

Article

Not peer-reviewed version

Intra-Atlas Node Size Effects on Graph Metrics in fMRI Data: Implications for Alzheimer's Disease and Cognitive Impairment

Sahithi Kolla , [Haleh Falakshahi](#) ^{*} , Anees Abrol , Zening Fu , [Vince D. Calhoun](#) ^{*}

Posted Date: 2 October 2023

doi: 10.20944/preprints202310.0041.v1

Keywords: functional network connectivity; fMRI; graph metrics; node size; brain networks



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Intra-Atlas Node Size Effects on Graph Metrics in fMRI Data: Implications for Alzheimer's Disease and Cognitive Impairment

Sahithi Kolla ¹, Haleh Falakshahi ^{2,*}, Anees Abrol ³, Zening Fu ⁴ and Vince D. Calhoun ⁵

¹ Tri-Institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS): {Georgia State University, Georgia Institute of Technology, and Emory University}, Atlanta, GA, USA; Department of Computer Science, Georgia State University, GA, USA; skolla@student.gsu.edu

² Tri-Institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS): {Georgia State University, Georgia Institute of Technology, and Emory University}, Atlanta, GA, USA; School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA; hfalakshahi@gatech.edu

³ Tri-Institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS): {Georgia State University, Georgia Institute of Technology, and Emory University}, Atlanta, GA, USA; aabrol@gsu.edu

⁴ Tri-Institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS): {Georgia State University, Georgia Institute of Technology, and Emory University}, Atlanta, GA, USA; zfu@gsu.edu

⁵ Tri-Institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS): {Georgia State University, Georgia Institute of Technology, and Emory University}, Atlanta, GA, USA; Department of Computer Science, Georgia State University, GA, USA; vcalhoun@gsu.edu

* Correspondence: hfalakshahi@gatech.edu

Abstract: Network neuroscience delves into comprehending the intricate arrangement and interactions within the neural networks of the brain. By merging insights from both neuroscience and network theory, it investigates the brain's structure and function through the lens of interconnected nodes (symbolizing brain regions) and edges (symbolizing the connections among these regions) forming a dynamic network. Graph metrics play a pivotal role in neuroscience by providing quantifiable insights into the intricate connectivity patterns within the brain's neural networks. Frequently, nodes are extracted from functional or structural atlases, which can lead to diverse shapes and sizes. Nevertheless, our understanding of how these differing node characteristics and definitions impact the computed graph metrics in neuroimaging data remains limited. This study adopts a data-driven methodology to delineate functional nodes, subsequently examining the influence of their sizes on resultant graph metrics. By employing the Neuromark framework, an entirely automated independent component analysis (ICA) is applied to resting state fMRI data. Through this, functional network connectivity (FNC) matrices are computed—capturing Pearson correlations between component-time courses—while employing a proportional threshold to streamline node connections based on these correlations. Subsequently, various global and local graph metrics are computed, offering insights into network characteristics. Global metrics provide an overarching summary of the network structure, while local measures delve into structural intricacies at the individual component level. Node sizes are computed based on voxel counts surpassing a designated threshold. Next, we calculated the Pearson correlation between the obtained node sizes and the graph metrics, which we term 'node-metric coupling' (NMC). Our findings revealed consistent and noteworthy correlations between the values of graph metrics and the dimensions of brain nodes. To delve deeper into the implications of this correlation, we examined the node-metric coupling within a dataset comprising Alzheimer's disease, mild cognitive impairment, and control subjects. The disparities observed in node-metric coupling among these groups underscore the necessity of accounting for this factor during analysis. The two principal outcomes of this study are as follows: Firstly, the substantial interplay between varying node sizes within a given atlas and the resultant graph metrics; and secondly, the potential utility of node-metric coupling as a viable biomarker for brain disorders. These discoveries hold significant ramifications. While comparing studies employing diverse atlases is already a challenge, our work highlights an additional, critical source of variability: accounting for node sizes. Consequently, the association between intra-atlas node size and graph metric should be thoughtfully acknowledged in future neuroimaging investigations.

Keywords: functional network connectivity; fMRI; graph metrics; node size; brain networks

1. Introduction

The human brain, a marvel of biological complexity, intricately weaves together a multitude of interconnected regions. For decades, researchers have been captivated by the profound challenge of unraveling the enigmatic workings of the mind. This pursuit has led us to the realm of brain networks, including regions as nodes within a graph, with their interconnections depicted as edges [1], [2]. Leveraging graph theory, a mathematical framework originally designed for analyzing networks and relationships, has proven invaluable in deciphering the brain's mysteries. Notably, it has emerged as an indispensable tool for decoding neuroimaging data, ushering in a new era of understanding [3], [1], [4], [5].

As we venture deeper into the study of brain function, a pivotal question arises: what constitutes a 'node' within these intricate brain networks? Some researchers construct intricate graphs using tiny brain units known as 'voxels,' a method that offers its own set of advantages. Conversely, many others opt for an alternative approach known as the atlas-based method [6], [7] In this method, atlases predominantly partition voxels into regions of interest (ROIs).

A large number of brain atlases are currently being utilized. Most of these have inherent variability in the size of the delineated ROIs within a given atlas. That is, across widely utilized atlases, the size of these regions can exhibit substantial disparities. For instance, the ratio of the largest ROI volume to the smallest ROI volume has been observed to vary from a modest 3.8 [8] to an astonishing 45 [7], or even an unprecedented 443 [6]. This striking heterogeneity in ROI size among different atlases underscores a dimension of the atlas-based neuroimaging pipeline that has received limited scrutiny.

While substantial research efforts have been dedicated to developing and validating various brain atlases, a noticeable gap remains in our understanding of how the size of ROIs/nodes when employed in graph-based analyses, impacts the computed graph metrics. Graph theory and network science heavily rely on the precise delineation and characterization of these nodes, providing a mathematical foundation for modeling the intricate interplay of brain regions and extracting meaningful insights from functional and structural connectivity data. Nevertheless, the influence of varying node sizes within the context of graph theory remains predominantly uncharted territory.

