Title

A Regression Framework for Brain Network Distance Metrics

Authors

Chal E Tomlinson[1], Paul J. Laurienti[2,3], Robert G. Lyday[2,3], Sean L. Simpson[2,4,*]

[1] Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[2] Laboratory for Complex Brain Networks, Wake Forest School of Medicine, Winston-Salem, NC, USA.

[3] Department of Radiology, Wake Forest School of Medicine, Winston-Salem, NC, USA.

[4] Department of Biostatistics and Data Science, Wake Forest School of Medicine, Winston-Salem, NC, USA.

* **Corresponding Author**. Department of Biostatistics and Data Science, Wake Forest School of Medicine, Winston-Salem, NC 27157, USA. Email: slsimpso@wakehealth.edu

Abstract

Analyzing brain networks has long been a prominent research topic in neuroimaging. However, statistical methods to detect differences between these networks and relate them to phenotypic traits are still sorely needed. Our previous work developed a novel permutation testing framework to detect differences between two groups. Here we advance that work to allow both assessing differences by continuous phenotypes and controlling for confounding variables. To achieve this, we propose an innovative regression framework to relate distances between brain network features to functions of absolute differences in continuous covariates and indicators of difference for categorical variables. We explore several similarity metrics for comparing distances between connection matrices, and adapt several standard methods for estimation and inference within our framework: Standard F-test, F-test with individual level effects (ILE), Generalized Least Squares (GLS), Mixed Model, and Permutation. We also propose a novel modification of the standard GLS approach for estimation and inference. We assess all approaches for estimation and inference via simulation studies, and illustrate the utility of our framework by analyzing the relationship between fluid intelligence and brain network distances in Human Connectome Project (HCP) data.

Keywords

graph theory, connectivity, fMRI, small-world, neuroimaging, Jaccard, Kolmogorov-Smirnov

Highlights

- Related distances between connection matrices to differences in covariates.
- Adapted methods for estimation and inference in this framework.
- Novel modification of the standard GLS approach.

- Assessment of methods and distance metrics via simulation.
- Analysis of the HCP data with the best approach for each distance metric.

1. Introduction

As brain network analyses have become popular in recent years, neuroimaging researchers often face the need to statistically compare brain networks (Simpson et al., 2013a). Despite the advances made, analysis methods are still needed that enable the comparison of networks while incorporating topological features inherent in each individual's network. In order to develop such an analysis, we can exploit the fact that brain networks often exhibit consistent organizations across subjects. For example, a number of studies have reported that nodes with particular characteristics (e.g., high degree) tend to coincide at the same spatial locations across subjects (Hagmann et al., 2008; Hayasaka and Laurienti, 2010; Moussa et al., 2011; van den Heuvel et al., 2008). Although the set of such nodes may not be exactly the same across subjects, there are large areas of overlap (Hagmann et al., 2008; Hayasaka and Laurienti, 2010). Furthermore, our study on network modules, or communities of highly-interconnected nodes, indicated that some building blocks of resting-state functional brain networks exhibited remarkable consistency across subjects (Moussa et al., 2012). It has also been shown that such consistent organizations differ under different cognitive states (Deuker et al., 2009; Moussa et al., 2011; Rzucidlo et al., 2013) or in different groups of subjects (Bassett and Bullmore, 2009; Burdette et al., 2010; Liu et al., 2008; Meunier et al., 2009a; Rombouts et al., 2005; Stam et al., 2007; Yuan et al., 2010). Thus, an analysis method sensitive to such differences in spatial locations or patterns can assess network differences across the entire brain (as opposed to univariate edge-by-edge or node-by-node comparisons). Toward this end, in previous work we developed a permutation testing framework that detects whether the spatial location network features (such as the location of high degree nodes) mapped back into brain space differs between two groups of networks, and whether distributions of topological properties vary by group (Simpson et al., 2013b). Despite the utility of this method, it has two major weaknesses. First, it cannot account for confounding variables. This means that while we can compare maps of hub regions, for example, between two populations, we cannot control for differences in characteristics such as age or education. Second, the method relies on dichotomous grouping to make comparisons. When making comparisons between groups with clear divisions (male vs. female), this is not an issue. However, it is impossible to assess if there is a significant relationship between network hub location and continuous measures, such as intelligence quotient (IQ) score or age.

To address these limitations, we propose an innovative regression framework to relate distances between brain network features to functions of absolute differences in continuous covariates and indicators of difference for categorical variables. We will consider several different types of metrics for establishing distances (i.e., similarity/dissimilarity) between networks. The first type compares degree distributions. We accomplish this by summarizing similarities in nodal cumulative degree distributions across multiple networks with the Kolmogorov-Smirnov statistic (K-S statistic), a measure that quantifies the distance between two cumulative distribution functions (Kolmogorov, 1933; Smirnov, 1948). The second type takes into account consistency of key node sets. We do so by summarizing similarities in node sets across multiple networks with the Jaccard index, a metric that quantifies similarity in partitions of a set (Joyce et al., 2010; Meunier et al., 2009b). The third type of metric measures similarity of nodal degree by employing Minkowski or Canberra distances between nodal degree vectors (Lance and Williams, 1966).

Within our regression framework we adapt several methods for estimation and inference: Standard F-test, F-test with individual level effects (ILE), Generalized Least Squares (GLS), Mixed Model, and Permutation. Each observation in the regression framework includes a "distance" between two individuals, so observations that share individuals are correlated. Thus, the standard F-test is generally not appropriate, but presented for comparison. The other methods attempt to deal with this correlation: 1) F-test with ILE—including fixed individual level effects within the regression, possibly rendering the F-test valid; 2) Mixed Model—including random individual level effects; 3) GLS—proposing a novel modification of the standard GLS approach, with the novelty being the estimation of the covariance matrix (Aitken, 1936); and 4) Permutation—similarity and distance metrics have unknown distributions; thus, a permutation test may be most appropriate (Simpson et al., 2013b). Permutation tests for these models require permuting the residuals, and we will adapt recent methods to implement this approach (Kherad-Pajouh and Renaud, 2014, 2010).

In this paper, we detail our proposed regression framework, and discuss several methods for estimation and inference to be used on a variety of network similarity/dissimilarity metrics. We assess all combinations of methods and metrics within this framework using simulated fMRI data with known differences in connectivity matrix distributions. We then apply our framework to functional brain networks derived from the HCP dataset to investigate the relationship between fluid intelligence and network distances after accounting for known confounders.

