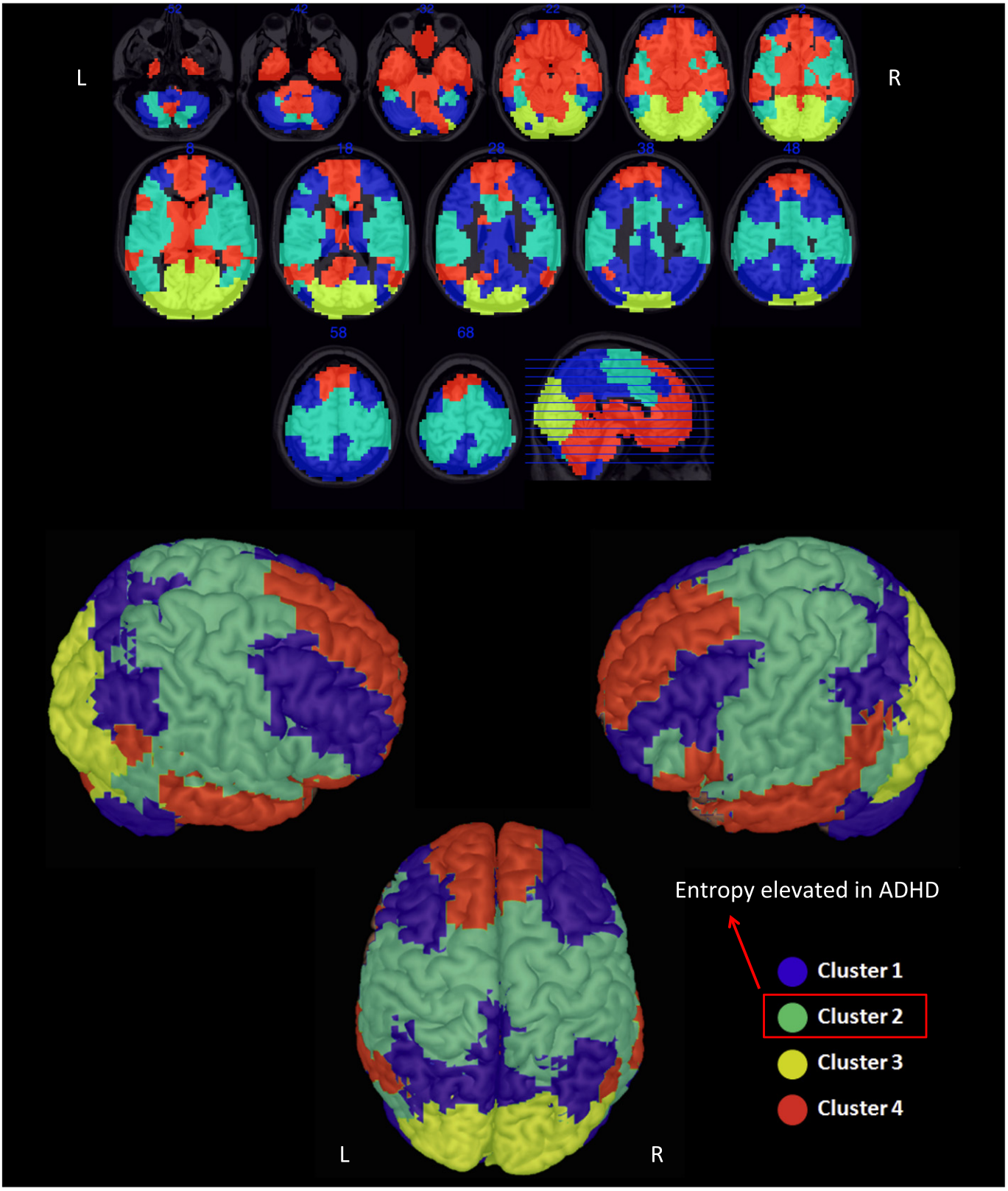

Fig. 3. Brain mapping of the functional ROIs' clusters. The entropy of cluster 2 is significantly reduced in ADHD when compared to typically developing children. This cluster is mostly composed of bilateral pre and post central cortex, superior temporal gyrus, and inferior frontal gyri.

one, two, three, and four are represented by the colors purple, green, yellow, and red, respectively. The number of time series is 97 for cluster 1, 99 for cluster 2, 38 for cluster 3, and 117 for cluster 4.

As for the previous analyses, 10 sets of 100 randomly chosen subjects did not affect the clustered regions significantly.