This paper endeavors to bridge this gap by conducting a systematic exploration of the influence of node size on computed graph metrics. Through our research, we aim to shed light on whether and to what extent the size of nodes affects the outcomes of graph-based analyses. By addressing this important but often overlooked aspect of brain network analysis, we seek to contribute to the refinement of neuroimaging studies, ensuring their accuracy and interpretability.

We investigate the relationship between intra-atlas node sizes and computed graph metrics. To ensure the temporal coherence of the included timecourses, we adopt a fully automated data-driven functional atlas based on independent component analysis (ICA). Subsequently, we construct a graph by computing cross-correlations among the ICA timecourses, a method known as functional network connectivity (FNC). The determination of strong node connections involves the application of a 'proportional thresholding' (PT) technique [9], wherein connections are retained based on a predefined PT%. To quantify the size of each component node, we perform voxel counting within each component and subject, considering only those voxels with values exceeding a specified threshold.

We proceed to assess the structure of the graph using a combination of global and local metrics. Firstly, we determine the degree of each node, representing the count of connections linked to that node. Additionally, we calculate the characteristic path length, which provides the average shortest path length between all pairs of nodes, as well as identifies densely interconnected clusters of brain regions. These clusters play a vital role in revealing the functional specialization within the network, showcasing areas where neural processing is distinctly segregated. These path lengths and cluster

formations are considered global metrics. Moreover, we identify specific regions within the brain known as 'hubs' for each node in the network. Hubs are central nodes that are more likely to participate in numerous shortest paths within the network. To quantify this centrality, we employ the closeness centrality metric, which measures a node's connectivity to all other nodes in the network. These node-specific graph metrics allow us to gain insights into the inner structural characteristics of individual brain regions.

To explore the relationship between the size of nodes in the graph network and the corresponding values of graph metrics, we employ statistical analysis. Specifically, we compute the Pearson correlation between these two variables. The outcome of this correlation analysis is referred to as 'node-metric coupling' (NMC). Additionally, we investigate whether variations among different subjects or groups might influence the NMC. To explore this aspect, we extend our analysis to include a dataset consisting of resting fMRI data collected from individuals diagnosed with Alzheimer's disease and mild cognitive impairment.

The remainder of the paper is organized as follows: In Section II, we delve into the intricacies of our methodology. Section III offers a comprehensive presentation of our findings derived from the analysis of resting-state fMRI data. Moving on to Section IV, we conduct an in-depth review of our results and discuss their implications. Finally, in Section V, we present our concluding remarks.

2. Materials and Methods

In this section, we begin by providing an overview of the dataset employed, followed by an exposition of our node estimation methodology. Subsequently, we detail our approach for establishing connections between graph metrics and node size.

2.1. Data Information

In this study, we leveraged functional magnetic resonance imaging (fMRI) data obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset. ADNI focuses on the investigation of Alzheimer's Disease (AD) and its progression. The fMRI dataset utilized herein encompasses a comprehensive cohort consisting of 213 subjects diagnosed with Alzheimer's disease, 368 with mild cognitive impairment, and 947 control subjects. Detailed subject characteristics for each group are provided in Table 1.

Table 1. Overview of Analyzed Participant Demographics.

Group Categories	Number of Subjects	M/F count		Average Subject Age
		Male	Female	
Healthy Controls (HC)	947	384	563	74
Mild Cognitive Impairment (MCI)	368	199	169	73
Alzheimer's Disease (AD)	213	115	98	75

2.2. Data Information

All participants underwent imaging using 3.0 Tesla scanners and underwent preprocessing via an SPM 12 pipeline. The preprocessing steps included rigid body motion correction to address subject head motion, correction for slice timing, normalization to the standard MNI space using the EPI template, resampling to achieve 3x3x3 mm isotropic voxels, and applying Gaussian smoothing with a kernel of FWHM = 6x6x6 mm.

2.3. Data-Driven Node Estimation

We employed independent component analysis (ICA) to identify the nodes of a graph, which involves decomposing whole-brain fMRI data into intrinsic connectivity networks (ICNs). Each ICN represents a functional unit comprising voxels that exhibit strong correlations in spontaneous BOLD signal fluctuations. However, when using fully-blind ICA to extract ICNs, different components can be identified across datasets, leading to challenges in replication and comparison. To address this

issue, we utilized the automated Neuromark pipeline [10]. This approach applies spatially constrained ICA guided by reliable network templates. To create the Neuromark templates and identify robust ICNs, we performed spatial ICA on two large datasets (the Human Connectome Project [HCP] and Genomics Superstruct Project [GSP]) of typical controls, resulting in group-level components. We then matched independent components from the HCP and GSP datasets based on spatial similarity (correlation ≥ 0.4) and selected reproducible ICNs as network templates. Five neuroscience experts performed ICN identification and functional domain labeling. Further details on the Neuromark templates can be found in [10]. Subsequently, we used the network templates as prior information within a spatially constrained ICA algorithm to obtain spatial maps and time courses (TCs) for the fBIRN dataset. For this purpose, we applied the multivariate objective optimization independent component analysis with reference (MOO-ICAR) method, implemented in the GIFT toolbox, specifically, the Neuromark_fMRI_1.0 template, from TReNDS (<http://trendscenter.org/software/gift>). Our choice of MOO-ICAR was based on its demonstrated performance in previous studies [11], [12].

2.4. Estimation of Edge and Local/Global Metrics

To elucidate the functional connectivity within the brain network previously computed, we executed the following procedures: (1) Calculating a functional network connectivity (FNC) matrix, and (2) Quantifying the FNC network through the calculation of graph metrics. FNC is derived by computing Pearson cross-correlation coefficients among the ICA time courses. This process yields a matrix from which we calculate graph metrics. Next, we concentrated our analysis on a selection of commonly employed local graph metrics (node degree, participation coefficient, and closeness centrality) as well as global graph metrics (global efficiency, characteristic path length, and clustering coefficient) within the graph.