2. Methods

2.1. Step 1: Network Construction

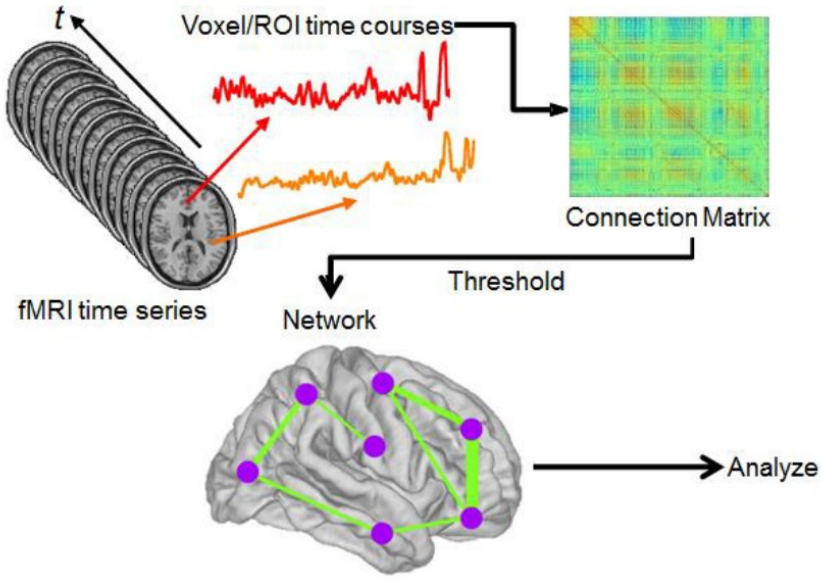

Figure 1: Schematic for generating brain networks from fMRI time series data—recreated from (Fornito et al., 2012; Simpson et al., 2013a)*. Functional connectivity between brain areas is estimated based on time series pairs to produce a connection matrix. A threshold is commonly applied to the matrix to remove negative and/or "weak" connections.*

It was assumed, for the purpose of this paper, that fMRI connection matrices had already been obtained (see Figure 1 recreated from (Fornito et al., 2012; Simpson et al., 2013a)). $\boldsymbol{C_i}$ represented a weighted connection matrix for individual

i, with matrix entries ranging from 0 (indicating no connection between the respective nodes) to 1 (strongest connection). We only considered undirected networks, so matrices were symmetric, with the *# of row = # of columns = # of nodes* (methods are easily adaptable if directed networks were desired).

b_i represented a binary key node vector for individual i, with an entry of 1 representing a key node and 0 a node that was not key. As stated in our previous work (Simpson et al., 2013b), key nodes could have been identified based on nodal characteristics such as high degree, high centrality, or other desired characteristics. Since key nodes were compared across subjects, it was important to employ the same criterion in all of the networks (e.g., top 10% highest centrality, node degree >200, etc.). Alternatively, key nodes could have been identified as those belonging to a particular module or network community, a collection of highly interconnected nodes. The resulting key nodes would form a set and the consistency of the spatial location of the nodes could be compared across groups. An example visualization of key node sets from voxel-based networks in brain space was given in Figure 2--recreated from (Simpson et al., 2013b).

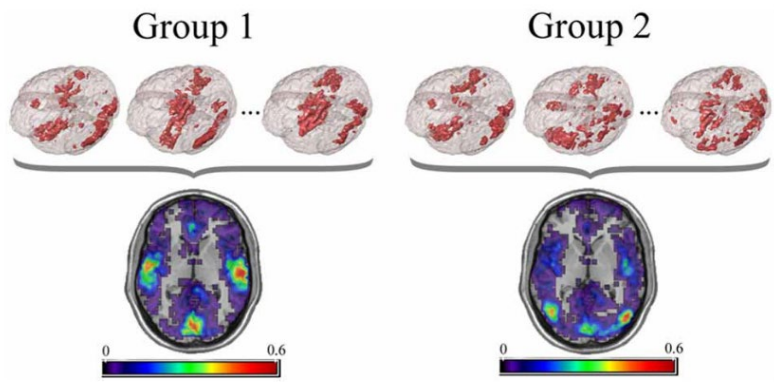

Figure 2: Example visualization of key node sets from voxel-based networks in brain space. The 3D brain images (Top) are 3 representative subjects from each group with the key node sets defined to be those with the top 20% highest degree. Overlap maps (Bottom) show the proportion of subjects with key nodes in given areas. Figure recreated from (Simpson et al., 2013b).

We constructed an $n \times 1$ weighted nodal degree vector d_i by summing $\boldsymbol{C_i}$ over its rows (or equivalently, columns). Each entry in vector d_i represented the "degree" of its respective node for individual i. As was described for the binary key node vector in the last paragraph, other weighted nodal vectors could have been used. For simplicity's sake, we focused on degree within the methods and simulation sections.

2.2. Step 2: Establish Similarity/Dissimilarity Between Networks

This section covered some of the metrics we used to gauge distance between individual networks.

2.2.1. KS Statistic

Degree distributions, which helped quantify the topology of networks, were likely more similar within distinctive groups than they were between these groups. We employed the KS statistic to quantify this potential dissimilarity.

$$KS_{ij} = sup_x \left| F_i(x) - F_j(x) \right|$$

KS_{ij} was the Kolmogorov-Smirnov statistic between nodal degree distributions for individuals i and j. $F_i(x)$ represented the empirical distribution function for observations from the nodal degree vector d_i. So, $sup_x|F_i(x) - F_j(x)|$ gave the biggest difference between the empirical degree distributions between individuals i and j. Bigger values indicated more dissimilarity.

2.2.2. Jaccard Index

$$J_{ij} = \frac{|b_i \cap b_j|}{|b_i \cup b_j|}$$

J_{ij} was the Jaccard index between key node vector i and key node vector j. Since $\boldsymbol{b_i}$ and $\boldsymbol{b_j}$ represented the (binary) key node vectors of individuals i and j, $|b_i \cap b_j|$ gave the number (count) of nodes that shared the same key status between the two networks and $|b_i \cup b_j|$ gave the total number of nodes. That is, J_{ij} gave the proportion of nodes that shared key status between individuals i and j. Values of J_{ij} ranged from 0 (no overlap) to 1 (perfect overlap--networks are the same).

2.2.3. Minkowski Distance (L^p norm distance)

$$M_{ij}^p = \|d_i - d_j\|_p = \left(\sum_{k=1}^{\#\, of\ nodes} |d_i[k] - d_j[k]|^p \right)^{\frac{1}{p}}$$

M_{ij}^p was the Minkowski distance of order p between degree distribution i and degree distribution j, where $d_i[k]$ represented the degree of node k for individual i. There were an infinite number of choices of Minkowski distances (the Manhattan distance when $p = 1$, the Euclidean distance when $p = 2$, etc.). Bigger values of Minkowski distances indicated more dissimilarity.

2.2.4. Canberra Distance

$$C_{ij} = \sum_{k=1}^{\#\, of\ nodes} \frac{|d_i[k] - d_j[k]|}{|d_i[k]| + |d_j[k]|}$$

C_{ij} was the Canberra distance of between degree distribution i and degree distribution j, where $d_i[k]$ represented the degree of node k for individual i. The Canberra Distance was a weighted version of the Manhattan distance (the Minkowski distance when $p = 1$). Bigger values of Canberra distances indicated more dissimilarity.