An interesting aspect of the clustering was the inclusion of anatomically contiguous and symmetric areas in the same clusters, although these constraints were not included *a priori* in the analyses. This is consistent with the hypothesis that the spectral clustering method groups areas with similar brain activities in the same cluster.

Identification of the sub-networks associated with ADHD

Finally, in order to identify sub-networks related to disease, we calculated the three commonly used descriptive measures namely betweenness centrality, clustering coefficient, shortest path length and the proposed graph entropy for each cluster of the control group and for the ADHD group. The values for the sub-networks' coefficients with the corresponding p-values are displayed on Table 1. The different descriptive measures (betweenness centrality, clustering coefficient, shortest path length, and the proposed GSE) of each sub-graph were statistically compared between TD and ADHD by using the General Linear Model. Gender, age and site of acquisition were included as covariates in order to eliminate their effects. The p-value of significance at 5% after Bonferroni correction for multiple tests is highlighted in bold. Only the entropy in cluster 2 showed statistically significant correlation in children diagnosed with ADHD and TD controls. Notice that the p-values shown on Table 1 are not corrected for multiple tests.

When evaluating the brain regions included in cluster 2, we observed that they were composed of: bilateral pre and post central cortex, superior temporal gyrus, and inferior frontal gyri. This cluster presents statistically higher entropy for children with ADHD than for children with TD (p-value = 0.002). The data show that the uncertainty or randomness in the sub-network structure of cluster 2 is significantly higher in children with ADHD than in children with TD, while other descriptive measures did not present evidences of statistical difference. Notice that the clustering of brain regions was fundamental for the identification of the differences between the two groups.

Thus, our results indicate that the use of GSE in combination with sub-graph identification, can find previously unobserved connections among brain regions, which are likely correlated to ADHD.

Discussion

Complex networks analysis is a promising tool to investigate the organization of the functional brain connectivity structure (Rubinov and Sporn, 2010). In the current study, we introduced a complexity measure to quantify the entropy of brain networks and sub-networks, as a complementary approach to known descriptors. In addition, we also demonstrate the usefulness of this novel approach in an illustrative application to fMRI data, obtained from a large database of TD controls and ADHD patients, under a resting state protocol.

It is important to discuss some limitations of the proposed approach. First, the graph of each subject was modelled using Spearman correlation coefficient, which is a pairwise measurement. Thus, the graph edges may not necessarily mirror direct connections but only indirect dependences. Furthermore, because Spearman correlation is a symmetrical measure, it only allows the construction of undirected graphs. Since the direction of information flow between neural nodes is useful to understand cognitive processing, an approach based on undirected graphs may not consider all the relevant information. Despite these limitations, the proposed approach was able to successfully identify differences in the network organization between controls and ADHD patients.

An expected challenge when using graph descriptors is the translation of the mathematical constructs into a meaningful neurophysiological interpretation in the context of brain networks. Entropy measures are traditionally associated to information, complexity and randomness (Cover and Thomas, 2006; Greven et al., 2003; Rissanen, 2007). In the case of networks, these characteristics mirror the ordering of randomness or organization of the edges connecting the nodes. For signalling pathways in breast cancer, for example, Teschendorff and Severini (2010) have used an entropy measure to quantify the degree of randomness in the pattern of information flow between different genes. The authors compared the entropies between non-metastatic and metastatic breast cancer networks, demonstrating that metastasis is characterized by a significant increase in the degree of randomness of local gene expression patterns. Analogously, the interpretation of entropy in the case of brain connectivity can be viewed as a coefficient to measure the degree of randomness of the patterns' information flow between distinct neural modules. High entropy levels could mean a lack of organization patterns, which may be associated to abnormal topologic structure or malfunctioning circuitry, which is expected in cases of neuropsychiatric disorders (Fornito et al., 2012).

In addition, as evidenced in the current study, the combination between graph clustering and entropy might be used to evaluate the organization of sub-graphs, which is suitable to investigate brain networks. This property is desired since neural networks can be described at different scales of resolution. Finally, it is important to emphasize that in the ADHD illustration, the entropy was the only descriptor to highlight significant differences between controls and ADHD patients, evidencing the importance of having a broad set of descriptors available. The analyses were carried out 10 times with different sets of data from children with TD as a training set in order to guarantee the reproducibility of the results.