Local graph metrics are instrumental in gaining insights into the attributes and roles of individual nodes, as well as the nature of their connections within the network. These metrics provide a detailed understanding of how each node contributes to the network. The degree of a node in a network refers to the number of edges (connections) that are directly linked to that node. In other words, it quantifies how many neighbors or connections a node has in the network. A node with a high degree is considered more central or influential within the network. The participation coefficient assesses the extent of intermodular connections of individual nodes within a network. Nodes with low participation coefficients, often referred to as provincial hubs, tend to be primarily connected within their own module, indicating modular segregation. In contrast, nodes with high participation coefficients, known as connector hubs, play a crucial role in facilitating intermodular integration by connecting multiple modules within the network. Closeness centrality is a network metric used to quantify how close or central a specific node is to all other nodes in a network. It is calculated based on the shortest path distances between the node of interest and all other nodes in the network. In essence, closeness centrality measures how quickly a node can interact with or reach all other nodes in the network. A node with a high closeness centrality is considered more central and influential because it can efficiently communicate with other nodes, serving as a bridge for the flow of information. Mathematically, closeness centrality is often expressed as the reciprocal of the sum of the shortest path distances from the node to all other nodes in the network. Nodes with higher closeness centrality values have shorter average distances to other nodes, indicating their greater proximity and connectivity within the network [13].

Global graph metrics are quantitative measures employed to assess the overall structural properties of a network or graph. They offer valuable insights into the organization, connectivity, and behavior of the network as a whole. In the context of brain network analysis, these metrics play a pivotal role in comprehensively describing various aspects such as network shape, functional integration, and communication (correlation) between different brain regions. Global efficiency is a graph metric used to quantify how efficiently information or signals can be exchanged or transferred throughout an entire network. It provides insight into how well a network facilitates global communication between its nodes.

Mathematically, global efficiency is typically defined as the reciprocal of the average shortest path length between all pairs of nodes in the network. The characteristic path length is a graph metric that quantifies the average distance between all pairs of nodes in a network. It provides insight into how well-connected or compact a network is. Mathematically, the characteristic path length is calculated as the average of the shortest path lengths (the minimum number of edges needed to travel from one node to another) between all possible pairs of nodes in the network. In simpler terms, it represents the average number of steps it takes to move from one node to any other node in the network. The clustering coefficient is a graph metric that quantifies the degree to which nodes in a network tend to form clusters or tightly interconnected groups. It measures the local density of connections within a network, providing insight into the presence of cliques or communities [13].

Table 2 presents a summary of the formulas employed for both local and global graph metrics.

Table 2. Formulas for Local and Global Graph Metrics.

Graph Metrics	Formula
Node degree	$k_i^w = \sum_{j \in N} w_{ij}$ w_{ij} represents the connection weights.
Participation coefficient	$y_i^w = 1 - \sum_{m \in M} \left(\frac{k_i^w(m)}{k_i^w} \right)^2$
Closeness centrality	$(L_i^w)^{-1} = \frac{n-1}{\sum_{j \in N, j \neq i} d_{ij}^w}$
Global efficiency	$E^w = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N, j \neq i} (d_{ij}^w)^{-1}}{n-1}$ d_{ij}^w is the shortest weighted path length between i and j
Characteristic path length	$L^w = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N, j \neq i} d_{ij}^w}{n-1}$
Clustering coefficient	$C^w = \frac{1}{n} \sum_{i \in N} \frac{2t_i^w}{k_i(k_i-1)}$ Where the number of triangles is $t_i^w = \frac{1}{2} \sum_{j, h \in N} (w_{ij}w_{ih}w_{jh})^{1/3}$ and w_{ij} represents the connection weights.

2.5. Quantification of Node Size

We determine node size using the individual component maps generated via the Neuromark pipeline. This process involves thresholding each component image, using a threshold set at 3.5 times the standard deviation of the image data. Subsequently, we tally the number of voxels in the component map that surpasses this threshold. This count is proportional to the volume of highly participating voxels within the specific brain network and serves as the basis for comparing the graph metrics. Notably, the maximum-to-minimum volume ratio for the Neuromark atlas stands at 14.2, slightly exceeding that of the Craddock atlas (3.8) [8], yet notably lower than that of the Glasser atlas (45) [14]

2.6. Node-Metric Coupling

We investigated the association between node size in functional connectivity and brain network metrics (i.e., NMC) by computing the correlation between each graph metric and node size for subjects in our dataset, which included individuals with Alzheimer's disease, mild cognitive impairment, and control participants. Additionally, we examined whether NMC varied across subjects and among different groups. To do so, we assessed the significance of correlations within each group and compared NMC between groups.

3. Results

We chose 53 ICNs for subsequent analysis and grouped them into seven distinct functional domains. These labeled domains encompass auditory, sub-cortical, sensorimotor, visual, cognitive control, cerebellar, and default mode networks, as illustrated in Figure 1.

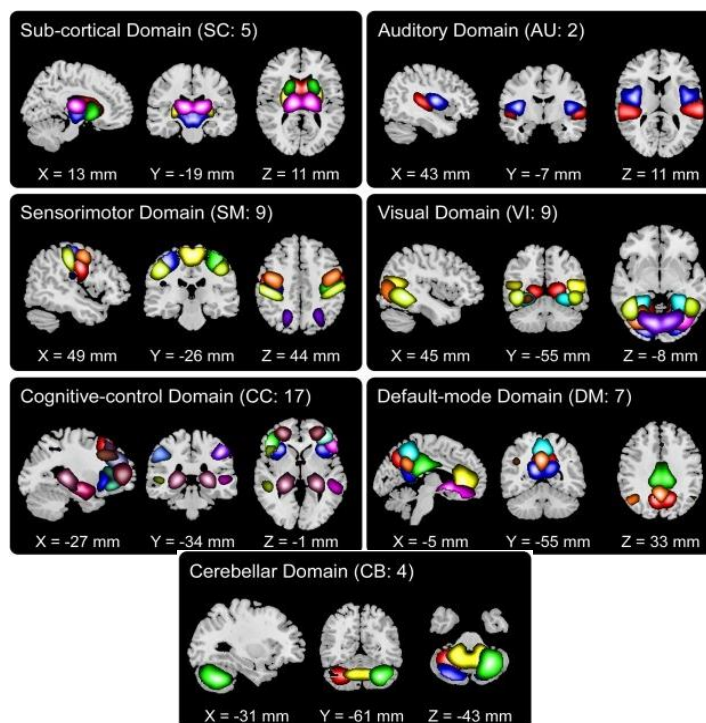


Figure 1. Fifty-three ICNs categorized into seven functional domains.