2.3. Step 3: Evaluating Differences between Networks

2.3.1. Model Formulation

$$Dist_{ij} = \boldsymbol{X_{ij,\mathrm{con}}^T \beta_{\mathrm{con}}} + X_{ij,\mathrm{coi}}^T \beta_{\mathrm{coi}} + \varepsilon_{ij}$$

$Dist_{ij}$ represented the distance between nodal vectors of individuals i and j. $Dist_{ij}$ was a generic placeholder for any metric outlined previously in Step 2, i.e., Jaccard Index (J_{ij}), KS Statistic (KS_{ij}), Minkowski Distance (M_{ij}^{p}), or Canberra Distance (C_{ij}).

$\boldsymbol{X_{ij,\mathrm{con}}^{T}}$ contained the intercept (optionally) and differences in confounding covariates (e.g., for our data, sex, educational attainment, age, and body mass index) between individuals i and j (with corresponding unknown parameter $\boldsymbol{\beta_{\mathrm{con}}}$) to control for differences that may have confounded the relationship between the covariate of interest and the given distance.

$X_{ij,\mathrm{coi}}^{T}$ contained the difference in the covariate of interest (or an indicator of group membership for group-based analyses) between individuals i and j (with corresponding unknown parameter β_{coi}).

ε_{ij} accounted for the random errors in the distance (Jaccard, KS, etc.) values. For ε_{ij} we assumed a common variance parameter, and a common covariance parameter between ε_{ij} and $\varepsilon_{i'j'}$, when either $i = i'$ or $j = j'$.

For example, if we tested whether there was an association between IQ (continuous) and the spatial consistency of hub nodes (top 20% highest degree) after controlling for age (continuous), sex (binary), and treatment (binary) status, our model would have been

$$J_{ij} = \beta_0 + \left|Age_i - Age_j\right|\beta_1 + \mathbb{1}\{Sex_i \neq Sex_j\}\beta_2 + \mathbb{1}\{Trt_i \neq Trt_j\}\beta_3 + \left|IQ_i - IQ_j\right|\beta_4 + \varepsilon_{ij}$$

with the associated null hypothesis $H_0: \beta_4 = 0$.

2.3.2. Standard F-test

If ε_{ij} were independent, homoscedastic, and approximately normally distributed, the F-test of a standard linear regression would have been an appropriate test. We expected correlation among observations from the same individual, so we included this standard testing procedure here mainly for comparison.

2.3.3. Standard F-test with Individual Level Fixed Effects (F-test with ILE)

$$\mathbf{Dist} = \boldsymbol{X^T}\boldsymbol{\beta} + \mathbf{ID_1}\alpha_1 + \cdots + \mathbf{ID_p}\alpha_p + \boldsymbol{\epsilon}$$

$\mathbf{Dist}$ was an $n \times 1$ vector of known distance metrics (as outlined in Step 2), $\boldsymbol{X^T}$ was the $n \times m$ design matrix of known covariates with corresponding $m \times 1$ unknown parameter vector $\boldsymbol{\beta}$, $\mathbf{ID_i}$ was the $n \times 1$ known indicator variable for individual i with corresponding unknown parameter α_i (where $n =$ the number of observations $= \binom{p}{2}$, $m =$ the number of covariates, and $p =$ the number of participants). By accounting for individual level effects, we hoped to induce independence, homoscedasticity, and normality for the error terms within the random vector $\boldsymbol{\epsilon}$. This (potentially) allowed for an F-test to appropriately evaluate the covariates of interest.

2.3.4. Mixed Model Approach

It should be mentioned that we attempted a mixed model formulation where each individual had their own random effect. Unfortunately, these calculations were too computationally intensive in the simulations we ran. Instead, we took the following approach:

$$\mathbf{Dist} = \boldsymbol{X^T\beta} + \mathbf{ID_1}\mathrm{b}_1 + \mathbf{ID_2}\mathrm{b}_2 + \boldsymbol{\epsilon}$$

$\mathbf{ID_1}$ and $\mathbf{ID_2}$ were $n \times 1$ categorical vectors representing the pair of individuals involved in each observation. $\mathbf{ID_1}$ represented one individual in the pair and $\mathbf{ID_2}$ the other. We attempted several methods to best assign which individual within an observation was placed either $\mathbf{ID_1}$ or $\mathbf{ID_2}$. Results did not vary meaningfully no matter how an individual was assigned, but a method was chosen that unbiasedly balanced an equal (or close to equal) number of individuals in each covariate. The corresponding unknown random effects were represented by b_1 and b_2, with mean $E\begin{pmatrix}\mathrm{b}_1\\ \mathrm{b}_2\end{pmatrix} = \begin{pmatrix}0\\ 0\end{pmatrix}$ and independent variance-covariance matrix $var\begin{pmatrix}\mathrm{b}_1\\ \mathrm{b}_2\end{pmatrix} = \boldsymbol{G} = \begin{pmatrix}g_{11} & 0\\ 0 & g_{22}\end{pmatrix}$. Parameter estimation was conducted via restricted maximum likelihood (REML), with the Satterthwaite approximation of the F-test for a Wald statistic employed for inference.

2.3.5. GLS Approach

$$\mathbf{Dist} = \boldsymbol{X^T\beta} + \boldsymbol{\epsilon}$$

$\mathbf{Dist}$ and $\boldsymbol{X^T\beta}$ were as before, but instead we assumed $\boldsymbol{\epsilon} \sim (\mathbf{0},\ \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ was the $n \times \mathrm{n}$ covariance matrix. Generalized least squares (GLS) allowed for estimating parameters when there was correlation among the residuals in ordinary least squares regression. However, GLS required $\boldsymbol{\Sigma}$ to be known. An unrestricted $n \times \mathrm{n}$ covariance matrix had $n(n+1)/2$ parameters to estimate. This was infeasible as we only had n observations. Thus, we restricted the form of $\boldsymbol{\Sigma}$ in order to estimate it.

Here we proposed a novel modification of the standard GLS approach, with the novelty being the estimation of the covariance matrix. Let $\boldsymbol{\Sigma} = \boldsymbol{\sigma^2}\mathbf{I_n} + \boldsymbol{\tau^2}\mathbf{ID_{mat}}\mathbf{ID_{mat}^T}$, where $\mathbf{I_n}$ was the $n \times \mathrm{n}$ identity matrix and $\mathbf{ID_{mat}} = [\mathbf{ID_1} \mid \mathbf{ID_2} \mid \cdots \mid \mathbf{ID_p}]$ was an $n \times p$ matrix such that each $\mathbf{ID_i}$ was the $n \times 1$ indicator variable for individual i. Now, in order to estimate $\boldsymbol{\Sigma}$, we just needed to estimate two parameters: σ^2 and τ^2. Taking a closer look at $\boldsymbol{\Sigma}$, we noticed that each diagonal element was $\sigma^2 + 2\tau^2$, some off-diagonal elements were τ^2, and all other off-diagonal elements were 0.

