Article

Variability and reproducibility of directed and undirected functional MRI connectomes in the human brain

Allegra Conti [1,*], Andrea Duggento [2], Maria Guerrisi [2], Luca Passamonti [3,4,*], Iole Indovina [1,5] and Nicola Toschi [2,6,*]

[1] Laboratory of Neuromotor Physiology, IRCCS Santa Lucia Foundation, Rome, Italy
[2] Department of Biomedicine and Prevention, University of Rome "Tor Vergata", Rome, Italy
[3] Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK
[4] Institute of Bioimaging & Molecular Physiology, National Research Council, Milano, Italy
[5] Saint Camillus International University of Health and Medical Sciences, Rome, Italy
[6] Department of Radiology, Athinoula A. Martinos Center for Biomedical Imaging, Boston, MA, USA
* Correspondence: a.conti@hsantalucia.it (A.C.), lp337@medschl.cam.ac.uk (L.P), toschi@med.uniroma2.it (N.T.)

Received: date; Accepted: date; Published: date

Abstract: A growing number of studies focus on methods to estimate and analyze the functional connectome of the human brain. Graph theoretical measures are commonly employed to interpret and synthesize complex network-related information. While resting state functional MRI (rsfMRI) is often employed in this context, is known to exhibit poor reproducibility, a key factor which is commonly neglected in typical cohort studies using connectomics-related measures as biomarkers. We aimed to fill this gap by analyzing and comparing inter- and intra- subject variability of connectivity matrices as well as graph-theoretical measures in a large (n=1003) database of young healthy subjects which underwent four consecutive rsfMRI sessions. We analyzed both directed (Granger Causality and Transfer Entropy) and undirected (Pearson Correlation and Partial Correlation) time-series association measures and related global and local graph-theoretical measures. While matrix weights exhibit a higher reproducibility in undirected as opposed to directed methods, this difference disappears when looking at global graph metrics and, in turn, exhibits strong regional dependence in local graphs metrics. Our results warrant caution in the interpretation of connectivity studies, and serve as a benchmark for future investigations by providing quantitative estimates for the inter- and intra- subject variabilities in both directed and undirected connectomic measures.

Keywords: Functional Networks; Functional magnetic resonance imaging; Connectome; Connectivity Matrices; Graphs; Reproducibility; Granger Causality; Transfer Entropy.

1. Introduction

The interest in studying directed and undirected interactions between different regions in the human brain (i.e. the functional 'connectome') is growing exponentially [1–4], and the advent of graph-theoretical applications to neuroscience has provided additional avenues to represent, analyze and interpret information contained in complex, possibly dynamic networks like the human connectome [1,2,5,6]. Functional connectome estimates are typically derived from neuromonitoring data (e.g. resting state functional MRI – rsfMRI), and established methods for computing whole-brain functional connectivity matrices include both undirected and directed estimators [7–9]. As is well known, reproducibility of MRI data varies widely across modalities, and rsfMRI data are known to exhibit significant inter- and intra-subject fluctuations [10–12]. While this can significantly bias and hamper the interpretation of functional connectivity studies, which do not typically include targeted scan-rescan experiments to assess reproducibility, the impact of this variability on connectome matrices and related measures has not yet been systematically investigated.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0031.v1

In this study, we explore, quantify and compare intra- and inter- subject variability (and hence reproducibility) of both directed and undirected resting state functional connectivity measures in a large (1003 subjects) database of high-quality rsfMRI data which comprises 4 scan sessions per subject. In detail, we investigate the reproducibility of whole-brain adjacency matrices derived from Pearson Correlation (PearC), Partial Correlation (PartC) as well as multivariate Granger Causality (mGC) and multivariate Transfer Entropy (mTE). In addition, we investigate the reproducibility of both global (whole-brain) and local (node-wise) graphs metrics calculated for all four types of adjacency matrices.

2. Materials and Methods

2.1. rsfMRI data

We employed rsfMRI data from 1003 subjects, part of the Human Connectome Project (HCP) (S1200 PTN release [13]). Each subject underwent a total of 4 resting state scans (2 sessions on 2 different days, 2 scans per session, 1200 timepoints/scan, TR 720 ms, TE 33.1 ms, flip angle 52, FOV 208x180, Thickness 2.0 mm; 72 slices; 2.0 mm isotropic voxel size, multiband factor 8, Echo spacing 0.58 ms, BW 2290 Hz/Px). In detail, we employed subject- and scan-wise wise fMRI timeseries (1200 points each) resulting from group independent component analysis (gICA) at dimensionality 15 (made available by the HCP consortium). Exemplary slices of the resulting node (i.e. component) maps, along with their physiological interpretations [14], are shown in Figure 1. In order to cater for confounds introduced by locally-varying hemodynamic response functions (HRFs) [15], fMRI timeseries were preprocessed using a the blind deconvolution approach [16] (maximum lag 10 s=14 time points, threshod: 1 standard deviation) implemented in a publicly available toolbox (https://users.ugent.be/~dmarinaz/HRF_deconvolution.html). Stationarity of all signals was verified using the Augmented Dickey-Fuller test. All signals were successively standardized prior to additional processing.