Our analysis of ADHD patients suggests that the organization of sub-networks comprising bilateral pre and post central cortices, superior temporal gyrus, and inferior frontal gyrus (cluster 2) has greater entropy levels when compared to TD controls. This possibly means an abnormal connectivity structure of this circuit in patients. These regions, including premotor cortex and supplementary motor area, are frequently associated with motor/sensitive functions, auditory processing (Wernicke's area and primary auditory cortex), and response inhibition. In accordance with our findings, abnormal functional connectivity encompassing frontal and temporal cortices was also detected in a non-clinical sample of young adult ADHD patients compared to controls (Cocchi et al., 2012). Taken together, these results suggest the involvement of other brain circuits outside the frontal–striatal pathway in the physiopathology of ADHD (Castellanos and Proal, 2012).

Differences in the organization of motor systems and its implications in ADHD were discussed in Castellanos and Proal (2012). One study using a motor tapping task has shown that ADHD children presented hypoactivation of the primary motor cortex compared to controls (Mostofsky et al., 2006). Likewise, in a study combining paced and unpaced finger tapping task with fMRI, adult ADHD patients

Table 1
Descriptive measures and entropy measure obtained for each cluster of TD control samples and for each cluster of ADHD patients.

Cluster	Average betweenness centrality	Average clustering coefficient	Average path length	Entropy
#1	(0.115)	(0.926)	(0.121)	(0.025)
TD	19.27 ± 11.81	0.68 ± 0.07	1.40 ± 0.25	−661.68 ± 140.74
ADHD	22.10 ± 16.03	0.67 ± 0.09	1.46 ± 0.34	−640.27 ± 169.05
#2	(0.381)	(0.256)	(0.371)	**(0.002)**
TD	19.77 ± 11.09	0.68 ± 0.07	1.40 ± 0.23	**−633.94 ± 138.43**
ADHD	21.37 ± 14.10	0.69 ± 0.08	1.44 ± 0.29	**−590.57 ± 142n83**
#3	(0.722)	(0.457)	(0.716)	(0.353)
TD	4.91 ± 2.35	0.78 ± 0.08	1.27 ± 0.13	−313.00 ± 730.52
ADHD	5.25 ± 2.99	0.78 ± 0.08	1.28 ± 0.16	−272.68 ± 201.68
#4	(0.238)	(0.774)	(0.214)	(0.396)
TD	25.24 ± 15.41	0.65 ± 0.08	1.44 ± 0.27	−792.97 ± 234.16
ADHD	27.65 ± 20.14	0.65 ± 0.10	1.48 ± 0.36	−764.74 ± 467.69
All data	(0.092)	(0.395)	(0.092)	(0.026)
TD	68.97 ± 51.74	0.68 ± 0.09	1.39 ± 0.30	−5668.3 ± 5226.5
ADHD	79.51 ± 66.79	0.67 ± 0.12	1.45 ± 0.38	−5184.6 ± 8111.8

p-values (between brackets) and mean ± standard deviation for TD and ADHD. The descriptive measures are: average betweenness centrality, average clustering coefficient, average path length, and entropy. "All data" represents the p-values obtained when the neural network of the whole brain was analyzed, i.e., without partitioning into clusters. After Bonferroni correction for multiple comparisons, only the entropy of cluster 2 (in bold) presented statistical significance at a threshold of 5%.

showed less activity within several brain regions, such as temporal gyrus and the prefrontal and precentral gyri relative to controls (Valera et al., 2010). Aron et al. (2004) described and discussed evidences for a relevant role of the right inferior frontal cortex in inhibition of responses and task-sets (e.g.: go/non-go tasks). Since one of the main symptoms of ADHD patients is impulsivity, the involvement of an abnormal sub-network of these regions is expected. A meta-analysis of fMRI studies reported hypoactivity of several frontal regions including the inferior prefrontal cortex in ADHD patients compared to controls (Dickstein et al., 2006). Likewise, frontal hypoactivation within the precentral gyrus and also within the inferior frontal cortex was detected in fMRI studies performed in ADHD using response-inhibition tasks (Dickstein et al., 2006).

The network complexity of task-positive networks (cluster 1), which is frequently described in attention engagement, was not statistically different between patients and controls (although it would be significant if multiple comparison correction was not applied). In contrast, a recent study that investigated a similar sample of subjects obtained from the ADHD-200 Consortium found higher connectivity in brain regions involved in reward-motivation and lower functional connectivity in regions related to attention processes in ADHD patients compared to controls (Tomasi and Volkow, 2012). The discrepancies between these two studies might be explained in part by methodological differences involving the computational algorithms. Whereas Tomasi and Volkow (2012) have used a functional connectivity density mapping approach, we have performed a seed-voxel correlation analysis. Moreover, the fraction of ADHD subjects versus TD controls was lower in our investigation (24%) compared to the study performed by Tomasi and Volkow (2012) (44%), which may have reduced our power to detect differences in other brain networks.

