Impact of normalization and segmentation on brain graphs

The impact of normalization and segmentation on resting state brain networks

Ricardo Magalhães¹,²,³,⁴, Paulo Marques¹,²,³, José Soares¹,²,³, Victor Alves⁴, Nuno Sousa¹,²,³

¹ Life and Health Sciences Research Institute (ICVS), School of Health Sciences, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal.
² ICVS/3B’s - PT Government Associate Laboratory, Braga/Guimarães, Portugal.
³ Clinical Academic Center – Braga, Portugal
⁴ Department of Informatics, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal.

Corresponding author: Nuno Sousa, Life and Health Sciences Research Institute (ICVS), School of Health Sciences, University of Minho, Campus Gualtar, 4710-057 Braga, Portugal.
Tel: 351-253-604806. Fax: 351-253-604847. Email: njcsousa@ecsaude.uminho.pt

Keywords: Neuroimaging, fMRI, Graph-Theory, Pre-Processing, Brain parcellation, Normalization;
Impact of normalization and segmentation on brain graphs

Abstract

Graph theory has recently received a lot of attention from the neuroscience community as a method to represent and characterize brain networks. Still, there is a lack of a gold standard for the methods that should be employed for the preprocessing of the data and the construction of the networks, as well as a lack of knowledge on how different methodologies can affect the metrics reported. We used graph theory analysis applied to resting-state functional Magnetic Resonance Imaging (rs-fMRI) to investigate the influence of different node-defining strategies and the effect of normalizing the functional acquisition on several commonly reported metrics used to characterize brain networks. The nodes of the network were defined using either the individual FreeSurfer segmentation of each subject or the FreeSurfer segmented MNI (Montreal National Institute) 152 template, using the Destrieux and sub-cortical atlas. The functional acquisition was either kept on the functional native space or normalized into MNI standard space. The comparisons were done at three levels: on the connections, on the edge properties and on the network properties levels. Our results reveal that different registration and brain parcellation strategies have a strong impact on all the levels of analysis, possibly favoring the use of individual segmentation strategies and conservative registration approaches. In conclusion, several technical aspects must be considered so that graph theoretical analysis of connectivity MRI data can provide a framework to understand brain pathologies.
Impact of normalization and segmentation on brain graphs

Acronyms

fMRI-functional Magnetic Resonance Imaging;
MNI-Montreal National Institute;
AAL-Automated Anatomic Labeling;
MPRAGE - magnetization prepared rapid gradient echo;
TR-Repetition Time;
TE-Echo Time;
FoV-Field of View;
WM-White matter;
CSF-Cerebrospinal fluid;
Dof-degrees of freedom;
ROIS-regions of interest;
ATL-atlas;
NAT-native;
FS-FreeSurfer;
Impact of normalization and segmentation on brain graphs

1 Introduction

The brain remains to this day as one of the most enigmatic organs in the human body. The inherent difficulty to understand its wiring patterns and how they relate to different functions brought the necessity to explore new methods that can bring new lights on this subject [1]. One such method that has received increasing attention in the last decade is graph theory and the study of complex networks. One of the core strengths of graph theory in brain network modeling lies in its ability to quantitatively characterize the topological organization of the brain [2]. It can have a key role in the interpretation of the complex dynamics detected in the brain across different studies using different modalities and models. Starting in the 90’s (see for a review reference [3]), several key moments revolutionized the study of the brain connectivity, namely the introduction of the functional Magnetic Resonance Imaging (fMRI) technique [4], the first studies that demonstrated the potential of the analysis of brain connectivity patterns [5-7], the association of memory mechanisms to different mental states and connectivity patterns [8,9] and the introduction of the concept of small-world networks [10]. More recently, multiple studies have shown the potential of the application of graph theory to represent and characterize brain networks. While there are reports of the application of graph theory with different imaging techniques, such as Electroencephalography, Magnetoencephalography [11] and Positron Emission Tomography [12], MRI is by far the most common [13,14].

A graph is a mathematical abstraction that can be understood as a group of nodes, which represent the different elements of the network, linked by edges, representing some relationship between them [15]. Translating these components into a brain network is fairly simple in concept: each node can represent an anatomic unit of the brain (can theoretically range from individual cells to full anatomic structures) and the edges can represent different relationships between regions, like the statistical correlation of fMRI activations, the number/density of connecting fibers obtained through diffusion tractography or structural properties such as cortical thickness correlations. Despite the many advantages of using graph theory as a model for brain connectivity, some key issues still ensue. Building a brain graph means making a lot of decisions and assumptions that can directly affect the outcome and the results arising from different processing strategies cannot be easily compared [16]. It
Impact of normalization and segmentation on brain graphs

has been demonstrated that different properties of the network behave differently with connection type [17], network size, edge density and degree distribution [18], specially the small world parameters, depending on the number of nodes and the average degree of the network [19].