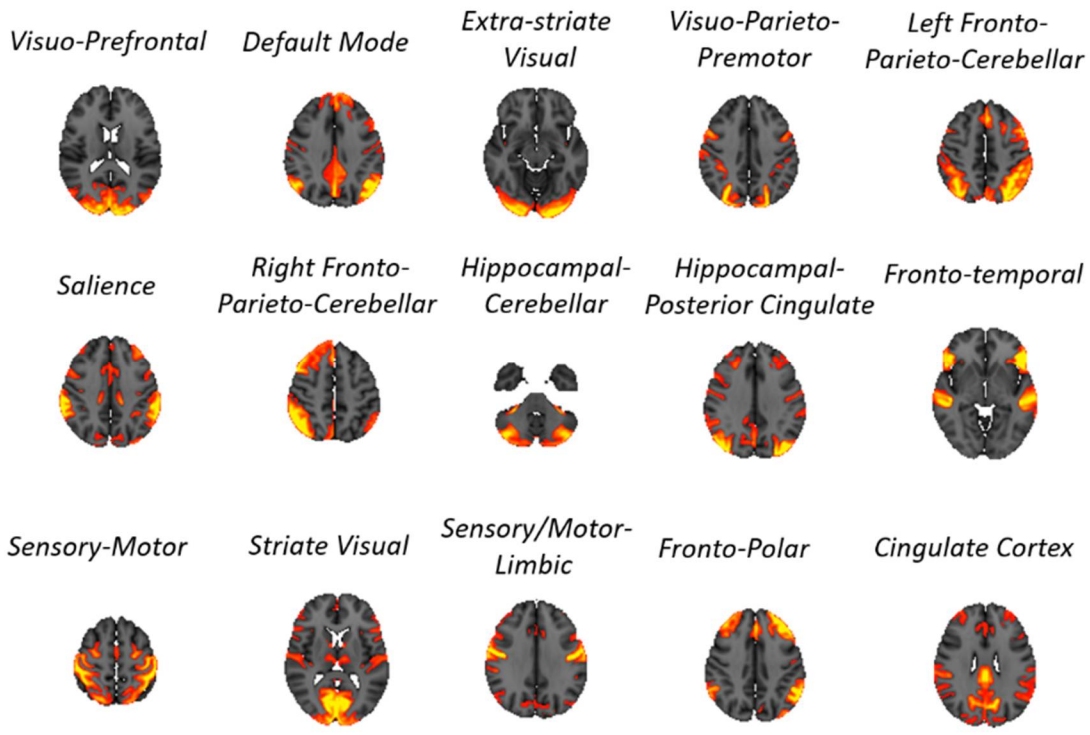
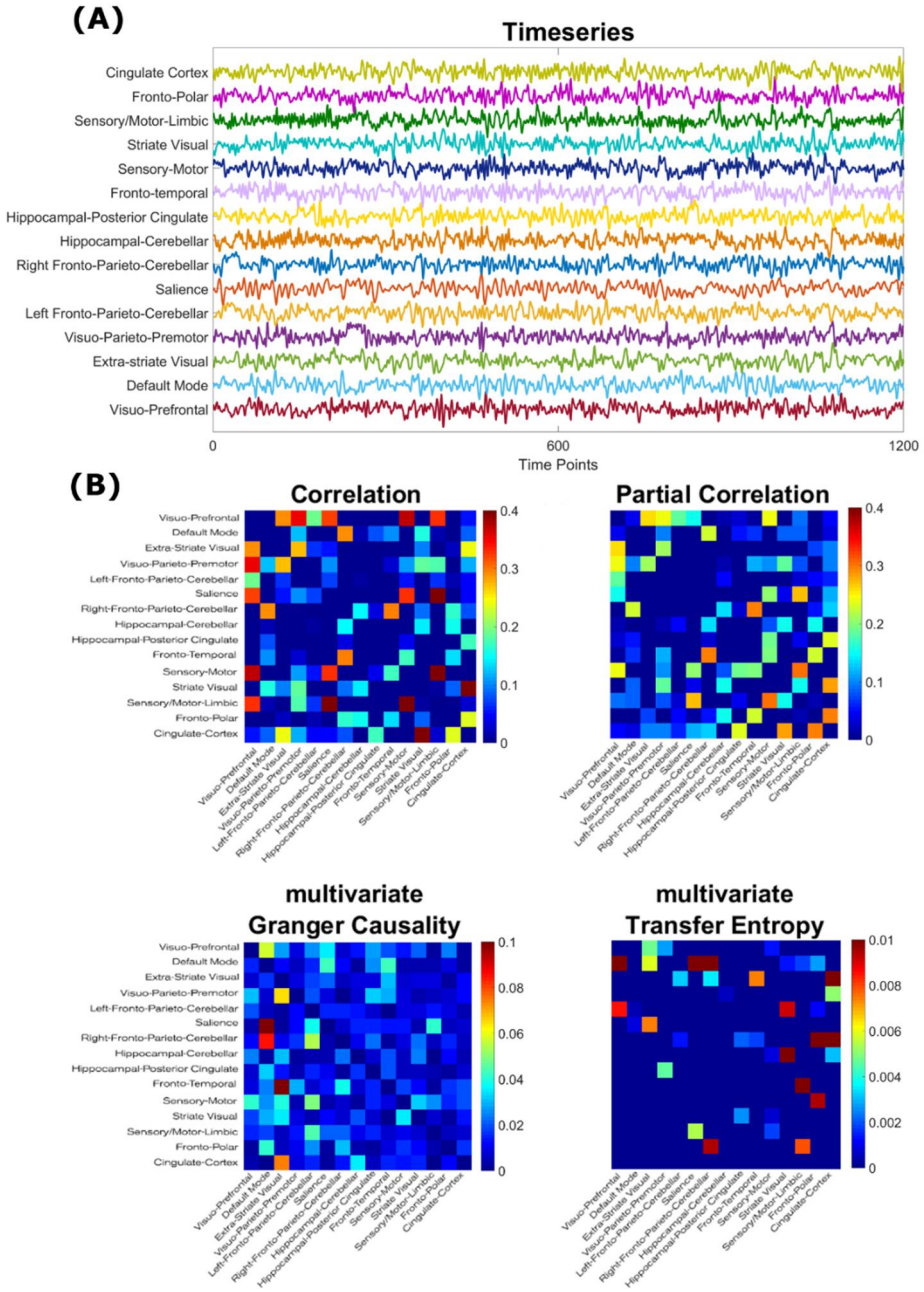