Figure 2 presents the FNC matrix for each group, organized by functional domain. In the matrix, darker colors signify stronger positive or negative correlations, indicating robust relationships within the network. Conversely, lighter colors denote weaker relationships. Notably, the AD group has slightly, but significantly, lower FNC values than either the control or MCI groups.

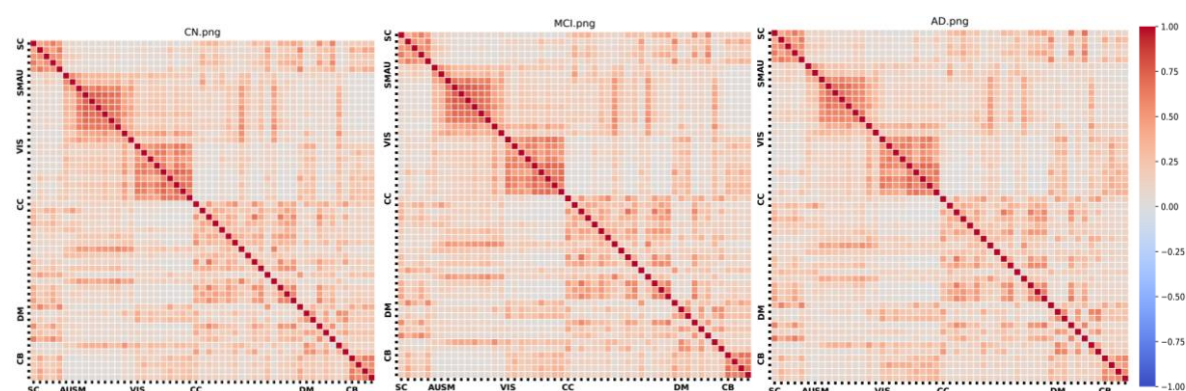


Figure 2. FNC matrices for each group.

Figure 3 illustrates the group-wise differences in cellwise Functional Network Connectivity (FNC) between the AD, HC, and MCI groups. In this depiction, we note that the AD-HC comparison reveals elevated correlation values, indicating stronger connections among brain components, particularly within the subcortical, auditory, and visual domains. Within the cognitive control domain, there is a mixture of higher and lower correlation values. Furthermore, the AD group exhibits slightly weaker correlations, notably in the visual domain.

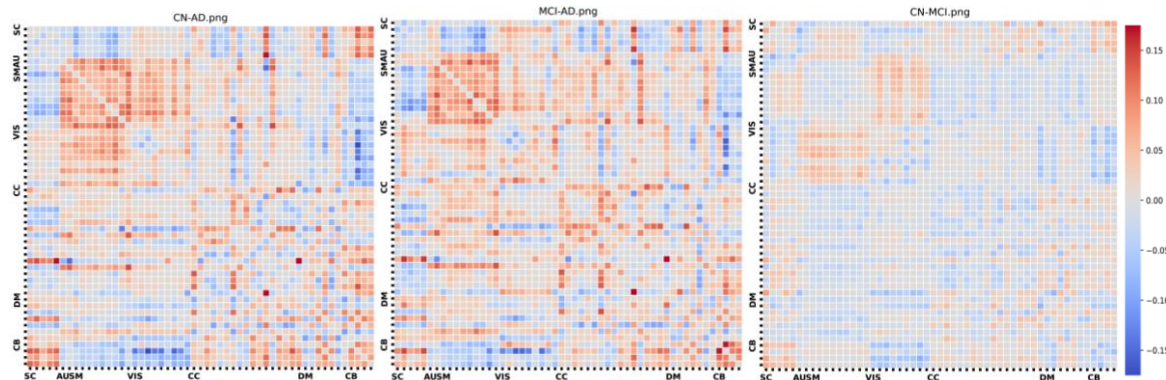


Figure 3. Group differences.

Table 3 displays group-level averages (HC, MCI, AD) for three global graph metric measures: global efficiency, characteristic path length, and clustering coefficient. When contrasting these metrics across different groups, we observe that both MCI and HC exhibit higher values for global efficiency compared to AD. Additionally, we assess brain network density using the clustering coefficient measure, revealing that node distances are lower in HC and AD when compared to MCI, along with similar findings for node density. Lastly, the segregation of nodes, as measured by the clustering coefficient, is lower in AD and HC relative to MCI.

Table 3. Average of calculated global graph metrics.

Graph Metric	HC	MCI	AD
Global Efficiency	0.646	0.639	0.634
Characteristic Path Length	1.823	1.858	1.858
Clustering Coefficient	0.689	0.713	0.693

Table 4 presents group differences in global metrics, including global efficiency, characteristic path length, and clustering coefficient, among HC, MCI, and AD.

Table 4. Global graph metric group differences.

Graph Metric	HC-MCI	HC-AD	MCI-AD
Global Efficiency	4.410	2.667	6.361
Characteristic Path Length	1.505	1.6557	1.0
Clustering Coefficient	2.0458	0.4387	0.0005

Table 5 displays the outcomes of two-sample t-tests conducted to assess group differences in global metrics, including global Efficiency, characteristic path length, and clustering coefficient, within the HC, MCI, and AD groups. All group comparisons yield significant results, even after applying the False Discovery Rate (FDR) correction for multiple comparisons, except for the global efficiency in the HC-MCI comparison. This underscores the significance of evaluating node-metric coupling, as it demonstrates its significance even at the global level.