Different node defining strategies can vary greatly on the number of nodes used, ranging from very small networks with 19 nodes [20], to the most common number of 90 nodes from the AAL (Automated Anatomic Labeling) atlas [21-23], and building up to much bigger networks ranging from 998 nodes [24] and up to 10000 [25]. The choice of the nodes definition can also vary from the use of pre-segmented atlas in a standard space [26,27], the use of segmentation tools to individually parcellate each subject’s brain [23,28], brain activation clusters extracted from fMRI analysis [26,29] or simply to the use of all voxels in the image [25]. Such diversity of possibilities demands a thorough knowledge on how these might affect the characteristics of the network in all its range of possibilities.

Another key difference between studies is the choice of fundamental preprocessing steps: some studies normalize the images to a standard space [20,22,27,28] while others do not [25,26]; some apply spatial smoothing steps with differently sized kernels [20,26,28], while others do not [22,27]; some apply signal regression to remove confound signals [20,27,28], while others do not [22,23,25]; and finally different frequency filtering strategies can also be found [17,30,31].

There is an obvious need to have the functional images (originally on its native space) and the node defining images (usually on a standard space) in a common space. While the most common approach to achieve this is to normalize the functional images to the standard space, the opposite is still a valid, rather unused, option. Although, as hypothesized in another study [25], the interpolation of the voxels that occurs during this step can introduce artificial correlations, the effect of the normalization step on the networks is still largely unexplored.

On the brain parcellation strategies, the most common approach is to simply co-register the functional images and a fixed template, resuming the segmentation to a normalization step. A more elaborate alternative is to employ tools to individually segment each subject, such as FreeSurfer [32]. This strategy takes more factors into consideration (such as surface matching and gyri and sulci...
Impact of normalization and segmentation on brain graphs

identification) and is known to produce more accurate results [33]. The impact of these different procedures on the resulting networks is still unknown.

All these different possibilities may be responsible for some of the differences found, in the metrics that characterize the networks, between different studies, having a direct impact on the conclusions that can be draw from its results. In this work we will focus on the study of two critical issues: the effect of normalizing the fMRI images and the effect that different node defining strategies (i.e. the use of individual parcellation methods versus the simple superposition of a fixed template) have on the resulting networks.

2 Materials and methods

2.1 Ethics statement

The study was conducted in accordance with the principles expressed in the Declaration of Helsinki and was approved by the Ethics Committee of Hospital de Braga (Portugal).

2.2 Subjects and acquisitions

In the current study, the MRI acquisitions of 59 healthy volunteers were used, from which 4 were discarded due to excess of movement during the acquisition, so that all subjects displayed head motion less than 2 mm in translation or 1 degree in rotation. Thus, a total of 55 subjects from the SWITCHBOX project were used in this study (31 males and 24 females aged between 51 and 82 years old and with mean age of 64.85 ± 8.82 years). The goals and tests of the study were explained to all participants, which provided informed written signed consent.

Two different acquisitions from each subject were used: as structural acquisition, a T1 magnetization prepared rapid gradient echo (MPRAGE) with repetition time (TR) = 2730 ms, echo time (TE) = 3.5 ms, field of view (FoV) = 256 mm x 256 mm, flip angle = 7°, in plane resolution of 1 mm x 1 mm and 1 mm slice thickness; as a resting-state functional acquisition, a T2* echo-planar imaging (EPI) acquisition with 180 volumes, TR = 2000 ms, TE = 30 ms, FoV = 224 mm x 224 mm, flip angle = 90°, in plane resolution of 3.5 mm x 3.5 mm and 4.5 mm slice thickness. All acquisitions were performed
Impact of normalization and segmentation on brain graphs

on a clinically approved Siemens Magnetom Avanto 1.5 T (Siemens Medical Solutions, Elangen, Germany) scanner, in Hospital de Braga, using a 12-channel receive-only Siemens head coil. During the resting-state acquisition all subjects were instructed to remain with their eyes closed and to not think of anything in particular.

2.3 Preprocessing strategies

All fMRI images were preprocessed using BrainCAT [34], and the Matlab (v.2009b, http://www.mathworks.com/) toolbox Conn [35]. fMRI preprocessing steps included the removal of the first 5 volumes, motion correction, slice timing correction, skull stripping, band-pass temporal filtering (0.01 – 0.08 Hz) using BrainCAT, and the removal of the white matter (WM) and cerebrospinal fluid (CSF) confound signals using Conn and the CompCor strategy [36].

While the usual procedure in functional studies is to normalize all images to a standard space (e.g. MNI) [37], this step will, no matter how good the interpolation function is, alter the original data. To evaluate the effects of the normalization, two different strategies were tested: the normalization of all images to MNI standard space with 2 mm isotropic voxel size and the registration of all the support data (both the T1 structural data and the masks defining the nodes) to the native space of the functional acquisition of each subject. These strategies will allow testing if the normalization procedures can directly affect the network properties. Normalization and registration of all images was done using FSL flirt [38]: for each subject from the native functional image to the T1 anatomical image using a rigid body transformation (6 dof) and from the T1 image native space to the MNI standard space using an affine transformation (12 dof). The concatenation and inversion of both transformation matrices allows for the registration from the functional native space to the MNI standard space and vice-versa.