Figure 1. Brain functional networks (i.e. independent components) identified by group independent component analysis (gICA) in 1003 subjects drawn from the Human Connectome Project database. The subject-wise 15 node-specific time-series relative to these components were employed for all analyses in this paper.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0031.v1

2.2. Estimation of adjacency matrices

Starting from the 15 subject-, scan- and node-specific timeseries (see Figure 2(A) for an example), subject-wise undirected adjacency matrices were obtained through both PearC and PartC between all pairs of timeseries through in-house code written in in MATLAB (v. 2018a). Negative correlations were set to zero. Directed adjacency matrices where obtained through mGC estimates, in its most recent state-space formulation [17,18], and through mTE. mGC and mTE were calculated through in-house modified versions of the Matlab Tools for the computation of multiscale Granger Causality (http://www.lucafaes.net/msGC.html) [18] and of multivariate Transfer Entropy (http://www.lucafaes.net/cTE.html)[19]. In the case of GC, the autoregressive order was chosen my minimizing the median Schwartz criterion across all subjects (order=5). In mTE, conditioning vectors are formed according to a non-uniform building scheme selecting past terms up to a maximum lag of 5 time points while minimizing the conditional entropy [19]. Examples of adjacency matrices obtained through all four methods are shown in figure 2(B).



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0031.v1

Figure 2. (A): Example node-wise timeseries from a single scan, used to compute subject-, session-scan-specific adjacency matrices. (B): Examples of adjacency matrices obtained from both undirected (first row) and directed (second row) connectivity estimators for one subject. For each method, the median across 4 rsfMRI sessions from a single subject is shown. Diagonal elements are set to zero.

2.3. Global and local graph metric estimation

Starting from the adjacency matrices, we computed both global and local graph-theoretical indices of functional connectivity for all three methods/estimators. All graph measures were computed via the Brain Connectivity Toolbox [20] (https://sites.google.com/site/bctnet/), by using functions available for weighted adjacency matrices for both directed (in case of mGC and mTE) and undirected (for pearC and partC) graph metrics.

For each subject and for each scan, we calculated graph metrics quantifying the centrality of a node within a network (local strength and betweenness centrality), its ability to transmit information at local level (local efficiency) and its integration properties (clustering coefficient) [20]. Also, since global graph metrics characterize the overall organization of a network, we computed the global strength, the global efficiency and the global clustering coefficient as the average of the respective local metrics of all nodes. In addition, we calculated graph transitivity (a property related to the existence of tightly connected communities of nodes). Figure 3 shows how these metrics change both locally and globally as a function of neighborhood connectivity.

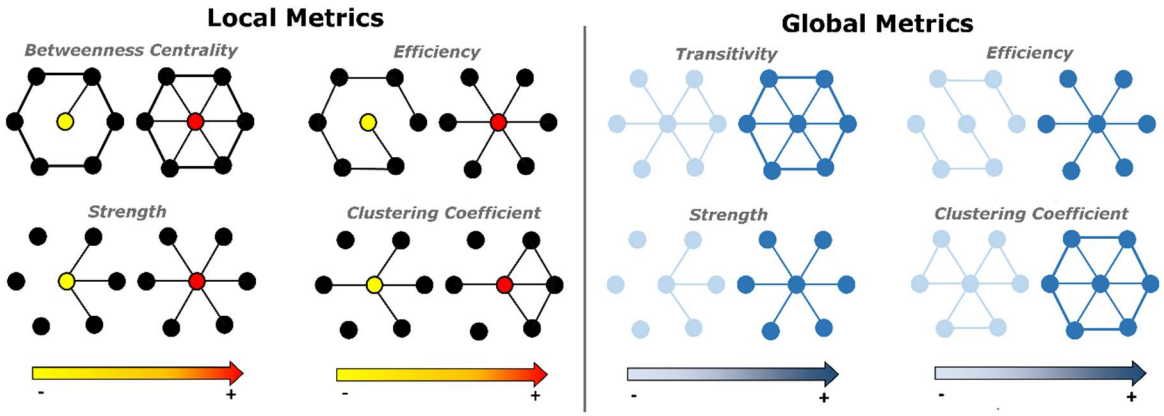


Figure 3. Exemplification of local (left) and global (right) graph metrics as a function of node neighborhood connectivity

2.4. Inter- and intra-subject variability distributions

In order to compare inter- and intra-subject fluctuations in adjacency matrices as well as graph-theoretical metrics, we employed dimensionless, normalized pairwise quantifiers of asymmetry, along with their estimated distributions. In detail, we defined the pairwise normalized difference/asymmetry (ND) as:

$$ND=\frac{a-b}{a+b} \quad (1)$$

where a and b refer to two distinct scans (either intra- or inter-subject). In the case of scalar quantities like global or local graph-theoretical metrics, Equation 1 can be applied directly. When comparing adjacency matrices, Equation 1 was modified by using Frobenius distances between (in this case) matrices a and b as follows:

$$ND=\sqrt{\frac{trace((a-b)\cdot(a-b)')}{trace((a+b)\cdot(a+b)')}} \quad (2)$$

In order to empirically estimate (and successively compare) the intra- and inter-subject distributions in ND, we proceeded as follows. From each set of 4 subject-wise scans, we sampled all 6 possible pairs of scans, computed the ND for each pair, and repeated this step for all 1003 subjects. For each

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0031.v1

method, this resulted in 1003x6=6018 distinct ND values referring to intra-subject variability. Successively, in order to build a comparable distribution related to inter-subject variability, we randomly and repeatedly sampled (without replacement) four scans from four distinct subjects, hence constructing set of 4 "inter-subject" scans from which 6 distinct pairs could be constructed and 6 distinct ND values could be calculated as above. In order to control for possible differences arising from intra-subject habituation, the random sampling was performed so that each "inter-subject" set of four scans contained exactly 1 sample from the 1003-sized set of first, second, third and fourth scans. This procedure was repeated 1003 times, resulting in 1003x6=6018 distinct ND values referring to inter-subject variability. Figure 4 summarizes the sampling process.