Table 5. Two-sample t-tests comparing group differences in global graph metrics.

Graph Metric	HC-MCI	HC-AD	MCI-AD
Global Efficiency	7.066(0.07)	*11.64(0.012)	*4.029(0.005)
Characteristic Path Length	*-9.304(0.035)	*-7.986(0.035)	0.0(0.0)
Clustering Coefficient	*-5.668(0.024)	*-0.775(0.004)	*3.459(0.020)

Figure 4 provides a component-level summary of the FNC matrix. We assess the significance of each node using three metrics: node degree, indicating nodes with numerous connections; closeness centrality, representing nodes in proximity to their neighbors; and participation coefficient, which

gauges node activity across the network. Notably, the HC group exhibits a higher level of interconnectedness among nodes in the FNC matrix compared to both AD and MCI.

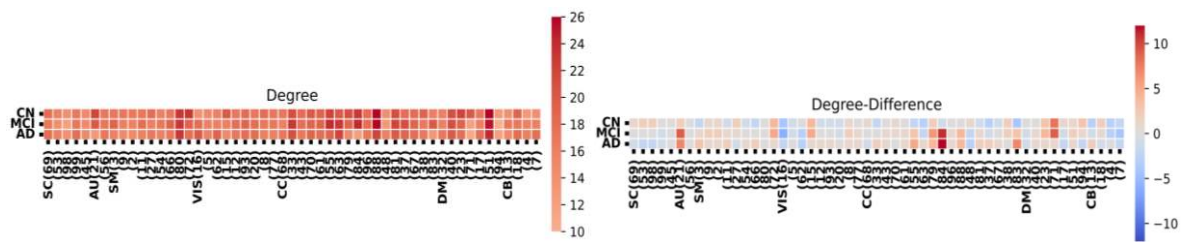


Figure 4. Node degree analysis and network differences among HC, MCI, and AD groups.

Figure 5 provides a component-level summary of the FNC matrix, emphasizing the evaluation of the shortest pathways from one node to all others based on their closeness centrality. Closeness centrality is commonly defined as the normalized inverse of the total topological distances across the graph, often referred to as the nodes' accessibility. Notably, in the FNC matrix, the HC group exhibits a greater level of integral connectivity among most nodes compared to MCI and AD.

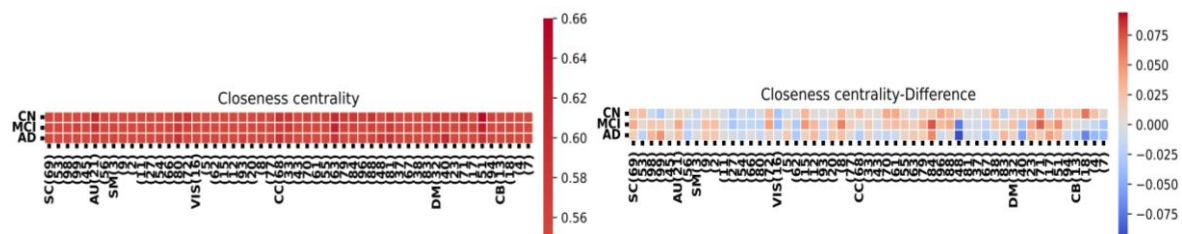


Figure 5. Closeness centrality analysis and network differences among HC, MCI, and AD groups.

This section In Figure 6, we present the results of the participation coefficient analysis in HC, MCI, and AD. The participation coefficient assesses how a node's edges are distributed among the various communities within a graph. A node's participation coefficient is zero if all its edges are exclusively connected to members of its own community. Notably, the figure reveals that after HC, AD exhibits the highest participation coefficient, while MCI shows the least difference in coefficients. We aim to identify the columns of significance among HC, MCI, and AD within the specified brain regions. Our goal is to discern whether AD and MCI exhibit significant values within these regions. Our observations indicate that MCI displays stronger significance within the considered regions compared to AD.

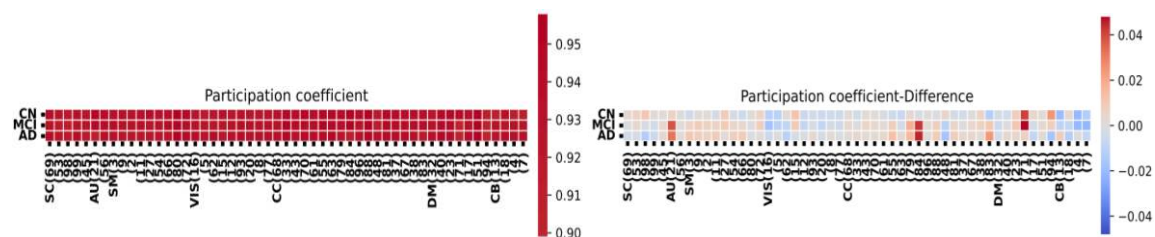


Figure 6. Participation coefficient analysis and network differences among HC, MCI, and AD groups.

This section Figure 7 presents the key findings, emphasizing the relationship between graph metrics and node size within different brain network groups. In Figure 7-a, it's evident that AD exhibits a robust correlation with global graph metrics compared to HC and MCI, while MCI displays a lower correlation with component graph metrics compared to AD and HC. These results underscore the highly significant relationship between node size and resulting graph metrics in each brain network. For instance, the characteristic path length shows a positive correlation in the control and MCI groups (meaning larger node size leads to greater characteristic path length), but a negative

correlation in the AD group (larger node size leads to lower characteristic path length). Regarding global efficiency, the coupling is predominantly negative in all three groups, indicating that larger nodes are associated with lower global efficiency. Other metrics, such as closeness centrality and participation coefficient, exhibit dependency on the specific network. For instance, subcortical networks display a positive node-metric coupling in all three groups, while most visual networks demonstrate a negative node-metric coupling. These differences may represent either a significant source of potential bias or a potential biomarker measure.