2.4 Graph construction

The construction of the graph encompasses two different stages, the definition of the nodes and the definition of the edges. On the node definition two different strategies were compared: the use of a
Impact of normalization and segmentation on brain graphs

fixed template atlas, which was overlaid into each subject’s acquisition, and the use of individually segmented images for each subject performed applying FreeSurfer’s v5.1 semi-automated workflow [32]. To eliminate the effect of comparing networks with different sizes [18] and different regions [19], derived from the use of different atlases, the MNI 152 brain template was segmented using the same FreeSurfer workflow, and thus creating the fixed template atlas. For both strategies, regions from two different segmentations were combined: 146 regions from the Destrieux cortical atlas [33] and 14 regions from the subcortical segmentation [39], resulting in a total of 160 regions (Table I; Figure 1). These two different node-defining strategies allowed us to test the effect of building networks with different parcellation strategies.

Insert Table I around here;

Insert Figure 1 around here;

Using these masks, the mean signal across the voxels of each region was extracted from each time point of the fMRI acquisition, allowing for the final extraction of the time series. To measure the similarity between time series, establishing a measure of functional connectivity (edges) between the corresponding ROIs, the Pearson correlation coefficient was used. The Pearson coefficient of each possible combination of ROIs was calculated in Conn and used to build a correlation matrix for each subject. The Fisher’s r-to-Z transformation was then used to transform the r-values of the correlations into Z values.

Combining the two different preprocessing and node defining strategies, four different networks were computed for each subject: with the functional images in the MNI standard space (MNI networks) using the fixed template atlas (ATL-MNI) and the individual FreeSurfer segmentation (FS-MNI); in the functional native space (NAT networks) using the fixed template atlas (ATL-NAT) and using the individual FreeSurfer segmentation (FS-NAT).

The workflow and all the possible strategies are represented in Figure 2.

Insert Figure 2 around here;
Impact of normalization and segmentation on brain graphs

2.5 Network analysis and comparison

The comparison of the resulting networks was done at three levels: on the edges level, on node level and on the global network level. Regarding the networks’ edges, differences in the Z transform correlation coefficients were tested for each pair of regions.

On the node level, two of the most commonly reported nodal metrics were used: betweenness centrality and local efficiency. The betweenness centrality, $C_i^B$, measures the fraction of shortest paths that pass through the node $i$ of the network [40],

$$C_i^B = \frac{1}{(n(n-1))} \sum_{s \neq t} \left( \frac{\sigma_{st}(i)}{\sigma_{st}} \right)^2$$

where $\sigma_{st}(i)$ denotes the number of shortest paths between $s$ and $t$ that pass through $i$ and $\sigma_{st}$ denotes the total number of shortest paths between $s$ and $t$. Measures of centralit evaluate the existence of nodes key to the communication of the network, through which a significant number of communication paths go. Metrics of segregation indicate the existence of highly interconnected groups, associated with functional specialization. Local efficiency ($E_{loc}$) is one of such metrics:

$$E_{loc,i} = \frac{1}{N_i} \sum_{i,j,h} a_{ij}a_{ih} d_{ij}(N_i) / k_i(k_i - 1)$$

where $N_i$ is the subgraph of the neighbours of node $i$, $a_{ij}$ is the existence of a connection between $i$ and $j$ and $k_i$ is the degree of the node $i$.

On the network level, three properties were tested: clustering coefficient, average path length and assortativity coefficient. As measure of network segregation, the mean clustering coefficient $C_p$ reflects the tendency to form clusters of interconnected nodes [10,41],

$$C_p = \frac{1}{N} \sum_i C_i, C_i = (\epsilon(G_i))/(d(i)(d(i)-1)/2),$$

where $N$ denotes the number of nodes in the network, $C_i$ denotes the clustering coefficient of the node $i$ in graph $G$, $d(i)$ denotes the degree (or number of edges connected to) of the node $i$, and $\epsilon(G_i)$ denotes the number of edges in $G_i$, the subgraph formed by all the node $i$ neighbours. Functional integration measures reveal how easily each brain region communicates with the rest of the brain and...
Impact of normalization and segmentation on brain graphs

combines specialized information. The average path length was used as a network wise metric of integration, here calculated as the inverse of the network efficiency, \( E_{\text{glob}} \), to allow its calculation in disconnected graphs\([41,42]\),

\[
L = \frac{1}{E_{\text{glob}}} \quad E_{\text{glob}} = \frac{1}{N(N-1)} \sum_{x,y \in G, x \neq y} d_{ij}^{-1}
\]

where \( d_{ij} \) denotes the distance, or the minimum number of connections, between the nodes \( i \) and \( j \).