2.5.Statistical analysis

After sampling and computation of all ND values for each graph-theoretical metric and for all four types of adjacency matrix, the resulting "inter" and "intra" distribution medians of ND were compared through nonparametric Mann-Whitney-U tests. Effect Size (ES) were evaluated as the difference between the medians of the "inter" and "intra" distributions in ND (i.e. ES = "inter"-"intra"). The result of each comparison is therefore a p-value with an associated ES. Additionally, in order to explore possible differences in variability across nodes (i.e. brain regions), we pooled ND values for all local graph-theoretical metrics across all four methods (4 metrics x four methods = 16 values per node), and tested the effect of anatomical localization using a nonparametric Kruskal-Wallis test. Whenever a significant effect was found, pairwise comparison between nodes were performed. This procedure was repeated separately for inter- and intra- subject variability.

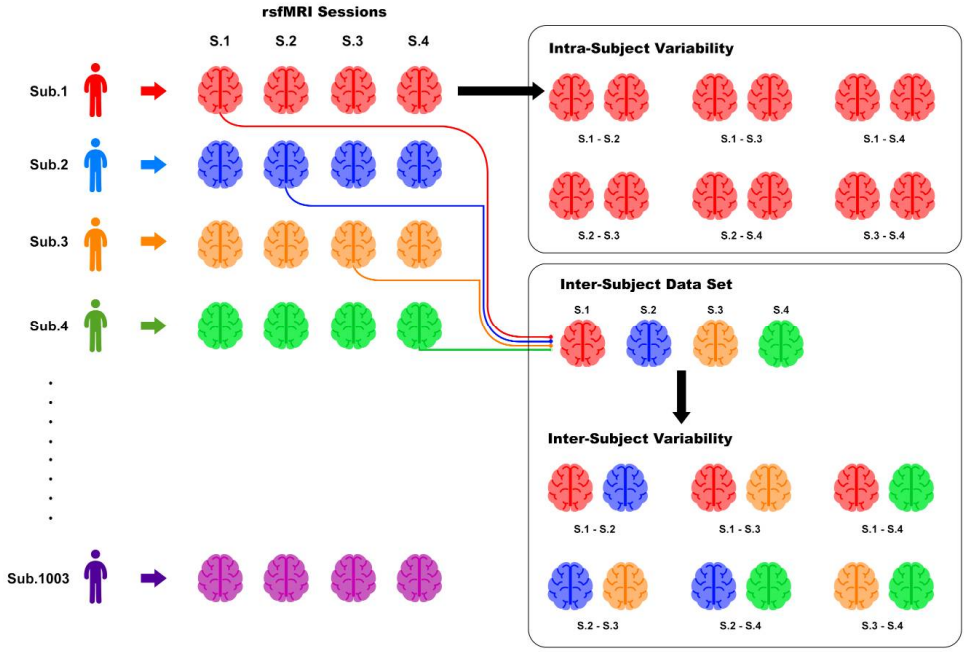


Figure 4. Sampling strategy used to construct intra- and inter-subject pairs and successively distributions of ND (Equations 1 and 2). The constructed samples for the basis for computing the distributions are shown e.g. in Figure 5 and Figure 6.

3. Results

Figure 5 shows the "inter" and "intra" distribution for ND obtained when comparing pairs of adjacency matrices (Equation 2) for all four methods (PearC, PartC, mGC and mTE), along with effect sizes. For all methods, as hypothesized, median intra-subject variability is significantly lower than median inter subject variability ($p<10^{-20}$ for all methods). The largest ES was associated with mGC and PearC (ES= 0.03 for both), followed by PartC and mTE (ES = 0.02 for both). Qualitatively, mTE

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0031.v1

showed the largest median variabilities (both intra- and inter-subject) whereas PearC and PartC showed similar median variabilities.