Since each diagonal element, $\sigma^2 + 2\tau^2$, was the same, we estimated it using sample variance:

$$\widehat{\sigma^2 + 2\tau^2} = \frac{1}{n}\left(\mathbf{Dist} - \boldsymbol{X_i^T}\beta\right)^T\left(\mathbf{Dist} - \boldsymbol{X_i^T}\beta\right)$$

Since we had that some off-diagonals of $\boldsymbol{\Sigma}$ are τ^2 and the rest are 0, each nonzero off-diagonal was the same value. Thus, we estimated τ^2 using the *idea* of sample covariance:

$$\widehat{\tau^2} = \frac{1}{m}\left(\mathbf{Dist} - \boldsymbol{X}_i^T\beta\right)^T\left(\mathbf{ID_{mat}ID_{mat}^T} - \mathbf{2}\cdot\mathbf{I_n}\right)\left(\mathbf{Dist} - \boldsymbol{X}_i^T\beta\right)$$

where m represented the sum of all entries in $\mathbf{ID_{mat}ID_{mat}^T} - \mathbf{2}\cdot\mathbf{I_n}$, which gave us the number of non-zero off-diagonal entries in the variance-covariance matrix.

Next, we pre-multiplied our original model with $\mathbf{L^{-1}}$, the inverse of Cholesky decomposition of $\boldsymbol{\Sigma}$ such that $\boldsymbol{\Sigma} = \mathbf{LL^T}$:

$$\mathbf{L^{-1}Dist} = \mathbf{L^{-1}}\boldsymbol{X^T}\boldsymbol{\beta} + \mathbf{L^{-1}}\boldsymbol{\epsilon} \xrightarrow{\textit{yields}} \mathbf{Dist}^* = \boldsymbol{X}^{*T}\boldsymbol{\beta} + \boldsymbol{\epsilon}^*$$

where $\boldsymbol{\epsilon}^* = \mathbf{L^{-1}}\boldsymbol{\epsilon} \sim \left(\mathbf{L^{-1}}\cdot\mathbf{0_n}, \mathbf{L^{-1}\Sigma L^{-T}}\right) = (\boldsymbol{0_n}, \boldsymbol{I_n})$. Finally, since post-transformation it is assumed entries of $\boldsymbol{\epsilon}$ were independent, homoscedastic, and approximately normally distributed, we applied the F-test.

2.3.6. Permutation Test

A permutation test requires no knowledge of how the test statistic of interest is distributed under the null hypothesis (e.g., H_0: no significant difference among IQ). The distribution under the null hypothesis is empirically "generated" by permuting data labels.

It is important to note that permutation tests are not assumption free and do assume exchangeability (i.e., that when the null hypothesis is true the joint distribution of the observations remains unchanged under permutation of the group labels). Thus, we were careful when permuting data labels. That is, we reasonably assumed ε_{ij} and $\varepsilon_{i'j'}$ were independent and from the same distribution if individuals i, j, i' and j' are all unique. However, if there was overlap between individuals, we should not expect this assumption to hold. To preserve exchangeability, we should not randomly permute across all observations, and instead permuted across individuals.

To illustrate how to randomly permute across individuals rather than all observations, consider a situation where we wanted to permute the residuals, r_{ij}. If we permuted across all observations, we would have randomly shuffled the r_{ij} column. To permutate across individuals (therefore preserving exchangeability), we instead would have randomly shuffled the individual labels. See Figure 3 for an example:

i	j	r_{ij}
Alpha	Bravo	1
Alpha	Charlie	2
Alpha	Delta	3
Bravo	Charlie	1
Bravo	Delta	2
Charlie	Delta	1

sample permutation:

Alpha -> Charlie

Bravo -> Alpha

Charlie -> Delta

Delta -> Bravo

i	j	r_{ij}
Charlie	Alpha	1
Charlie	Delta	2
Charlie	Bravo	3
Alpha	Delta	1
Alpha	Bravo	2
Delta	Bravo	1

Figure 3: To illustrate how to randomly permute across individuals rather than all observations, consider a situation where we wanted to permute the residuals, r_{ij}. If we permuted across all observations, we would have randomly shuffled the r_{ij} column. To permutate across individuals (therefore preserving exchangeability), we instead would have randomly shuffled the individual labels.

3. Simulation Studies

3.1. Data

We conducted four simulation studies to assess how well our proposed approaches could detect various relationships between brain network properties and covariates of interest. Each simulation contained 100 subjects, with four covariates of interest. A fair coin was flipped for each subject to determine their gender (SEX = male or female). Half of subjects were assigned to treatment and the other half were assigned to placebo (TRT = treatment or placebo). IQ and Age were both simulated from a normal distribution with mean of 100 and a standard deviation of 15 (rounded to the nearest integer). This resulted in 2 binary (SEX, TRT), and 2 continuous (AGE, IQ) covariates--variables were given names purely for purposes of explication.

For simulations 1-3, we simulated fMRI connectivity matrices with 268 nodes each to mimic the experimental data detailed in the next section. In each simulation, a 268x268 symmetric matrix (with entries ranging from 0 to 1) was generated for each subject. Entries within each connectivity matrix were drawn from three types of distributions: a) a noise distribution, $Beta\left(\frac{4}{3},6\right)$; b) a "same-for-all" signal distribution, $Beta(4,6)$; and c) a signal distribution dependent on covariates and signal percentage, $Beta\left(s_p \cdot a_i + \left(1 - s_p\right) \cdot \frac{4}{3}, 6\right)$ where $a_i = max\left(\frac{4}{3}, 4 + (IQ_i - 100) * .2 + (Trt_i = = "Treatment") * 6)\right)$ represented the covariate-dependent parameter and s_p represented the signal percentage (from 0 to 100%). When the signal percent (s_p) was 100%, $Beta\left(s_p \cdot a_i + \left(1 - s_p\right) \cdot \frac{4}{3}, 6\right) = Beta(a_i, 6)$. Similarly, when signal percent was 0%, $Beta\left(s_p \cdot a_i + \left(1 - s_p\right) \cdot \frac{4}{3}, 6\right) = Beta\left(\frac{4}{3}, 6\right)$, and was therefore identical to the noise distribution and no longer dependent on covariates.

Simulation 1 had a 15x15 region where all individuals had the same signal distribution, a 15x15 region where the signal distribution was dependent on covariates, and the rest of the 256x256 connection matrix was drawn from the noise distribution. Simulation 2 was the same as 1, but the signal distribution dependent on covariates was moved to expand the border of the 15x15 same signal distribution to a combined 21x21 signal region. Simulation 3 was the same as 1, with two additional 15x15 regions where all individuals had the same signal distribution. For a drawn to scale representation of these simulations, see Figure 4. It should be noted that we assumed the connectivity matrices to be symmetric, with 0 entries along the diagonal.