Measures of network resilience quantify the network ability to resist to direct or random attacks. One possible measure is the assortativity coefficient \( r \), which measures the correlation between a node degree and the degree of its neighbours, in other words, shows how likely it is for a node with a high degree to be connected to nodes with a similar degree\([43,44]\),

\[
r = \left( M^{1-\sum_{j=1}^{M} j k_i - \left[ M^{1-\sum_{j=1}^{M} \frac{1}{2}(j_i + k_i)\right]^2 / \left( M^{1-\sum_{j=1}^{M} \frac{1}{2}(j_i + k_i)\right)^2 \right] \right)\]

\( M \) denotes the number of edges in the graph, \( j_i \) and \( k_i \) denote the degree of the nodes at both ends of the edge \( i \).

Commonly, networks can be classified according to different architectures. One such architecture is the small-world topology. These networks are usually defined as having a higher \( C_p \) than those found on random networks \( (C_p/C_{p \text{ rand}} > 1) \), and a similar \( L \) \( (L/L_{\text{rand}} \approx 1) \)\([10,41]\). This translates into the formation of highly connected sub-networks while keeping a high level of connectivity\([45]\). As so, \( C_p/C_{p \text{ rand}} \) and \( L/L_{\text{rand}} \) were also called small-world parameters and were used to characterize the networks.

The nodal and network metrics for all networks and subjects were calculated along a range of densities, from 5 to 40% in steps of 2.5%. In order to reduce the complexity of the analysis, the nodal properties were condensed across the density range through the use of the integrated version of the measure, calculated as follows\([22]\):

\[
X(i) = \sum_{s=5}^{k=45} X(i; k \Delta d) \Delta d,
\]

where \( X \) is the property of the node \( i \), along 15 density levels, and \( \Delta d \) is the density interval (0.025).
Impact of normalization and segmentation on brain graphs

Only positive correlation coefficients were considered and all networks were binarized prior to calculation of the metrics. All calculations were done using the Brain Connectivity Toolbox [41] and Matlab.

2.6 Statistical analysis

In order to compare the different strategies, a 2x2 repeated measures ANOVA was done using Matlab, on the Z-transformed correlation coefficients, metrics $C^b$, LocEf, $C_p$, L, r and using the registration approach, parcellation strategy and registration*parcellation interactions as within subject contrasts. The necessary assumptions for the analysis were previously verified and met. Differences were considered significant at $p \leq 0.05$ corrected for multiple comparisons using the family-wise error (FWE) procedure.

3 Results

3.1 Network Characterization

The average network matrices for each strategy, thresholded at a density of 7.5%, can be found in Figure 3. In these matrices, similar connection tendencies were observed for all the strategies. From the division in cortical and sub-cortical regions, it is possible to observe a tendency for stronger connections between regions of the same atlas: cortical regions revealed higher connectivity with other cortical regions than with subcortical regions and the same pattern was observed for subcortical regions. Moreover, regions tend to present stronger connections to regions within the same hemisphere than to regions of the contralateral hemisphere. A symmetry effect can also be observed, meaning that, regions showed strong tendencies to connect to the matching region on the contralateral hemisphere, as evidenced by the diagonal crossing both the second and third quadrant of the matrices (Figure 3). From the distribution of Z-transformed correlation coefficients (Figure 4 a)) it is possible to observe a roughly normal distribution, with the majority of values located between 0 and 0.2, while in the distance distribution (Figure 4 b)) it is noticeable that most values are located between 15 and 55 mm.
Impact of normalization and segmentation on brain graphs

272 Insert figure 3 around here.

273 Insert figure 4 around here.

274 Regarding the small-world properties of the networks (Figure 5), all presented high clustering values
275 for the full range of densities, with $1.5 < C_p/C_{p\text{rand}} < 6$, and a tendency for the values to become closer
276 to those of random networks as the network density increased (Figure 5 a)). The characterization of
277 the average path length of each network revealed values close to those of random networks with $1 <
278 L/L_{\text{rand}} < 1.35$ and once again a tendency to grow closer to those of random networks as the density
279 increased (Figure 5 b)). Concerning the assortativity of the networks, all values were found to be
280 positive and $0.24 < r < 0.41$ (Figure 5 c)).

281 Insert Figure 5 around here.

282 3.2 Effect of registration and parcellation strategies

283 For the isolated effect of the registration and parcellation strategies significant differences were found
284 on the three levels of analysis.

285 In the edges comparison, differences were found in the registration test (Figure 6 a)), in a total of 243
286 edges. All these edges, when computed in native space, presented higher values than the same edges
287 computed in standard space. When considering the parcellation test (Figure 6 b)), differences were
288 found in a total of 112 edges, 26 of them having higher value on the ATL networks and 86 on the FS
289 networks.

290 From the distribution of correlations of the edges with significant differences, on the registration
291 comparison (Figure 6 c)) it was possible to observe that the NAT networks Z-transformed correlation
292 values lied between 0.2 and 0.4, while on the MNI networks these were mainly located between -0.1
293 and 0.2. On the parcellation comparison (Figure 6 d)) the values of Z-transformed correlation
294 significant differences from all networks were located mainly between 0.2 and 0.4. From the distance
295 distribution of differences, on the registration comparison (Figure 6 e)), it was possible to observe a
296 tendency for differences to be located on either long or short distance edges with a higher tendency
Impact of normalization and segmentation on brain graphs

for the short range. On the parcellation comparison, differences were mainly located over shorter
connections (Figure 6 f).