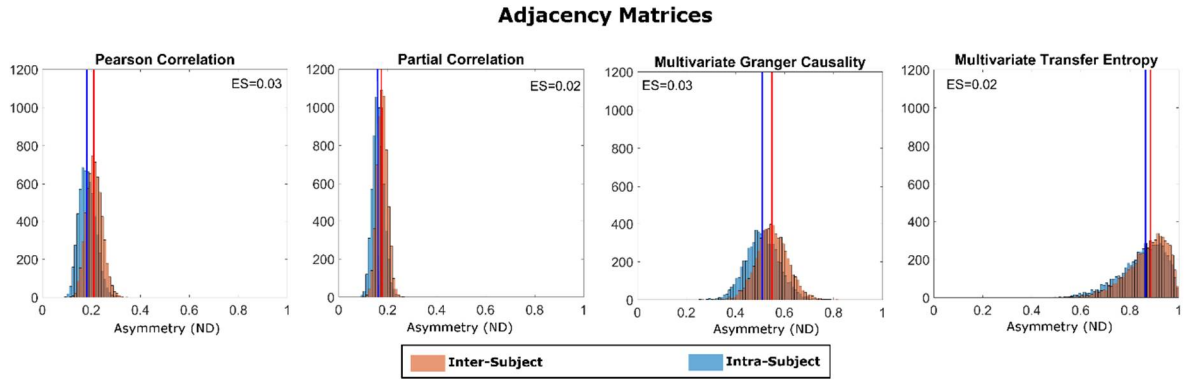
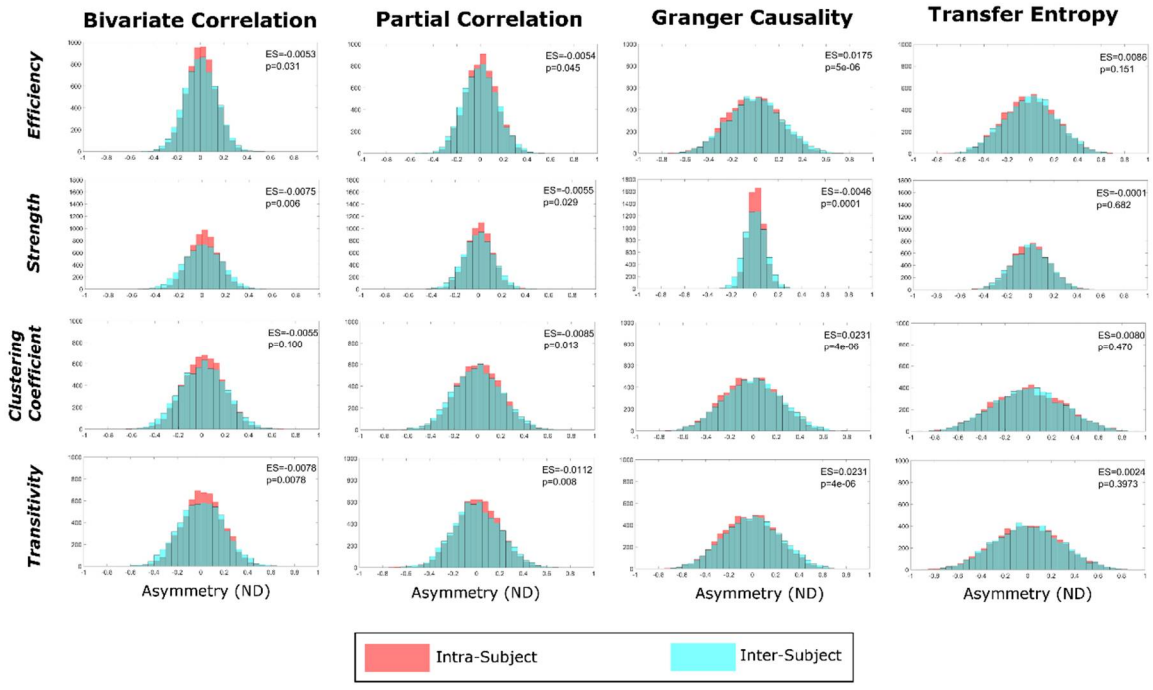


Figure 5. Intra- (blue) and inter- (orange) subject distributions of the Normalized Differences (ND), for all four connectivity estimation methods. Effect size (ES) (i.e. the difference between the median values of the inter- and intra- subject ND distributions) are shown in the insets. Mann-Whitney U-tests for comparing medians returned $p<10^{-20}$ for all four methods.

Figure 6 shows the "inter" and "intra" distribution for ND (Equation 1) obtained when comparing pairs of global graph-theoretical metrics, along with the results of statistical testing in terms of p-values and ES. Interestingly, intra- and inter-subject fluctuations are distinguishable only in some cases. In particular, we found statistically significant differences ($p<0.05$) for all metrics estimated through PearC (named "Bivariate Correlation" in figure) and PartC. However, in these cases intra- subject fluctuations were larger than inter-subject fluctuations (i.e. negative ES). On the other hand, when looking global metrics calculated through mGC, significant differences between intra- vs. inter-subject fluctuations were mainly associated with positive ES. No differences were found between ND distributions for global-graphs metrics evaluated through mTE. Notably, in this analysis all significant differences were associated with minimal (compared to unity) effect sizes (of the order of between 10^{-2} and 10^{-3}). In addition, the widths intra- and inter- subject variability distributions are qualitatively similar for all methods.



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0031.v1

Figure 6. Intra- (pink) and inter- (light blue) subject distributions of the Normalized Differences (ND), for all four connectivity estimation methods, along with the p-value resulting from the corresponding Mann-Whitney-U test. Effect size (ES) (i.e. the difference between the median values of the inter- and intra- subject ND distributions) are shown in the insets.

Figure 7 summarizes the results of comparing ND distributions computed from local graph-theoretical metrics, along with the results of statistical testing as well as ES. In this case, an intricate, anatomically dependent pattern emerged, where the statistical difference between median intra- and inter-subject fluctuations depends on the specific node, with some nodes exhibiting higher intra-subject (as compared to inter-subject) fluctuations. Asterisks indicate nodes in which differences between inter- and intra-subject variabilities resulted to be different from 0 ($p<0.05$, Mann-Whitney U test). These results highlight the strong anatomical dependence in reproducibility of all local graph-theoretical indices. Only the betweenness centrality of the Hippocampal-Posterior Cingulate network show statistically higher intra-subject (as opposed to inter-subject) variability across all methods. Still, ES in local efficiency and strength estimated in the same network tend to the same direction in all methods. Higher intra-subject (as opposed to inter-subject) variability of the betweenness centrality and local efficiency and strength was also found for the salience and the sensory motor networks. On the other hand, the default mode network displayed higher intra-subject (as opposed to inter-subject) stability in both betweenness centrality and local clustering coefficient across methods. The same hold for the betweenness centrality of the fronto-temporal network. The results shown in Fig.7 allow also to qualitatively compare the variability of the local metrics obtained though undirected (PearC and PartC) and directed (mGC and mTE) methods. In particular, local efficiency, strength and clustering coefficient in the visuo-prefrontal network are highly reproducible within the same subject when estimated using mGC and mTE.