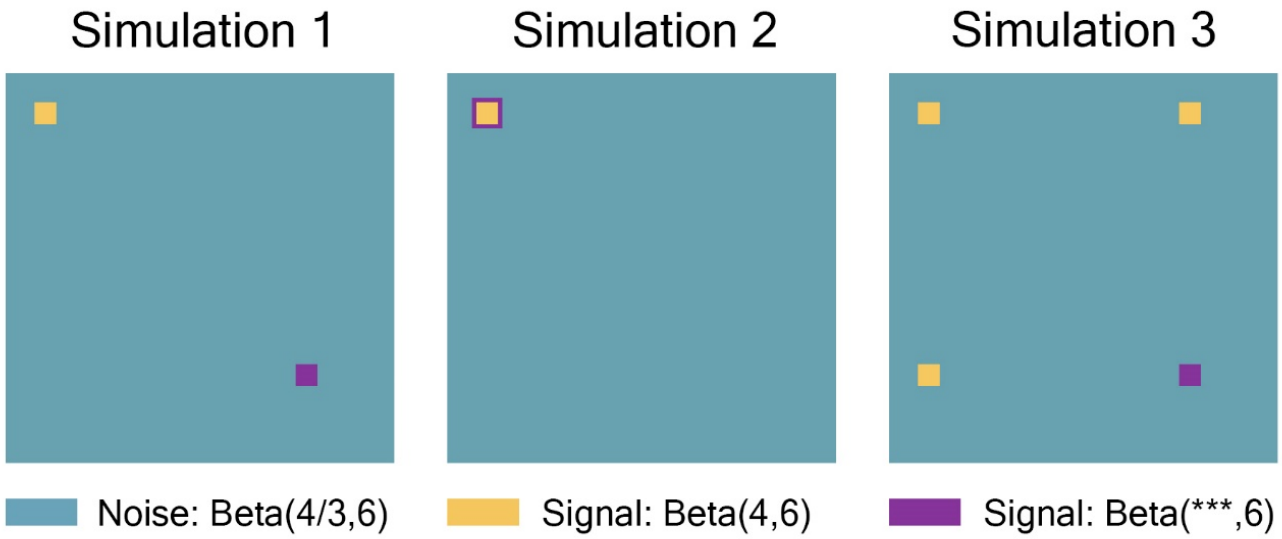

Figure 4: Simulation 1 had a 15x15 region where all individuals had the same signal distribution, a 15x15 region where each the signal distribution was dependent on covariates, and the rest of the 256x256 connection matrix was drawn from the noise distribution. Simulation 2 was the same as 1, but the signal distribution dependent on covariates was moved to expand the border of the 15x15 same signal distribution to a combined 21x21 signal region. Simulation 3 was the same as 1, with two additional 15x15 regions where all individuals had the same signal distribution. The plots were drawn to scale.

Represented by the same colors from Figure 4 simulated connectivity matrices, Figure 5 displays the distributions used for those matrices as signal percentage increased. The noise distribution was distributed Beta(4/3,6) and was shown in teal (not affected by signal percentage). The “same-for-all” signal region was distributed Beta(4,6) and is in yellow (not affected by signal percentage). The “covariate-dependent” signal region can be seen in purple and was distributed Beta(***,6). The *** parameter had some distribution based upon the underlying covariate distribution and signal percentage. There are five purple distributions in each plot, representing the 0.01, 0.25, 0.5, 0.75, and 0.99 quantiles from the *** distribution. Further, the “covariate-dependent” (purple) signal region’s distribution goes from being the same as the noise region’s distribution (at 0% signal) to more and more different than the noise region’s distribution as signal percentage increases. Note that the 0.01 quantile curve overlays the noise curve for all signal intensities.

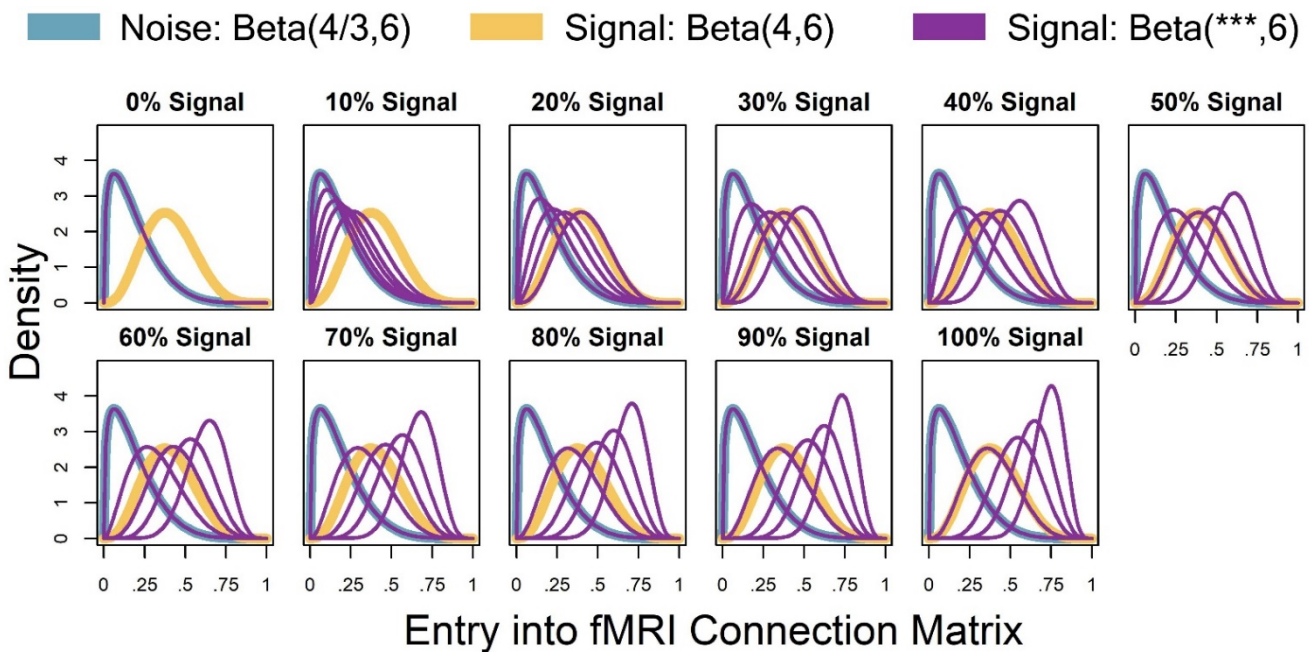

*Figure 5: Represented by the same colors from Figure 4 simulated connectivity matrices, Figure 5 displays the distributions used for those matrices as signal percentage increased. The noise distribution was distributed Beta(4/3,6) and was shown in teal (not affected by signal percentage). The "same-for-all" signal region was distributed Beta(4,6) and is in yellow (not affected by signal percentage). The "covariate-dependent" signal region can be seen in purple and was distributed Beta(***,6). The *** parameter had some distribution based upon the underlying covariate distribution and signal percentage. There are five purple distributions in each plot, representing the 0.01, 0.25, 0.5, 0.75, and 0.99 quantiles from the *** distribution. Further, the "covariate-dependent" (purple) signal region's distribution goes from being the same as the noise region's distribution (at 0% signal) to more and more different than the noise region's distribution as signal percentage increases. Note that the 0.01 quantile curve overlays the noise curve for all signal intensities.*