On the nodal metrics, differences were found on both metrics for the registration test: one on C^B, with
MNI networks having higher value and 5 on the Eloc with the NAT networks always having higher
values. On the parcellation test, significant differences were found on the C^B metric on 3 vertices with
FS networks always having higher values and no differences were found on the Eloc (Table 2).

On the global metrics level, several differences were found: on the registration test, differences were
found on the clustering metric for values of density between 15 and 40%, with the NAT networks
always having higher values; on the average path length, differences were found on the full range of
densities, with the NAT networks having higher values; and no differences were found on the
assortativity metric (Table 3).

For the parcellation test, differences were found for the clustering coefficient between densities of
22.5 and 40%, on the average path length for density values of 5, 17.5 and 20% and on the
assortativity metric a density of 12.5%. In all instances FS networks showed higher values than ATL
networks (Table 4).

No significant results were found for the registration*parcellation interaction.

4 Discussion
The results herein obtained are in accordance to the expected characteristics and architecture of brain networks, as demonstrated in several studies [18,19,22,46,47]. All networks were shown to be assortative ($r>0$) and efficient ($L/L_{\text{rand}} \approx 1$) in the full range of densities. All of them also showed a strong small world like tendency to constitute clusters ($C_p/C_{p\text{ rand}} > 1$) on the lower range of densities (5% to 20%).

The thresholded correlation matrices demonstrated a clear tendency for regions to connect to other regions within the same anatomical division (cortical/sub-cortical and left/right), evidencing not only a tendency for lateralization of function in the human brain [22], but also reflecting a tendency to have more efficient networks by having stronger connections over smaller anatomical distances [48]. The diagonals over the L/R quadrants reflect the natural correlation and sharing of functions between symmetric regions [49,50].

For the isolated effect of the registration strategies several significant differences were found in all three levels of the analysis. From the differences matrix and the corresponding correlation and distance distribution it is possible to observe a strong tendency for the networks built in the functional native space to have higher values of correlation and for these differences to be mainly located on either shorter or longer range connections. These differences translate in the network metrics into a higher clustering coefficient and longer average path distance in the native space networks, with a very large effect size [51]. Overall, this seems to point that networks built in this space favor the detection of stronger local paths of communication, not restrained to physically proximal, or inter-hemispheric, regions, in detriment of more evenly spread connections. On the nodal metrics, it could be expected that this would translate into a higher tendency for the native space networks to have vertices with a stronger centrality role. Instead we found differences in only one vertex in B⁴, favoring the MNI space networks (although with a small effect size). In the local efficiency, the differences favor again the native space networks, with the existence of nodes with a more important role on facilitating the communication on their local networks, which should be associated with the higher clustering coefficient, but again, with a small effect size. The differences found can be the result of multiple factors: the resampling and interpolation of the functional data in the normalization step, the
resampling of the parcellation images and loss of detail, or a mixture of both. As no interaction was found between the registration and parcellation strategies, the first option emerges as the most likely.

In the parcellation strategies comparison several differences were also found. At the correlation level it is possible to observe differences in both directions, mostly with higher values in the individual segmentation networks. From the correlation distribution it is possible to observe a small tendency for the individual segmentation networks to have higher values of correlation, and from the distance distribution for the differences to be located mainly over shorter-range edges and within hemispheres. We found higher values of $C_p$ in the individual segmentation strategies, with large to very large effect sizes, $L$, with high effect sizes, $C_B$, with small to very small effect sizes and $r$ with large effect sizes. These differences can be explained by the higher precision of the strategy in segmenting the different areas according to sulci and gyri [33], possibly establishing higher and more accurate correlations. This strategy seems to favor the detection of strong and short connections associated with the formation of local clusters (as seen in the increased $C_p$). The lack of ability of the fixed template strategy networks to detect these strong correlations favors the survival on the threshold procedure of the weaker, long distance connections, establishing extra paths of communication between distant points in the network, reducing the average shortest path in these networks. Furthermore on the nodal metrics comparison, three individual segmentation network vertices were found to have higher value of $B^C$. As the main hubs of communication in a network are usually associated with the necessity to establish bridges of communications between strong and independent clusters, these seem a logic finding, supported by the large effect size.

This study presents some limitations that should be noted. While we reveal the impact of the normalization step in several aspects of the network, it is, in this study, impossible to show which inner step of the normalization is the exact responsible for these differences. Further exploration of the effect of resampling the functional data and the use of different interpolation functions are needed to characterize the exact effect of this step. It is also important to notice that several of the differences found in the network metrics were present on higher values of density. This can be explained by observing the distributions of correlations of differences, as most edges with significant differences
Impact of normalization and segmentation on brain graphs

have medium to low levels of correlation, only once these survive the thresholding procedure their
effect will be noticeable on the network. There is a tendency in neuroimaging studies to focus on
lower density networks, possibly reducing the relevance of some of the present findings. Yet, there is
no direct evidence of which density threshold is more appropriate, which could be critical for the
overall interpretation of the present finding and of complex brain networks studies.