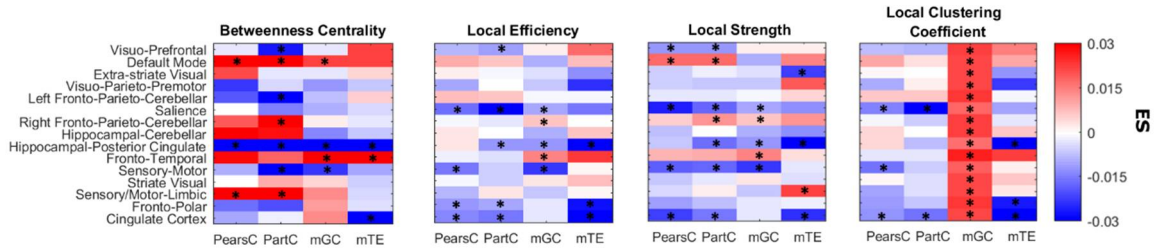


Figure 7. Effect size, i.e. differences between median intra- and inter- subject distributions in local graph metrics for each node and for each connectivity estimation method (PearC, PartC, mGC and mTE). For each metric darker red and blue cells correspond to higher and lower ES values, respectively. The nodes in which differences between inter- and intra-subject variabilities are statistically different from 0 ($p<0.05$) are indicated with an asterisk.

With reference to Figure 7, Table I summarizes the number of nodes where statistically significant intra- vs. inter-subject differences were found. Overall, across all four methods (PearC, PartC, mGC and mTE) the percentage of nodes (out of 15) where inter- and intra- subject variability are statistically different varied between 28% and 38% (bottom row of Table 1), and in 3%-35% of the 15 nodes the inter-subject variability is significantly higher than the intra-subject variability. mGC is the method with the lowest average number of nodes (across metrics) where the above mentioned difference is significantly different from zero. Also, node strength is the metric where the above mentioned difference is significantly different from zero in the highest average number of nodes (across methods). Further, it is interesting to note that, across metrics and methods, whenever the "intra" vs "inter" difference is significantly different from zero, this difference mainly results from a higher intra-subject variability. This Highlights the volatility and anatomical dependence of some local graph theoretical measures.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0031.v1

Table 1. Number of nodes showing statistically different intra- and inter- subject variabilities. In brackets: (number of nodes showing higher inter- vs intra subject variability/number of nodes showing higher intra- vs inter subject variability). The right column shows the average percentage of nodes (standard deviation in brackets), across metrics, where the inter-subject variability is significantly higher than the intra-subject one. The bottom row shows the percentage count of the first number in each cell.

Method	Betweenness Centrality	Local Efficiency	Strength	Clustering Coefficient	*Mean(sd) % out of 15*
PearsC	3(2/1)	4(0/4)	5(1/4)	3(0/3)	*5%(6%)*
PartC	7(3/4)	5(0/5)	7(2/5)	2(0/2)	8%(9%)
mGC	4(2/2)	5(2/3)	5(2/3)	15(15/0)	35%(38%)
mTE	3(1/2)	3(0/3)	4(1/3)	3(0/3)	*3%(3%)*
Mean(sd) % out of 15	*28%(13%)*	*28%(6%)*	*35%(8%)*	*38%(41%)*	

Figure 8(A) shows the results the Kruskal-Wallis test performed separately on the "intra" and on the "inter" ND values. In both intra- and inter- subject variability we found a significant effect of node ($p=1.1e-4$). In post-hoc comparisons (Figure 8(B)) we found that overall intra-subject variability in local graph-theoretical metrics of the fronto-temporal network is lower as compared to the hippocampal-posterior cingulate, the cingulate cortex and the fronto-polar networks. Also, overall inter-subject variability in local graph-theoretical metric is lower in the salience network as compared to the hippocampal-cerebellar, fronto-temporal and sensory/motor-limbic networks.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2019 doi:10.20944/preprints201905.0031.v1