For simulation 4, we simulated nodal degree vectors of length 268 (instead of 268 x268 connectivity matrices as in simulations 1-3) to assess the method's ability to detect distributional differences rather than location differences. All entries of each individual's degree vector were simulated $\mathrm{Normal}(100 + s_p \cdot a_i, 1)$, where s_p (signal percent) and a_i (covariate-dependent parameter for individual i) were as described with simulations 1-3. Within Figure 6, there were five purple distributions in each plot, representing the 0.01, 0.25, 0.5, 0.75, and 0.99 quantiles from the distribution of the mean parameter, $100 + s_p \cdot a_i$. At 0% signal, all individual's nodal degree vectors were drawn from $Normal(100,1)$. As signal percentage increased, the mean parameter of the covariate-dependent distributions became more and more distinct.

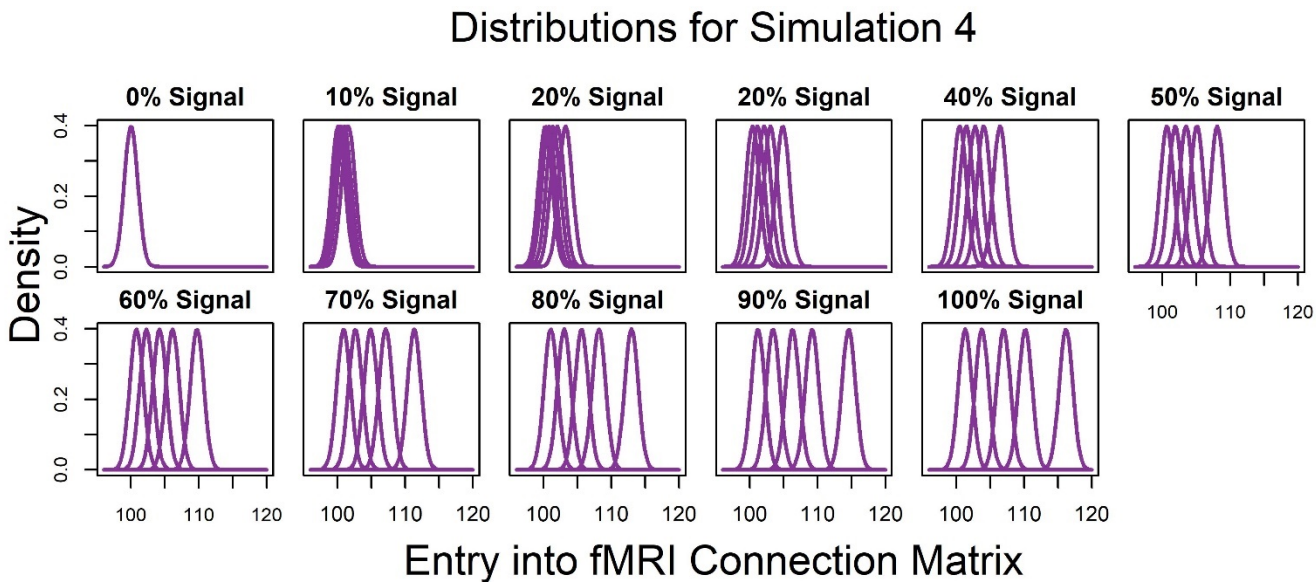

Figure 6: For simulation 4, we simulated nodal degree vectors of length 268 (instead of 268 x268 connectivity matrices as in simulations 1-3) to assess the method's ability to detect distributional differences rather than location differences. All entries of each individual's degree vector were simulated $Normal(100 + s_p \cdot a_i, 1)$, where s_p (signal percent) and a_i (covariate-dependent parameter for individual i) were as described with simulations 1-3. There were five purple distributions in each plot, representing the 0.01, 0.25, 0.5, 0.75, and 0.99 quantiles from the distribution of the mean parameter, $100 + s_p \cdot a_i$. At 0% signal, all individual's nodal degree vectors were drawn from $Normal(100,1)$. As signal percentage increased, the mean parameter of the covariate-dependent distributions became more and more distinct.

3.2. Results

We assessed methods with the four simulation scenarios detailed in the previous section. Each simulation was run 10,000 times. The percentage of p-values less than $\alpha = 0.05$ for the covariates of interest were recorded for each combination of signal percent (0%, 10%, …, 100%), distance metric (KS, Jaccard, Minkowski of order 0.5, Minkowski of order 1.4, Euclidean [Minkowski of order 2], Minkowski of order 3), and testing framework (F-test, Permutation, GLS, F-test with individual level effects, and Mixed Model). In this section, it was discussed whether type I error rate was controlled and at what signal percent the 80% power threshold was reached. For a full visual display of the results, see Figures 7-14.

The standard F-test did not control type I error when testing Age. It was included in the Figures for reference, but was not mentioned any further in this section.

3.2.1. KS

The Mixed Model approach did not control type I error in any simulation for IQ or Treatment (continuous or binary covariates of interest). Simulations 1-3: None of the methods that did control type I error (Permutation, GLS, and F-test with individual level effects) reached 80% power at any level of covariate-dependent signal. Simulation 4: The GLS and F-test with ILE methods reached the power threshold within 20% of signal. Permutation reached the threshold within 30% of signal. All three methods controlled type I error.

3.2.2. Jaccard

All methods (not including F-test) adequately controlled for type I error. Simulations 1-3, IQ: The GLS and F-test with ILE methods reached the power threshold within 20-30% of signal; the permutation method never reached the minimum power requirement. Simulations 1-3, Treatment: The GLS, F-test with ILE, and Permutation methods reached the power threshold within 20% of signal. **Surprisingly, as signal percentage increased, some situations had decreasing power with some even falling far back below the power threshold**. **This only occurred with the Jaccard distance metric.** Simulation 4: No methods could detect significance. All methods power levels hovered around $\alpha = 0.05$, the significance level.

3.2.3. Canberra and Minkowski

All remaining distance metrics were encapsulated in this section: Canberra and Minkowski of order 0.5, 1.4, 2 (Euclidean), and 3. All methods (not including F-test) adequately controlled for type I error. Simulations 1-3, IQ: The GLS, F-test with ILE, and Mixed Model methods reached the power threshold within 20% signal. The permutation method reached the minimum power requirement within 40-60% of signal. Simulations 1-3, Treatment: All methods reached the power threshold within 20% signal. Simulation 4, IQ: The GLS, F-test with ILE, and Mixed Model methods reached the power threshold within 50% signal. The permutation method reached the minimum power requirement within 80-90% of signal. Simulation 4, Treatment: All methods reached the power threshold within 40-50% signal.

