# Test-Retest Reliability of Computational Network Metrics Derived from the Structural Connectome of the Human Brain

Julia Parsosn Owen[1], Etay Ziv[1], Polina Bukshpun[2], Nicholas Pojman[2], Mari Wakahiro[2], Jeffrey Berman[3], Timothy Roberts[3], Elliott Sherr[2], and Pratik Mukherjee[1]
*[1]Radiology, UCSF, San Francisco, CA, United States, [2]Neurology, UCSF, San Francisco, CA, United States, [3]Radiology, CHOP, Philadelphia, PA, United States*

**Purpose:** Structural MR connectomics holds promise for improving diagnosis, outcome prediction and treatment monitoring of many common neurodevelopmental, psychiatric and neurodegenerative disorders for which there is currently no clinical utility of MR imaging [1]. Before computational network metrics from the human connectome can be applied in a clinical setting, their precision and their normative inter-subject variation must be understood in order to guide study design and the interpretation of longitudinal data. The reproducibility of commonly used graph theoretic measures is investigated, as applied to the structural connectome of healthy adult volunteers, using standard MRI acquisition parameters and a widely-used connectome processing pipeline.

**Methods:** Ten healthy control subjects were scanned twice at Site 1 and five healthy controls were scanned once at Site 1 and once at Site 2. All MR imaging was performed on a 3T TIM Trio MR scanner (Siemens, Erlangen, Germany) at each site, using 32-channel phased-array radio-frequency head coils. High-resolution structural MR imaging of the brain was performed with an axial 3D magnetization prepared rapid acquisition gradient-echo (MPRAGE) T1-weighted sequence. Whole-brain diffusion was performed with 30 diffusion-encoding directions, a diffusion-weighting strength of $b$=1000 s/mm$^2$ interleaved 2 mm axial sections and an in-plane resolution of 2 x 2 mm. FSL tools were used to perform motion correction, brain extraction, and probabilistic HARDI tractography seeded from the gray-white matter boundary [2] to create a whole-brain connectome. The processing pipeline is based on the M2 method in [3]. Global graph metrics: mean degree (K), characteristic path length (L), clustering coefficient (C), mean betweenness (B), global efficiency (E), and mean local efficiency ($E_{loc}$) are calculated for unweighted and weighted connectomes[4] and two levels of granularity of the connectome are evaluated, one based on the 82-node cortical and subcortical parcellation from FreeSurfer and one based on an atlas-free parcellation of the gray-white matter boundary consisting of 1000 cortical nodes. Intra-class correlation coefficient (ICC) and coefficient of variation (CV%) were used as measures of precision and variability. We use nonparametric permutation testing to determine p-values for these measures. The consistency of the unweighted and weighted edges using the correlation coefficient of the networks, and the module assignments using the Hubert rand index, are also computed for the 82-node connectomes.

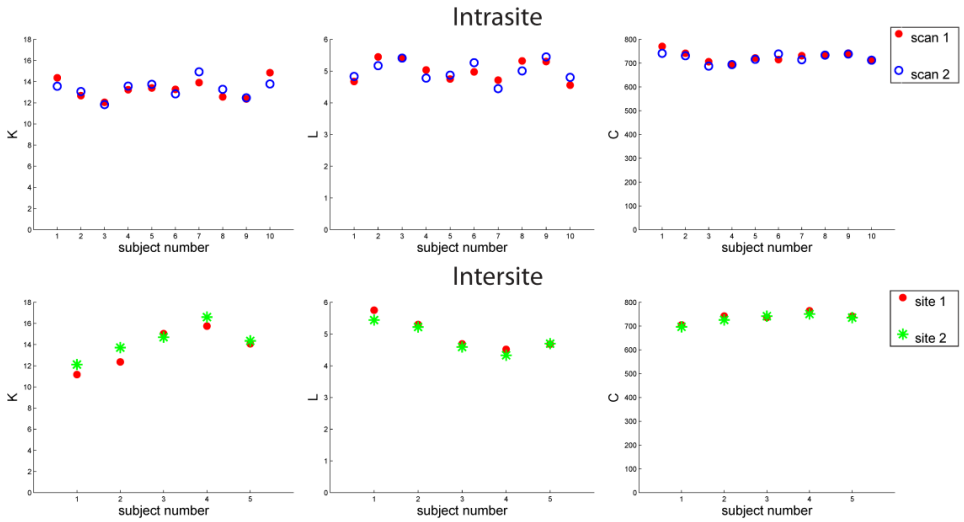


Figure 1: Demonstration of the test-retest reliability of unweighted metrics applied to the 1000-node cortical connectomes for the intra- and intersite data sets. Mean degree (K), characteristic path length (L), and mean clustering coefficient (C) are plotted.

**Results:** In Table 1, we report low CV% and high ICC values for the unweighted network metrics applied to both the 82-node and 1000-nodes connectomes. Only the ICC and CV% values for C and E calculated for the 82-node connectomes (intrasite) are not significant. In Figure 1, the global metric values are plotted to illustrate the test-retest reliability. The weighted metrics have similar CV% and ICC values, although weighted characteristic path length stands out as a less reproducible metric with a CV% of 34.75 for the 82-node connectome. We also find that the unweighted and weighted edges and module assignments are reproducible intra- and inter-site.

| | 82-Node Connectome | | 1000-Node Connectome | |
|---|---|---|---|---|
| **Intrasite** | CV% | ICC | CV% | ICC |
| K | 3.21* | 0.89* | 2.89* | 0.78* |
| L | 1.44* | 0.92* | 2.96* | 0.79* |
| C | 3.40 | 0.51 | 1.12* | 0.78* |
| B | 2.73* | 0.90* | 4.89* | 0.79* |
| E | 1.18* | 0.92* | 2.58* | 0.78* |
| $E_{loc}$ | 1.79 | 0.54 | 0.42* | 0.75* |
| Intersite | CV% | ICC | CV% | ICC |
| K | 6.26* | 0.79* | 3.81* | 0.89* |
| L | 2.26* | 0.82* | 1.98* | 0.94* |
| C | 3.05 | 0.59* | 0.94* | 0.88* |
| B | 4.00* | 0.82 | 3.31* | 0.94* |
| E | 1.94* | 0.84* | 1.74* | 0.94* |
| $E_{loc}$ | 1.61 | 0.63* | 0.42* | 0.87* |

Table 1: CV% and ICC values for unweighted network metrics , * denotes p<0.05.

**Discussion and Conclusion:** Overall, the results demonstrate good to excellent test-retest reliability for the graph analytics in both the intrasite and intersite datasets. The weighted and unweighted global network metrics applied to the 1000-node connectome yield better reproducibility than as applied to the 82-node connectomes. The ICC and CV% values indicate, in general, better reproducibility and precision than the findings of other connectome variability papers [5,6,7]. This improvement is most likely due to methodological differences, including tractography algorithm and thresholding. The findings reported here indicate that computational network metrics from the structural connectome have sufficient precision to be tested as potential biomarkers for diagnosis, prognosis, and monitoring of interventions in neurological and psychiatric diseases.

**References:** [1] Sporns, et al., PLoS Comput Bio, 2005; [2] Behrens, et al., Neuroimage, 2007; [3] Li, et al., Human Brain Mapping, 2012; [4] Rubinov, et al., Neuroimage, 2010; [5] Vaessen, et al., Neuroimage, 2010; [6] Bassett, et al., Neuroimage, 2010.; [7] Cheng, Neuroimage, 2012.