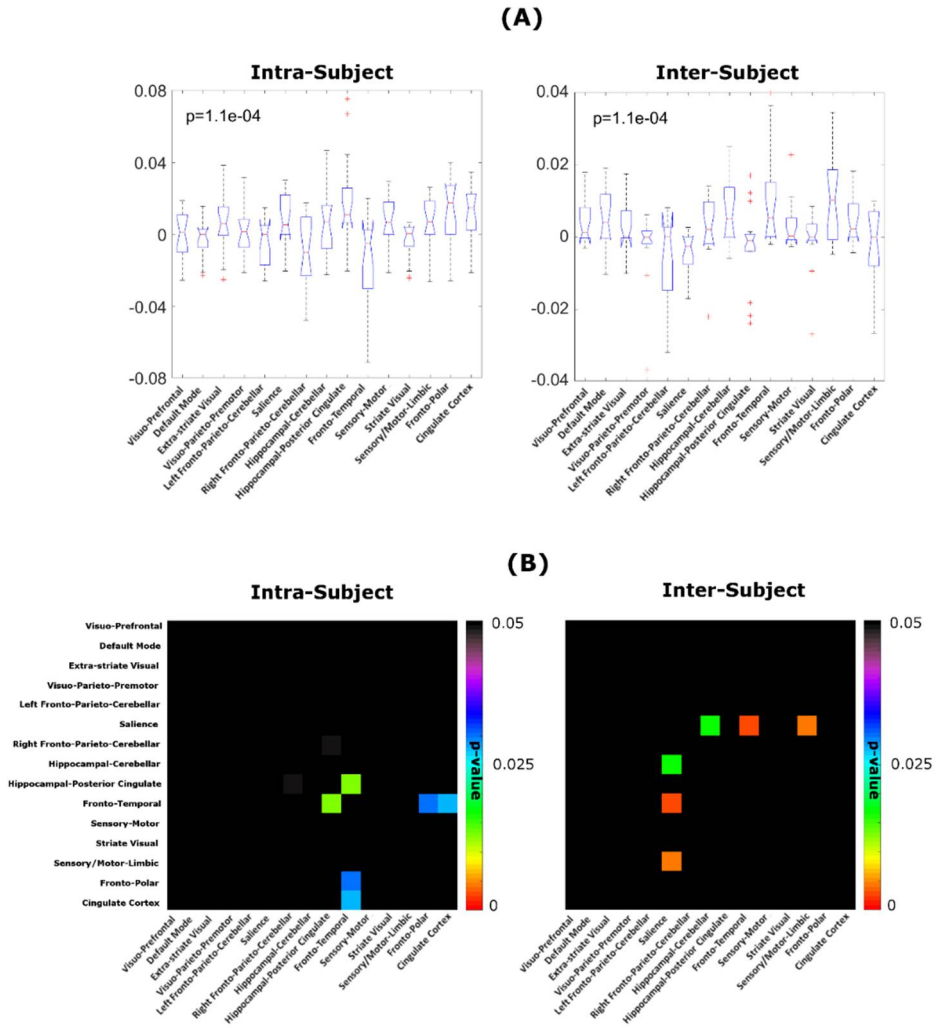


Figure 8. (A): Intra- and on the Inter- ND values (across 1003 subjects; left: intra-subject variability, right: inter-subject variability) for all local metrics (B): p-values obtained from pairwise comparison of node-specific, intra- and inter-subject variability.

4. Discussion and Conclusions

In this paper, we have studied inter- and intra-subject variability of directed and undirected connectivity matrices as well as of global as well as local graph-theoretical metrics derived from resting state functional MRI data in a large sample (n=1003) of exceptionally high quality resting state fMRI data, which also has the unique property of including scan-rescan sessions. While previous reproducibility studies exist, they are commonly focused on specific data processing [21] or statistical [22] aspects, and are commonly based on limited sample sizes. In order to eliminate unwanted sources of variability and improve the comparability between distributions, we have adopted a bootstrap approach followed by nonparametric statistics on a normalized difference metric, which also allowed us to pool across methods and/or nodes and make qualitative inter-method comparisons. Our results show that while adjacency matrix weights generated using conventional methods like PearC or PartC exhibit 1) a qualitatively lower variability as compared to directed metrics and 2) a lower within-subject v.s. between-subject variability, these differences virtually disappear when looking at global and local graphs-metrics values, where re-scanning the same subject results in the same statistical fluctuation as scanning a different subject. The overall higher variability in adjacency matrices obtained from directed metrics may be partially due to the higher complexity of both estimators and the multiple degrees of freedom (e.g. model order, embedding dimensions) in their computation, whose exploration was beyond the scope of this study. Also, the

appropriateness of using of mGC in fMRI has been the object of discussion [23–29], mainly because of the confounds introduced by locally varying neurovascular coupling and the interaction of these confounds with e.g. parameter choices and hence overall validity of the estimators. To this end, we applied a blind deconvolution approach which estimates a local HRF for each signal and returns an estimate of neural activity. Still, the application of this type of pipeline is not yet widespread in fMRI studies using causality methods, and that it also involves parameter choices whose effect on reproducibility merit investigation in a separate paper. Similarly, while partial and Pearson correlation do not involve parameter optimization, both mGC and mTE estimations are based on parameter choices inherent to each method. In this sense, the main aim of our work was to not to directly compare across methods, but rather to provide an estimation, in a high quality and unique data sample, of overall the intra- and inter- subject reproducibility of a variety of connectomic metrics which are commonly used in the fMRI community.

Interestingly, we found an articulate pattern of "intra" vs "inter" variability differences as a function of metrics, well as node (i.e. anatomical localization), which points towards a certain degree of unpredictability in the reproducibility of results obtained using each possible composite pipeline. Interestingly, in several nodes and metrics median intra-subject variability was found to be larger than median inter-subject variability. While we cannot exclude that the HCP database may intrinsically contain high intra-subject variability in raw data (e.g. due to the switching of phase encode direction in same-day rsfMRI scans), our findings warrant caution in the interpretation of graph-based connectomics studies, which should, whenever possible, include a scan-rescan validation arm at least in subset of subjects. They also serve as benchmarks for future investigations aiming to estimate the variability in other directed and undirected connectomic databases, and can also be employed for prospective power calculations in planning functional connectomics experiments.

Author Contributions: Conceptualization, N.T.; methodology, N.T., A.C. and A.D.; analysis, A.C.; investigation, N.T., A.C. A.D, L.P. and M.G.; writing—original draft preparation, N.T. and A.C.; supervision, N.T.; funding acquisition N.T., L.P. and I.I. All authors contributed to manuscript revision and editing. All authors contributed to paper validation and approved the final draft for submission.

Funding: Research supported by the Italian Ministry of Health (PE- 2013-02355372) and by the Medical Research Council (MRC) (MR/P01271X/1) at the University of Cambridge.

Acknowledgments: Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. This work was supported by the Italian Ministry of Health (PE-2013-02355372). Luca Passamonti is funded by the Medical Research Council (MRC) (MR/P01271X/1) at the University of Cambridge

Conflicts of Interest: The authors declare no conflict of interest.