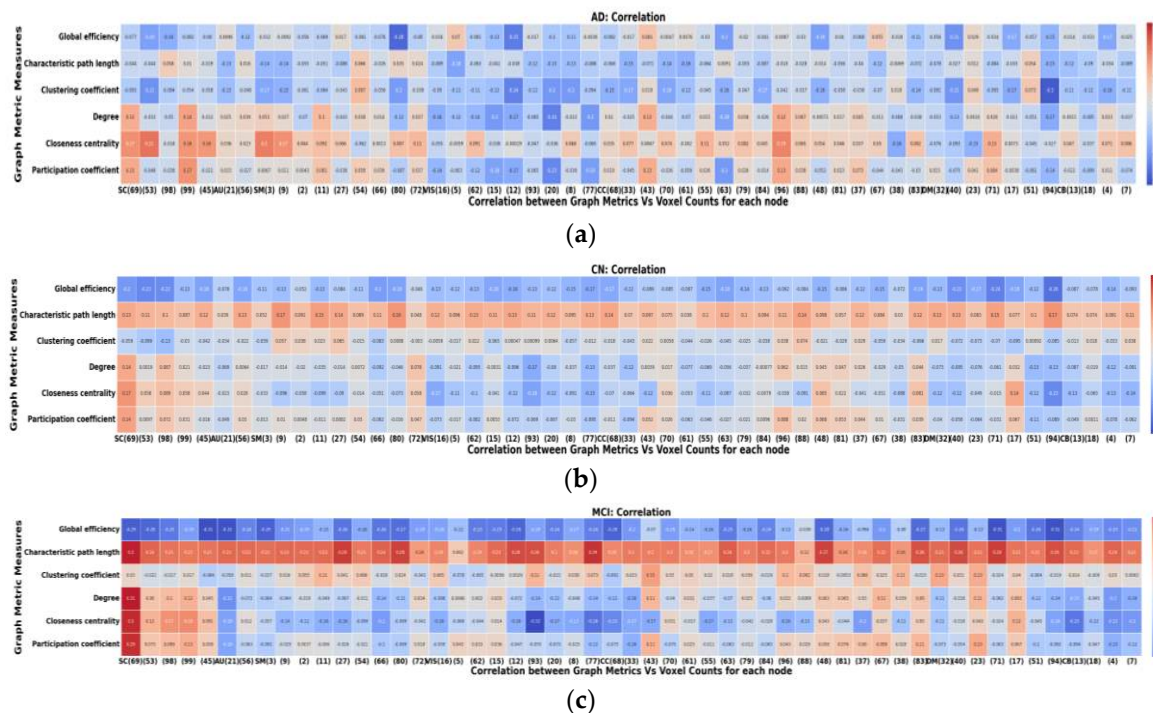
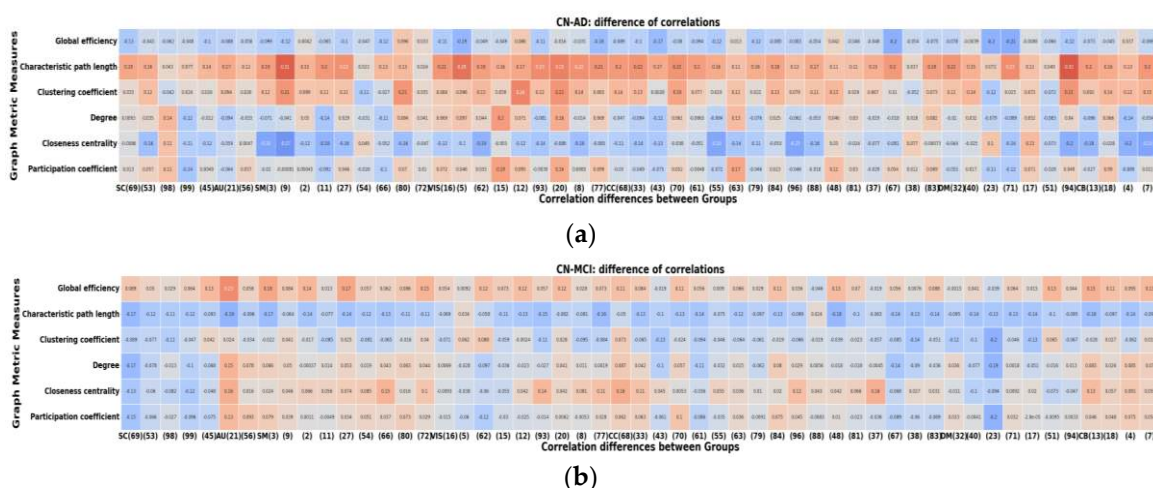


Figure 7. (a) correlation between node size and the global and local metrics in AD group. (b) Correlation between node size and the global and local metrics in HC. (c) Correlation between node size and the global and local metrics in MCI.

Figure 8 illustrates the differences in correlations between different groups. Notably, the AD group exhibits more pronounced differences in their correlations with graph metrics compared to both HC and MCI, whereas the differences between HC and MCI are comparatively smaller. This highlights the significance of the relationship between node size and graph metrics, which is also notably influenced by the specific graph metric under consideration.



(b)

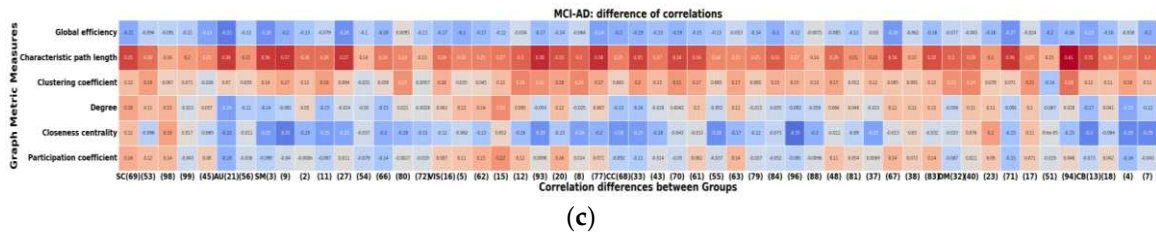


Figure 8. Differences of correlations between the global and local metrics. (a) AD-HC, (b) MCI- HC, (c)AD-MCI.