We believe that results can vary according to the considered atlas for brain parcellation that was used to define the nodes of the graph. Since there is no best atlas for a general case, we considered the CC400 atlas (Craddock et al., 2012, available at the NeuroBureau website) for our studies due to its granular parcellation (351 regions than conventional ones (e.g., AAL) and its defined functional perspective. By using the graph spectral clustering, the whole brain graph based on this atlas was then split into four sub-graphs (the number of sub-graphs was determined in an objective manner by using the silhouette method) from which graph descriptors were extracted. This approach allowed us to obtain a spatial description of these sub-networks.

Future development of the method is still necessary in order to extend its use to direct and directed description. These applications may result in the increase of statistical power to detect differences, and can provide additional valuable information about brain functioning. Furthermore, graph entropy measure can be considered as a complementary approach to the analysis of resting state functional networks, when compared to other methods. An integration of the entropy coefficient with the Network Based Statistics (NBS) and Spatial Pairwise Clustering (SPC) approach (Zalesky et al., 2012a, 2012b) may be promising considering the analyses of brain functional connectivity (hypothesis testing), because this combination could permit network-based multiple comparison corrections (control family wise errors).

Conclusions

In our study we found a statistically significant difference in the graph entropy of the sub-networks comprising mostly bilateral pre and post central cortex, superior temporal gyrus, and inferior frontal gyrus (cluster 2) in ADHD patients compared to TD controls. We demonstrate the proposed method to successfully identify sub-graph's entropy. The method was able to differentiate brain regions with connectivity abnormalities in ADHD that other measures could not identify.

The main contributions of our work can be summarized as following: (i) we developed a framework to identify brain regions that are potentially related to medical disorders; (ii) we applied comparisons among standard descriptive measures and the proposed measure based on the graph spectral entropy to actual biological data, and (iii) we were able to identify brain regions that might be related to ADHD. Thus, the results indicate that the use of graph spectral entropy, in combination with sub-graph identification, can be a useful tool to investigate neuronal network abnormalities in neuropsychiatric disorders.

Acknowledgments

This work was supported by FAPESP, CNPq, and Pew Latin American Fellowship.

Conflict of interest

The authors declare that there are no conflicts of interest.