 doi:10.20944/preprints201905.0031.v1

References

1. Duggento, A.; Passamonti, L.; Guerrisi, M.; Toschi, N. A realistic neuronal network and neurovascular coupling model for the study of multivariate directed connectivity in fMRI data. *Conf Proc IEEE Eng Med Biol Soc* **2018**, *2018*, 5537–5540.
2. Duggento, A.; Passamonti, L.; Valenza, G.; Barbieri, R.; Guerrisi, M.; Toschi, N. Multivariate Granger causality unveils directed parietal to prefrontal cortex connectivity during task-free MRI. *Sci Rep* **2018**, *8*, 5571.
3. Schmidt, C.; Pester, B.; Schmid-Hertel, N.; Witte, H.; Wismüller, A.; Leistritz, L. A Multivariate Granger Causality Concept towards Full Brain Functional Connectivity. *PLoS ONE* **2016**, *11*, e0153105.
4. Schoffelen, J.-M.; Hultén, A.; Lam, N.; Marquand, A.F.; Uddén, J.; Hagoort, P. Frequency-specific directed interactions in the human brain network for language. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 8083–8088.
5. Toschi, N.; Riccelli, R.; Indovina, I.; Terracciano, A.; Passamonti, L. Functional connectome of the five-factor model of personality. *Personal Neurosci* **2018**, *1*.
6. Vecchio, F.; Miraglia, F.; Maria Rossini, P. Connectome: Graph theory application in functional brain network architecture. *Clinical Neurophysiology Practice* **2017**, *2*, 206–213.
7. Deshpande, G.; LaConte, S.; James, G.A.; Peltier, S.; Hu, X. Multivariate Granger causality analysis of fMRI data. *Hum Brain Mapp* **2009**, *30*, 1361–1373.
8. Baccalá, L.A.; Sameshima, K. Partial directed coherence: a new concept in neural structure determination. *Biol Cybern* **2001**, *84*, 463–474.
9. Vicente, R.; Wibral, M.; Lindner, M.; Pipa, G. Transfer entropy--a model-free measure of effective connectivity for the neurosciences. *J Comput Neurosci* **2011**, *30*, 45–67.
10. Zhang, N.; Wang, F.; Turner, G.H.; Gore, J.C.; Avison, M.J.; Chen, L.M. Intra- and inter-subject variability of high field fMRI digit maps in somatosensory area 3b of new world monkeys. *Neuroscience* **2010**, *165*, 252–264.
11. Smith, S.M.; Beckmann, C.F.; Ramnani, N.; Woolrich, M.W.; Bannister, P.R.; Jenkinson, M.; Matthews, P.M.; McGonigle, D.J. Variability in fMRI: a re-examination of inter-session differences. *Hum Brain Mapp* **2005**, *24*, 248–257.
12. Seghier, M.L.; Price, C.J. Interpreting and Utilising Intersubject Variability in Brain Function. *Trends Cogn. Sci. (Regul. Ed.)* **2018**, *22*, 517–530.
13. Van Essen, D.C.; Smith, S.M.; Barch, D.M.; Behrens, T.E.J.; Yacoub, E.; Ugurbil, K. The WU-Minn Human Connectome Project: An Overview. *Neuroimage* **2013**, *80*, 62–79.
14. Toschi, N.; Duggento, A.; Passamonti, L. Functional connectivity in amygdalar-sensory/(pre)motor networks at rest: new evidence from the Human Connectome Project. *Eur. J. Neurosci.* **2017**, *45*, 1224–1229.
15. Deshpande, G.; Sathian, K.; Hu, X. Effect of hemodynamic variability on Granger causality analysis of fMRI. *Neuroimage* **2010**, *52*, 884–896.
16. Wu, G.-R.; Liao, W.; Stramaglia, S.; Ding, J.-R.; Chen, H.; Marinazzo, D. A blind deconvolution approach to recover effective connectivity brain networks from resting state fMRI data. *Med Image Anal* **2013**, *17*, 365–374.
17. Barnett, L.; Seth, A.K. Granger causality for state-space models. *Phys Rev E Stat Nonlin Soft Matter Phys* **2015**, *91*, 040101.
18. Faes, L.; Nollo, G.; Stramaglia, S.; Marinazzo, D. Multiscale Granger causality. *Phys Rev E* **2017**, *96*, 042150.
19. Faes, L.; Nollo, G.; Porta, A. Information Domain Approach to the Investigation of Cardio-Vascular, Cardio-Pulmonary, and Vasculo-Pulmonary Causal Couplings. *Front Physiol* **2011**, *2*.
20. Rubinov, M.; Sporns, O. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* **2010**, *52*, 1059–1069.

21. Song, X.; Panych, L.P.; Chou, Y.-H.; Chen, N.-K. A Study of Long-Term fMRI Reproducibility Using Data-Driven Analysis Methods. *Int J Imaging Syst Technol* **2014**, *24*, 339–349.
22. Chen, X.; Lu, B.; Yan, C.-G. Reproducibility of R-fMRI metrics on the impact of different strategies for multiple comparison correction and sample sizes. *Hum Brain Mapp* **2018**, *39*, 300–318.
23. Chen, Y.; Rangarajan, G.; Feng, J.; Ding, M. Analyzing multiple nonlinear time series with extended Granger causality. *Physics Letters A* **2004**, *324*, 26–35.
24. Benhmad, F. Modeling nonlinear Granger causality between the oil price and U.S. dollar: A wavelet based approach. *Economic Modelling* **2012**, *29*, 1505–1514.
25. Montalto, A.; Stramaglia, S.; Faes, L.; Tessitore, G.; Prevete, R.; Marinazzo, D. Neural networks with non-uniform embedding and explicit validation phase to assess Granger causality. *NEURAL NETWORKS* **2015**, *71*, 159–171.
26. Schreiber, T. Measuring Information Transfer. *Phys. Rev. Lett.* **2000**, *85*, 461–464.
27. *Directed Information Measures in Neuroscience*; Wibral, M., Vicente, R., Lizier, J.T., Eds.; Understanding Complex Systems; Springer-Verlag: Berlin Heidelberg, 2014; ISBN 978-3-642-54473-6.
28. Wen, X.; Rangarajan, G.; Ding, M. Is Granger Causality a Viable Technique for Analyzing fMRI Data? *PLOS ONE* **2013**, *8*, e67428.
29. Ramsey, J.D.; Hanson, S.J.; Hanson, C.; Halchenko, Y.O.; Poldrack, R.A.; Glymour, C. Six problems for causal inference from fMRI. *Neuroimage* **2010**, *49*, 1545–1558.