Overall, we have shown a strong impact of two different pre-processing steps on the resulting
networks. It is, without a prior-knowledge of the expected network properties, impossible to
determine which strategy better represents the underlying biological brain network, making it only
possible to theoretically argue in favor of those strategies that have either less impact on the data, and
are more precise in defining the regions of interest. While further exploration of the normalization
step is needed, the present results favor the use of the more conservative approach, avoiding the use of
resampling and interpolation functions on the data, whenever possible. Similar to this, the use of a
more precise segmentation method for defining the ROIs also seems to have a moderate impact on the
networks obtained.

Acknowledgments: We are thankful to all study participants.

Financial disclosures: RM is supported by a fellowship of the project FCT-ANR/NEU-
OSD/0258/2012 founded by FCT/MEC (www.fct.pt) and by Fundo Europeu de Desenvolvimento
Regional (FEDER). PM was supported by a fellowship of the project SwitchBox-FP7-HEALTH-
2010-grant 259772-2 (www.switchbox-online.eu). The funders had no role in study design, data
collection and analysis, decision to publish, or preparation of the manuscript.

References

   Nonlinear Biomed Phys 1: 3.
   341-350.
4. Ogawa S, Lee TM, Kay AR, Tank DW (1990) Brain magnetic resonance imaging with contrast
dependent on blood oxygenation. Proc Natl Acad Sci U S A 87: 9868-9872.
Impact of normalization and segmentation on brain graphs


Impact of normalization and segmentation on brain graphs

Impact of normalization and segmentation on brain graphs

**Legends and titles:**

**Figure 1.** Axial view of the ROIs defined for 1 subject after preprocessing: a) in native functional space; b) in MNI space.

**Figure 2.** Complete image processing workflow: from the raw images to the correlation matrices. In blue, the workflow of functional images and, in red, the workflow of the node defining atlases.

**Figure 3.** Correlation matrices for the four networks thresholded at 7.5% density averaged for all subjects: networks built using the pre-segmented atlas (ATL), the individual segmentation (FS); built on the standard MNI space and on the functional native space (NAT). White lines divide the matrices in sub-cortical (Sub Cort) and cortical (Cort) regions as well as left and right.
Impact of normalization and segmentation on brain graphs

Figure 3. Correlation matrices for the four networks thresholded at 7.5% density averaged for all subjects: networks built using the pre-segmented atlas (ATL), the individual segmentation (FS); built on the standard MNI space and on the functional native space (NAT). White lines divide the matrices in sub-cortical (Sub Cort) and cortical (Cort) regions as well as left and right.
Impact of normalization and segmentation on brain graphs

Figure 4. Networks distributions: a) distribution of Z-transformed correlation values; b) distribution of distances;

Figure 5. Global metrics network characterization. $C_p/C_{prand}$: network average clustering coefficient compared to random networks; $L/L_{rand}$: network average path length compared to random networks; $r$: network assortativity; For the networks built using the pre-segmented atlas (ATL), the individual segmentation (FS); built on the standard MNI space (MNI) and on the functional native space (NAT);
**Figure 6:** Edge level significant differences for the registration and parcellation comparisons for $p < 0.05$ corrected for multiple comparisons with FWE procedure: a) and b) location and direction of differences; c)
Impact of normalization and segmentation on brain graphs

and d) distribution of the Z-transformed correlation values for the edges with significant differences for each network; e) and f) distribution of physical distances for the edges with significant differences.

Table 1- Regions of Interest used from the Destrieux and Sub-Cortical atlas.