We assessed the significance of our results using Bonferroni correction for multiple comparisons at a significance level of 0.05, considering a total of 6 graph metrics and 53 components. Table 6 and Table 7 display the correlation values at which significance was achieved, along with the number of networks that exhibited significant correlation for each graph metric.

Table 6. The correlation values at which significance was achieved.

	correlation
HC (significant value)	0.12
AD (significant value)	0.26
MCI (significant value)	0.20
MCI-CN	0.11
AD-CN	0.12
AD-MCI	0.16

Table 7. Number of networks that exhibited significant correlation for each graph metric.

Graph Metric	HC	AD	MCI	HC-MCI	AD-HC	AD-MCI
Global Efficiency	0	3	0	24	4	0
Characteristic Path Length	53	4	53	0	48	53
Clustering Coefficient	5	3	17	5	30	46
Degree	8	10	16	10	9	17
Closeness Centrality	9	26	8	12	3	5
Participation Coefficient	14	8	14	10	15	17

In general, our analysis reveals numerous significant differences in node-metric coupling across all tested graph metrics. Specifically, when examining the effects within the AD group, we observe that closeness centrality exhibits the most pronounced and significant differences. Meanwhile, characteristic path length and clustering coefficient show the most substantial differences when comparing AD with HC and MCI groups. Thus, results underscore the substantial influence of node size on the resulting graph metrics.

4. Discussion

We conducted an initial study to explore the potential relationship between intera-atlas node size and resulting graph metrics. Our findings revealed a robust relationship, which also interacts with the comparisons between patients and controls. These results carry significant implications for the utilization of graph measures in neuroimaging research. Network science holds substantial promise for advancing our understanding of human brain development, pathology, and aging, and such analyses are increasingly prevalent in the field [15]. Nevertheless, despite its widespread adoption, challenges persist in its application to neuroimaging data. One such challenge is the ongoing debate surrounding the definition of a network node in functional brain network analysis. An important caveat in previous studies lies in the absence of an accepted standard for defining brain nodes. The diverse array of approaches employed in this regard can significantly influence the

outcomes. For instance, in the context of atlas-based node definition, which is commonly employed in most studies, researchers have the option to choose between anatomically predefined nodes, functionally informed nodes [16], or data-driven nodes [10], [17]. However, more recent studies have begun to compare multiple atlases [18]. Notably, the published and shared atlases exhibit significant variations in the volumes of included brain regions. To date, there has been a lack of systematic research into whether the volume of brain regions included has an impact on the resulting graph metrics.

In this study, we adopted a data-driven approach to identify functionally defined nodes and investigate how their size influences the resulting metrics. We processed resting fMRI data via a fully automated ICA using the `neuromark_fmri_1.0` template. Subsequently, we applied a proportional threshold to prune node connections based on correlation, leading to the computation of FNC matrices, which represent the correlations among component-time courses. Based on these matrices, we derived various global and local graph metrics. Local metrics capture structural details at the component level, while global metrics provide an overall summary of the network structure. To assess the relationship between node size and graph metrics, we estimated node size by counting the number of voxels above a specified threshold. Larger node sizes corresponded to a higher voxel count.

Employing our node-metric coupling technique, we computed the cross-correlation between the obtained node sizes and graph metrics. To assess the impact of this relationship, we conducted a comparative analysis using a dataset comprising individuals from three categories: controls, individuals with mild cognitive impairment, and those diagnosed with Alzheimer's disease. Our findings revealed a significant disparity in node-metric coupling across these groups. We observed that node size exhibited a substantial correlation with the resulting graph metrics. These two key findings suggest that node-metric coupling could potentially serve as a valuable biomarker for mental diseases. In summary, our findings underscore a robust connection between graph node size and graph metrics. Additionally, we identify noteworthy disparities in these correlations when comparing patient and control groups. This not only underscores the potential of our proposed joint approach, involving both node volume and graph metrics, as a potential brain-based marker but also emphasizes the necessity of meticulous evaluation when interpreting results in previous and forthcoming studies that rely on atlas-based methods employing nodes of irregular sizes. Interestingly, we observed variations in these correlations among different groups. Our findings specifically indicated that the Alzheimer's group exhibited predominantly positive correlations between FNC, and node properties compared to HC and MCI groups. This observation opens the possibility of considering NMC as a potential biomarker.

This study comes with some limitations. Firstly, our initial assessment focused on only a subset of graph metrics, and there exist numerous alternative approaches for characterizing a node's role within a network. Some of these metrics operate at a local level, considering connections with neighboring nodes, while others operate at a global level, taking the node's position within the entire network into account. Secondly, the dataset we utilized from the ADNI exhibited an imbalance between the number of Alzheimer's and control subjects, with more data available for controls. This imbalance can lead to a higher likelihood of negative correlation values. To address this issue and enable meaningful comparisons across varying subject sizes, we implemented a proportional thresholding mechanism to eliminate low-impact areas of nodes within the dataset. It's worth noting that alternative thresholding strategies could be explored, and a more comprehensive investigation into the impact of thresholding on the results is a goal for future research. Lastly, this study centered on a single data-driven node parcellation approach. The choice of a data-driven atlas was deliberate, as it enables the grouping of temporally coherent and distinguishable brain regions, offering several advantages such as heightened sensitivity and the ability to capture dynamic changes within nodes [19]. In future research, it would be valuable to conduct a more extensive analysis encompassing various atlas-based approaches, each characterized by distinct node-size ratios. Furthermore, exploring the potential linkage between these changes and different types of Mild Cognitive

Impairment (e.g., early MCI, late MCI, stable and progressive MCI) to determine if they hold predictive value for future cognitive decline could be an intriguing avenue for investigation.