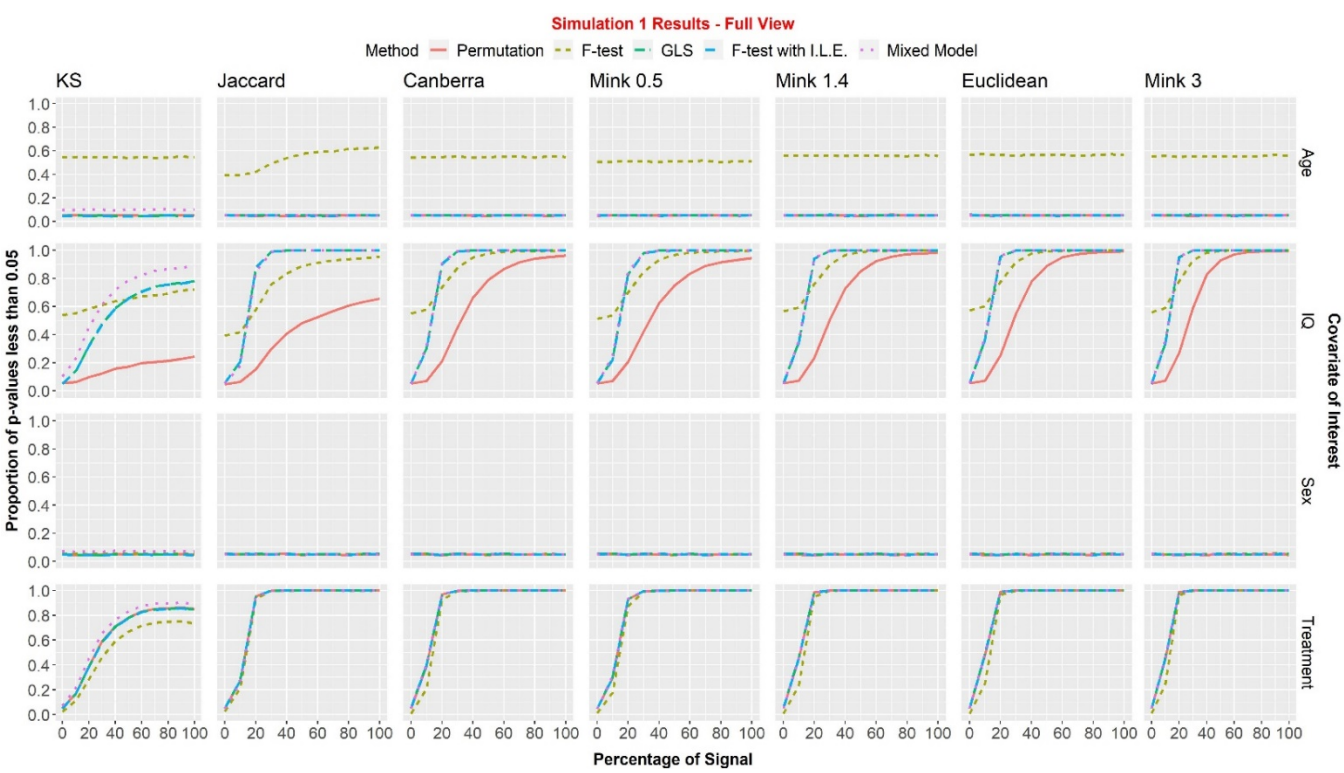

Figure 7: Simulation 1 Results - Full View

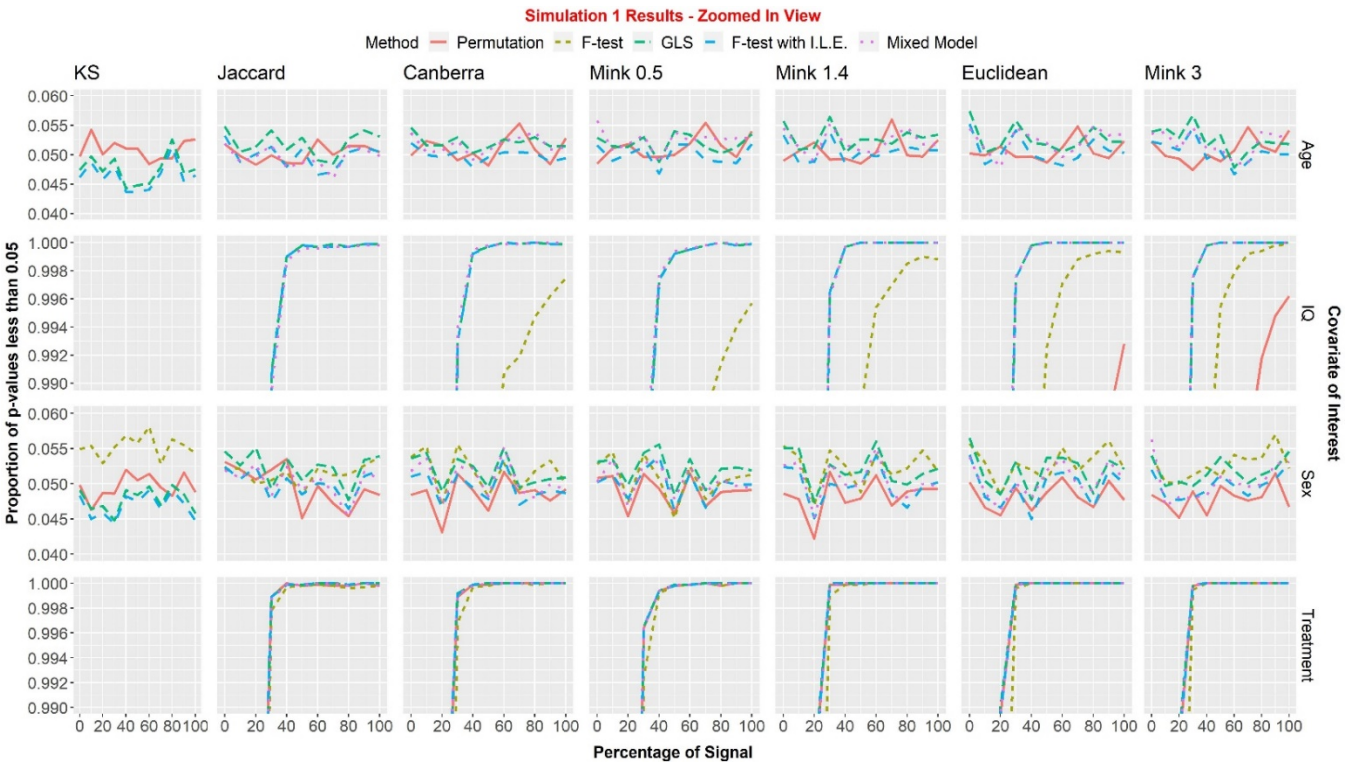

Figure 8: Simulation 1 Results - Zoomed In View