References

Aron, A.R., Robbins, T.W., Poldrack, R.A., 2004. Inhibition and the right inferior frontal cortex. Trends Cogn. Sci. 8, 170–177.
Barabási, A.-L., Albert, R., 1999. Emergence of scaling in random networks. Science 286, 509–512.
Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B 57, 289–300.
Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10, 186–198.
Castellanos, F.X., Proal, E., 2012. Large-scale brain systems in ADHD: beyond the prefrontal–striatal model. Trends Cogn. Sci. 16, 17–26.
Cocchi, L., Bramati, I.E., Zalesky, A., Furukawa, E., Fontenelle, L.F., Jorge, M., Tripp, G., Mattos, P., 2012. Altered functional brain connectivity in a non-clinical sample of young adults with attention-deficit/hyperactivity disorder. J. Neurosci. 32, 17753–17761.
Cover, T.M., Thomas, J.A., 2006. Elements of information theory, 2nd edition. Wiley-Interscience, NJ.
Craddock, R.C., James, G.A., Holtzheimer, P.E., Hu, X.P., Mayberg, H.S., 2012. A whole brain fMRI atlas generated via spatial constrained spectral clustering. Hum. Brain Mapp. 33, 1914–1928.
Dickstein, S.G., Bannon, K., Castellanos, F.X., Milham, M.P., 2006. The neural correlates of attention deficit hyperactivity disorder: an ALE meta-analysis. J. Child Psychol. Psychiatry 47, 1051–1062.
Fair, D.A., Dosenbach, N.U., Church, J.A., Cohen, A.L., Brahmbhatt, S., 2007. Development of distinct control networks through segregation and integration. Proc. Natl. Acad. Sci. U. S. A. 104, 13507–13512.
Fornito, A., Zalesky, A., Pantelis, C., Bullmore, E.T., 2012. Schizophrenia, neuroimaging and connectomics. Neuroimage 62, 2296–2314.
Freeman, L.C., 1978. Centrality in social networks: conceptual clarification. Soc. Networks 1, 215–239.
Greven, A., Keller, G., Warnecke, G., 2003. Entropy (Princeton Studies in Applied Mathematics). Princeton University Press, NJ.
Hagmann, P., Grant, P.E., Fair, D.A., 2012. Connectomics: a conceptual framework for studying the developing brain. Front. Syst. Neurosci. 6, 43.
Luxburg, U., 2007. A tutorial on spectral clustering. Stat. Comput. 17, 395–416.
Lynall, M.E., Bassett, D.S., Kerwin, R., McKenna, P.J., Kitzbichler, M., et al., 2010. Functional connectivity and brain networks in schizophrenia. J. Neurosci. 30, 9477–9487.
Mostofsky, S.H., Rimrodt, S.L., Schafer, J.G., Boyce, A., Goldberg, M.C., et al., 2006. Atypical motor and sensory cortex activation in attention-deficit/hyperactivity disorder: a functional magnetic resonance imaging study of simple sequential finger tapping. Biol. Psychiatry 59, 48–56.
Nadaraya, E.A., 1964. On estimating regression. Theory Probab. Appl. 10, 186–190.
Nair, J., Ehimare, U., Beitman, B.D., Nair, S.S., Lavin, A., 2006. Clinical review: evidence-based diagnosis and treatment of ADHD in children. Mo. Med. 103, 617–621.
Ogawa, S., Lee, T.M., Kay, A.R., Tank, D.W., 1990. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. Proc. Natl. Acad. Sci. U. S. A. 87, 9868–9872.
Pastor, P.N., Reuben, C.A., 2008. Diagnosed attention deficit hyperactivity disorder and learning diability: United States, 2004–2006. Vital Health Stat. 10, 237.
Power, J.D., Fair, D.A., Schlaggar, B.L., Petersen, S.E., 2010. The development of human functional brain networks. Neuron 67, 735–748.
Rissanen, J., 2007. Information and Complexity in Statistical Modelling (Information Science and Statistics). Springer, NY.
Rousseeuw, P., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Comput. Appl. Math. 20, 53–65.
Rubinov, M., Sporn, O., 2010. Complex network measures of brain connectivity: uses and interpretations. Neuroimage 52, 1059–1069.
Sturges, H.A., 1926. The choice of a class interval. J. Am. Stat. Assoc. 21, 65–66.
Takahashi, D.Y., Sato, J.R., Ferreira, C.E., Fujita, A., 2012. Discriminating different classes of biological networks by analyzing the graph spectra distribution. PLoS One 7, e49949.
Teschendorff, A.E., Severini, S., 2010. Increased entropy of signal transduction in the cancer metastasis phenotype. BMC Syst. Biol. 4, 104.

The ADHD-200 Consortium, 2012. The ADHD-200 consortium: a model to advance the translational potential of neuroimaging in clinical neuroscience. Front. Syst. Neurosci. 6, 62.
Tomasi, D., Volkow, N.D., 2012. Abnormal functional connectivity in children with attention-deficit/hyperactivity disorder. Biol. Psychiatry 71, 443–450.
Valera, E.M., Spencer, R.M., Zeffiro, T.A., Makris, N., Spencer, T.J., 2010. Neural substrates of impaired sensorimotor timing in adult attention-deficit/hyperactivity disorder. Biol. Psychiatry 68, 359–367.
van den Heuvel, M., Mandl, R., HH, 2008. Normalized cut group clustering of resting-state fMRI data. PLoS One 3, e2001.
Wang, L., Zhu, C., He, Y., Zang, Y., Cao, Q., et al., 2009. Altered small-world brain functional networks in children with attention-deficit/hyperactivity disorder. Hum. Brain Mapp. 30, 638–649.
Wasserman, S., Faust, K., 1994. Social Network Analysis: Methods and Applications. Cambridge University Press, Cambridge.
Watson, G.S., 1964. Smooth regression analysis. Sankya A 26, 359–372.
Zalesky, A., Fornito, A., Bullmore, E., 2012a. On the use of correlation as a measure of network connectivity. Neuroimage 60, 2096–2106.
Zalesky, A., Cocchi, L., Fornito, A., Murray, M.M., Bullmore, E., 2012b. Connectivity differences in brain networks. Neuroimage 60, 1055–1062.
Zang, Y., Jiang, T., Lu, Y., He, Y., Tian, L., 2004. Regional homogeneity approach to fMRI data analysis. Neuroimage 22, 394–400.

WEB REFERENCES

ADHD-200, A. http://fcon_1000.projects.nitrc.org/indi/adhd200/ (Accessed 2011 Aug 20).
NITRC, A. http://www.nitrc.org/plugins/mwiki/index.php/neurobureau:AthenaPipeline (Accessed 2011 Aug 25).
ADHD-200 Preprocessed data. http://neurobureau.projects.nitrc.org/ADHD200 Accessed 2011 Aug 20.