<table>
<thead>
<tr>
<th>Designation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>lh/rh-Thalamus-Proper</td>
<td>ctx_lh/rh_G_temp_sup-Lateral</td>
</tr>
<tr>
<td>lh/rh-Caudate</td>
<td>ctx_lh/rh_G_temp_sup-Plan_polar</td>
</tr>
<tr>
<td>lh/rh-Putamen</td>
<td>ctx_lh/rh_G_temp_sup-Plan_tempo</td>
</tr>
<tr>
<td>lh/rh-Pallidum</td>
<td>ctx_lh/rh_G_temporal_inf</td>
</tr>
<tr>
<td>lh/rh-Hippocampus</td>
<td>ctx_lh/rh_G_temporal_middle</td>
</tr>
<tr>
<td>lh/rh-Amygdala</td>
<td>ctx_lh/rh_Lat_Fis-ant-Horizont</td>
</tr>
<tr>
<td>lh/rh-Accumbens-area</td>
<td>ctx_lh/rh_Lat_Fis-ant-Vertical</td>
</tr>
<tr>
<td>ctx_lh/rh_G_and_S_frontomargin</td>
<td>ctx_lh/rh_Lat_Fis-post</td>
</tr>
<tr>
<td>ctx_lh/rh_G_and_S_occipital_inf</td>
<td>ctx_lh/rh_Pole_occipital</td>
</tr>
<tr>
<td>ctx_lh/rh_G_and_S_paracentral</td>
<td>ctx_lh/rh_Pole_temporal</td>
</tr>
<tr>
<td>ctx_lh/rh_G_and_S_subcentral</td>
<td>ctx_lh/rh_S_calcarine</td>
</tr>
<tr>
<td>ctx_lh/rh_G_and_S_transv_frontopol</td>
<td>ctx_lh/rh_S_central</td>
</tr>
<tr>
<td>ctx_lh/rh_G_and_S_cingul-Ant</td>
<td>ctx_lh/rh_S_cingul-Marginalis</td>
</tr>
<tr>
<td>ctx_lh/rh_G_and_S_cingul-Mid-Ant</td>
<td>ctx_lh/rh_S_circular_insula_ant</td>
</tr>
<tr>
<td>ctx_lh/rh_G_and_S_cingul-Mid-Post</td>
<td>ctx_lh/rh_S_circular_insula_inf</td>
</tr>
<tr>
<td>ctx_lh/rh_G_cingul-Post-dorsal</td>
<td>ctx_lh/rh_S_circular_insula_sup</td>
</tr>
<tr>
<td>ctx_lh/rh_G_cingul-Post-ventral</td>
<td>ctx_lh/rh_S_collat_transv_ant</td>
</tr>
<tr>
<td>ctx_lh/rh_G_cuneus</td>
<td>ctx_lh/rh_S_collat_transv_post</td>
</tr>
<tr>
<td>ctx_lh/rh_G_front_inf-Opercular</td>
<td>ctx_lh/rh_S_front_inf</td>
</tr>
<tr>
<td>ctx_lh/rh_G_front_inf-Orbital</td>
<td>ctx_lh/rh_S_front_middle</td>
</tr>
<tr>
<td>ctx_lh/rh_G_front_inf-Triangul</td>
<td>ctx_lh/rh_S_front_sup</td>
</tr>
<tr>
<td>ctx_lh/rh_G_front_middle</td>
<td>ctx_lh/rh_S_interm_prim-Jensen</td>
</tr>
<tr>
<td>ctx_lh/rh_G_front_sup</td>
<td>ctx_lh/rh_S_intrapariet_and_P_trans</td>
</tr>
<tr>
<td>ctx_lh/rh_G_ins_lg_and_S_cent_ins</td>
<td>ctx_lh/rh_S_oc_middle_and_Lunatus</td>
</tr>
<tr>
<td>ctx_lh/rh_G_insular_short</td>
<td>ctx_lh/rh_S_oc_sup_and_transversal</td>
</tr>
<tr>
<td>ctx_lh/rh_G_occipital_middle</td>
<td>ctx_lh/rh_S_occipital_ant</td>
</tr>
<tr>
<td>ctx_lh/rh_G_occipital_sup</td>
<td>ctx_lh/rh_S_oc-temp_lat</td>
</tr>
<tr>
<td>ctx_lh/rh_G_oc-temp_lat-fusifor</td>
<td>ctx_lh/rh_S_oc-temp_med_and_Lingual</td>
</tr>
<tr>
<td>ctx_lh/rh_G_oc-temp_med-Lingual</td>
<td>ctx_lh/rh_S_orbital_lateral</td>
</tr>
</tbody>
</table>
Impact of normalization and segmentation on brain graphs

<table>
<thead>
<tr>
<th>Networks</th>
<th>Registration</th>
<th>Parcellation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATL-MNI</td>
<td>ATL-NAT</td>
<td>FS-MNI</td>
</tr>
<tr>
<td>Node</td>
<td>C²p</td>
<td></td>
</tr>
<tr>
<td>Left-Pole_temporal</td>
<td>59.554</td>
<td>62.079</td>
</tr>
<tr>
<td>Right-G_and_S_occipital_inf</td>
<td>53.345</td>
<td>55.552</td>
</tr>
<tr>
<td>Right-Lat_Fis-ant-Horizont</td>
<td>60.639</td>
<td>33.838</td>
</tr>
<tr>
<td>Right-Pole_temporal</td>
<td>72.203</td>
<td>67.579</td>
</tr>
<tr>
<td>Eloc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left-S_postcentral</td>
<td>0.272</td>
<td>0.287</td>
</tr>
<tr>
<td>Right-G_and_S_cingul-Mid-Post</td>
<td>0.257</td>
<td>0.280</td>
</tr>
<tr>
<td>Left-G_postcentral</td>
<td>0.272</td>
<td>0.285</td>
</tr>
<tr>
<td>RightS_postcentral</td>
<td>0.272</td>
<td>0.290</td>
</tr>
<tr>
<td>Right-S_precentral-sup-part</td>
<td>0.255</td>
<td>0.276</td>
</tr>
</tbody>
</table>

Table 2 - Significant differences in the nodal metrics in bold for the registration and parcellation comparisons, with p<0.05 corrected for multiple comparisons with the FWE procedure.