In brief, our findings emphasize the substantial impact of node size on resulting graph metrics. This underscores both the challenge of comparing existing graph metric studies and the potential opportunity for node size to serve as an informative marker of brain health and disorders. Further research is essential to delve deeper into the node-metric coupling and its potential applications.

5. Conclusions

In the realm of neuroimaging research, it has been a common practice to disregard the influence of node size on graph metrics. However, our research has unveiled a compelling revelation: the size of a node plays a pivotal role in shaping the resulting graph metrics, and this interplay is not to be underestimated, particularly in the context of the statistical tests being employed. Our study introduces the concept of Node-Metric Coupling (NMC), shedding light on its potential to provide invaluable insights into the comparative analyses we undertake. Furthermore, NMC could potentially serve as a biomarker, highlighting its significance in understanding brain health and disorders. Looking ahead, future investigations should delve deeper into this intricate relationship. This exploration may involve conducting comparisons with voxel-level graphs, exploring the utility of uniformly sampled graphs, and conducting systematic assessments of the contrasts between atlas-based and data-driven approaches. These endeavors will contribute to a more comprehensive understanding of the dynamic interplay between node size and graph metrics in the context of neuroimaging studies.

References

- [1] A. Fornito, "Graph theoretic analysis of human brain networks," in *FMRI techniques and protocols*, Springer, 2016, pp. 283-314.
- [2] Q. K. Telesford, S. L. Simpson, J. H. Burdette, S. Hayasaka and P. J. Laurienti, "The brain as a complex system: using network science as a tool for understanding the brain," *Brain connectivity*, vol. 1, no. 4, pp. 295-308, 2011.
- [3] A. Fornito, A. Zalesky and E. Bullmore, *Fundamentals of brain network analysis*, Academic Press, 2016.
- [4] M. P. van den Heuvel and A. Fornito, "Brain networks in schizophrenia," *Neuropsychology review*, vol. 24, no. 1, pp. 32-48, 2014.
- [5] H. Falakshahi, V. M. Vergara, J. Liu, D. H. Mathalon, J. M. Ford, J. Voyvodic, B. A. Mueller, A. Belger, S. McEwen, S. G. Potkin and others, "Meta-modal Information Flow: A Method for Capturing Multimodal Modular Disconnectivity in Schizophrenia," *IEEE Transactions on Biomedical Engineering*, 2020.
- [6] E. M. Gordon, T. O. Laumann, B. Adeyemo, J. F. Huckins, W. M. Kelley and S. E. Petersen, "Generation and evaluation of a cortical area parcellation from resting-state correlations," *Cerebral cortex*, vol. 26, no. 1, pp. 288-303, 2016.
- [7] M. F. Glasser, T. S. Coalson, E. C. Robinson, C. D. Hacker, J. a. Y. E. a. U. K. a. A. J. Harwell, C. F. Beckmann, M. Jenkinson and others, "A multi-modal parcellation of human cerebral cortex," *Nature*, vol. 536, no. 7615, pp. 171-178, 2016.
- [8] N. Craddock and I. Jones, "Genetics of bipolar disorder," *Journal of medical genetics*, vol. 36, no. 8, pp. 585-594, 1999.
- [9] M. P. van den Heuvel, S. C. de Lange, A. Zalesky, C. Seguin, B. T. Yeo and R. Schmidt, "Proportional thresholding in resting-state fMRI functional connectivity networks and consequences for patient-control connectome studies: Issues and recommendations," *Neuroimage*, vol. 152, pp. 437-449, 2017.
- [10] Y. Du, Z. Fu, J. Sui, S. Gao, Y. Xing, D. Lin, M. Salman, A. Abrol, M. A. Rahaman, J. Chen and others, "NeuroMark: An automated and adaptive ICA based pipeline to identify reproducible fMRI markers of brain disorders," *NeuroImage: Clinical*, vol. 28, p. 102375, 2020.

- [11] Y. Du and Y. Fan, "Group information guided ICA for fMRI data analysis," *Neuroimage*, vol. 69, pp. 157-197, 2013.
- [12] Y. Du, E. A. Allen, H. He, J. Sui, L. Wu and V. D. Calhoun, "Artifact removal in the context of group ICA: A comparison of single-subject and group approaches," *Human brain mapping*, vol. 37, no. 3, pp. 1005-1025, 2016.
- [13] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: uses and interpretations," *Neuroimage*, vol. 52, no. 3, pp. 1059-1069, 2010.
- [14] M. F. Glasser, T. S. Coalson, E. C. Robinson, C. D. Hacker, J. Harwell, E. Yacoub, K. Ugurbil, J. Andersson, C. F. Beckmann, M. Jenkinson and others, "A multi-modal parcellation of human cerebral cortex," *Nature*, vol. 536, no. 7615, pp. 171-178, 2016.
- [15] O. Sporns, "Graph theory methods: applications in brain networks," *Dialogues in clinical neuroscience*, 2022.
- [16] A. S. Blevins and D. S. Bassett, "Reorderability of node-filtered order complexes," *Physical Review E*, vol. 101, no. 5, p. 052311, 2020.
- [17] R. J. Anderson, C. M. Long, E. D. Calabrese, S. H. Robertson, G. A. Johnson, G. P. Cofer, R. J. O'Brien and A. Badea, "Optimizing diffusion imaging protocols for structural connectomics in mouse models of neurological conditions," *Frontiers in physics*, vol. 8, p. 88, 2020.
- [18] C. Staiger, S. Cadot, R. Kooter, M. Dittrich, T. Muller, G. W. Klau and L. F. Wessels, "A critical evaluation of network and pathway-based classifiers for outcome prediction in breast cancer," *PloS one*, vol. 7, no. 4, p. e34796, 2012.
- [19] Q. Yu, Y. Du, J. Chen, H. He, J. Sui, G. Pearlson and V. D. Calhoun, "Comparing brain graphs in which nodes are regions of interest or independent components: A simulation study," *Journal of neuroscience methods*, vol. 291, pp. 61-68, 2017.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.