Table 3 - Isolated effect of the registration strategies for the graph metrics across all threshold values;
Significant results in bold corrected for multiple comparisons with the FWE procedure;
Impact of normalization and segmentation on brain graphs

<table>
<thead>
<tr>
<th>th(%)</th>
<th>p</th>
<th>F_{(1,54)}</th>
<th>\eta^2</th>
<th>p</th>
<th>F_{(1,54)}</th>
<th>\eta^2</th>
<th>p</th>
<th>F_{(1,54)}</th>
<th>\eta^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.58</td>
<td>0.31</td>
<td>0.006</td>
<td>&lt;.001</td>
<td>15.107</td>
<td>0.222</td>
<td>0.005</td>
<td>8.721</td>
<td>0.141</td>
</tr>
<tr>
<td>7.5</td>
<td>0.015</td>
<td>6.29</td>
<td>0.106</td>
<td>0.002</td>
<td>11.028</td>
<td>0.172</td>
<td>0.006</td>
<td>8.373</td>
<td>0.136</td>
</tr>
<tr>
<td>10</td>
<td>0.008</td>
<td>7.604</td>
<td>0.125</td>
<td>0.002</td>
<td>10.404</td>
<td>0.164</td>
<td>0.001</td>
<td>13.110</td>
<td>0.198</td>
</tr>
<tr>
<td>12.5</td>
<td>0.165</td>
<td>1.984</td>
<td>0.036</td>
<td>0.002</td>
<td>10.131</td>
<td>0.160</td>
<td>0.000</td>
<td>15.192</td>
<td>0.223</td>
</tr>
<tr>
<td>15</td>
<td>0.067</td>
<td>3.485</td>
<td>0.062</td>
<td>0.003</td>
<td>9.614</td>
<td>0.154</td>
<td>0.008</td>
<td>7.639</td>
<td>0.126</td>
</tr>
<tr>
<td>17.5</td>
<td>0.016</td>
<td>6.159</td>
<td>0.104</td>
<td>&lt;.001</td>
<td>19.658</td>
<td>0.271</td>
<td>0.003</td>
<td>9.782</td>
<td>0.156</td>
</tr>
<tr>
<td>20</td>
<td>0.038</td>
<td>4.514</td>
<td>0.078</td>
<td>&lt;.001</td>
<td>14.613</td>
<td>0.216</td>
<td>0.040</td>
<td>4.440</td>
<td>0.077</td>
</tr>
<tr>
<td>22.5</td>
<td>&lt;.001</td>
<td>13.808</td>
<td>0.207</td>
<td>0.002</td>
<td>10.568</td>
<td>0.166</td>
<td>0.166</td>
<td>1.969</td>
<td>0.036</td>
</tr>
<tr>
<td>25</td>
<td>&lt;.001</td>
<td>19.346</td>
<td>0.267</td>
<td>0.002</td>
<td>10.783</td>
<td>0.169</td>
<td>0.403</td>
<td>0.710</td>
<td>0.013</td>
</tr>
<tr>
<td>27.5</td>
<td>&lt;.001</td>
<td>21.622</td>
<td>0.290</td>
<td>0.005</td>
<td>8.533</td>
<td>0.139</td>
<td>0.498</td>
<td>0.467</td>
<td>0.009</td>
</tr>
<tr>
<td>30</td>
<td>&lt;.001</td>
<td>28.434</td>
<td>0.349</td>
<td>0.009</td>
<td>7.348</td>
<td>0.122</td>
<td>0.482</td>
<td>0.502</td>
<td>0.009</td>
</tr>
<tr>
<td>32.5</td>
<td>&lt;.001</td>
<td>29.846</td>
<td>0.360</td>
<td>0.016</td>
<td>6.228</td>
<td>0.105</td>
<td>0.656</td>
<td>0.200</td>
<td>0.004</td>
</tr>
<tr>
<td>35</td>
<td>&lt;.001</td>
<td>29.312</td>
<td>0.356</td>
<td>0.045</td>
<td>4.202</td>
<td>0.073</td>
<td>0.592</td>
<td>0.290</td>
<td>0.005</td>
</tr>
<tr>
<td>37.5</td>
<td>&lt;.001</td>
<td>25.243</td>
<td>0.323</td>
<td>0.181</td>
<td>1.839</td>
<td>0.034</td>
<td>0.694</td>
<td>0.157</td>
<td>0.003</td>
</tr>
<tr>
<td>40</td>
<td>&lt;.001</td>
<td>25.713</td>
<td>0.327</td>
<td>0.119</td>
<td>2.505</td>
<td>0.045</td>
<td>0.781</td>
<td>0.078</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Table 4- Isolated effect of the parcellation strategies for the graph metrics across all threshold values; Significant results in bold corrected for multiple comparisons with the FWE procedure;
Impact of normalization and segmentation on brain graphs