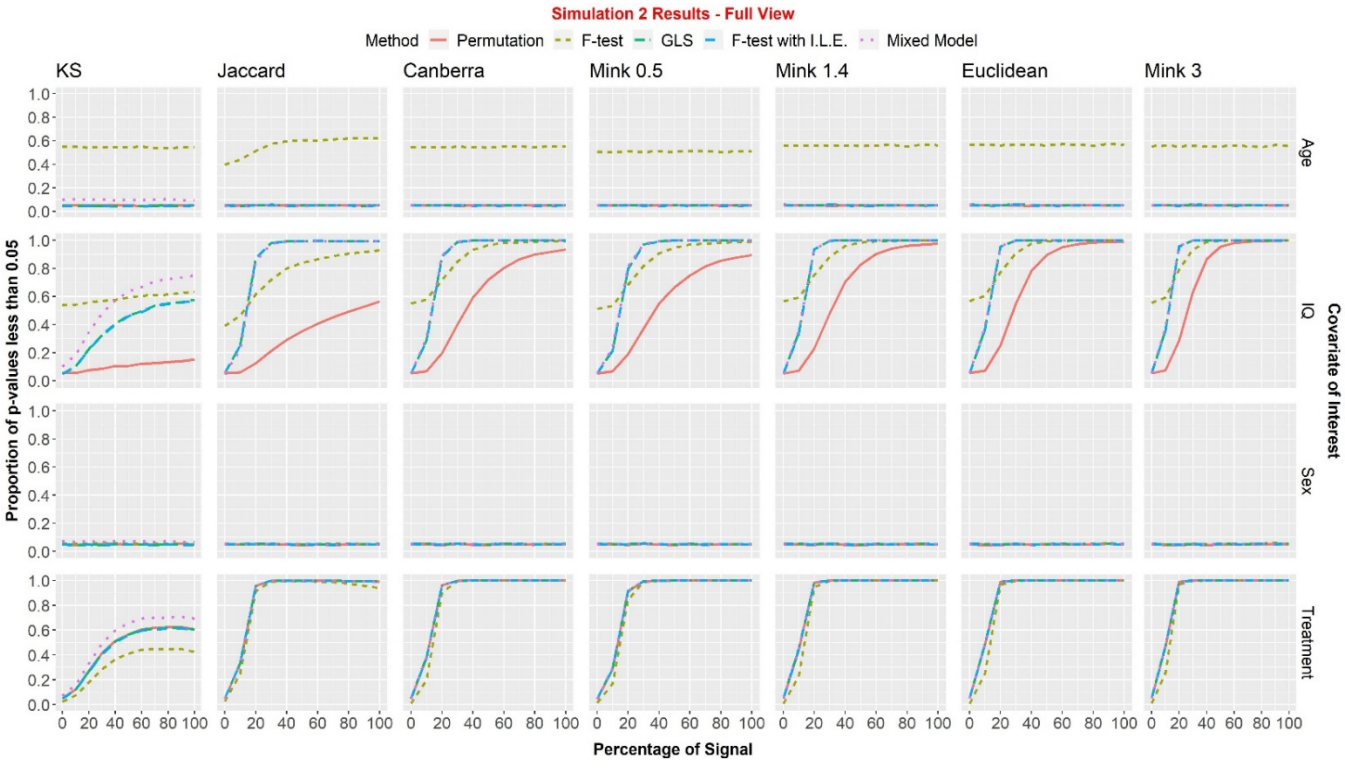

Figure 9: Simulation 2 Results - Full View

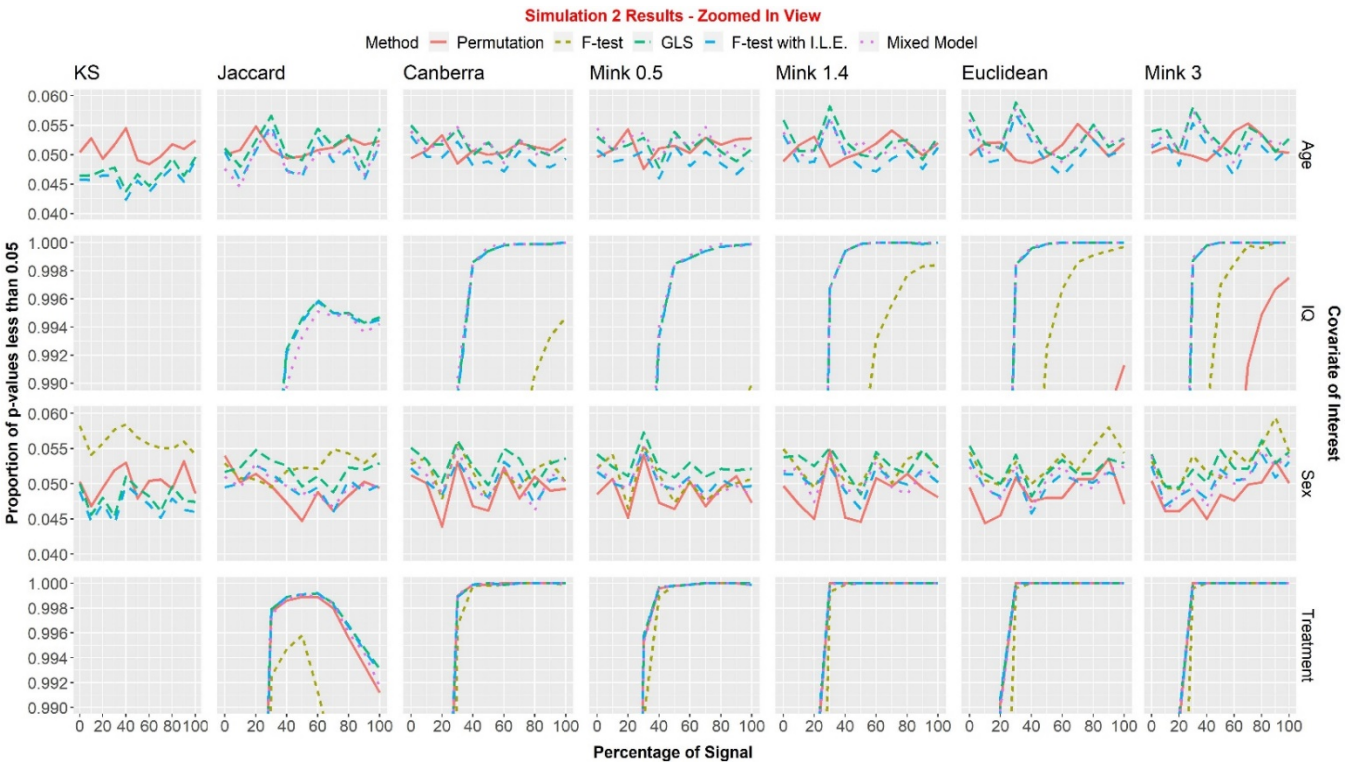

Figure 10: Simulation 2 Results - Zoomed In View

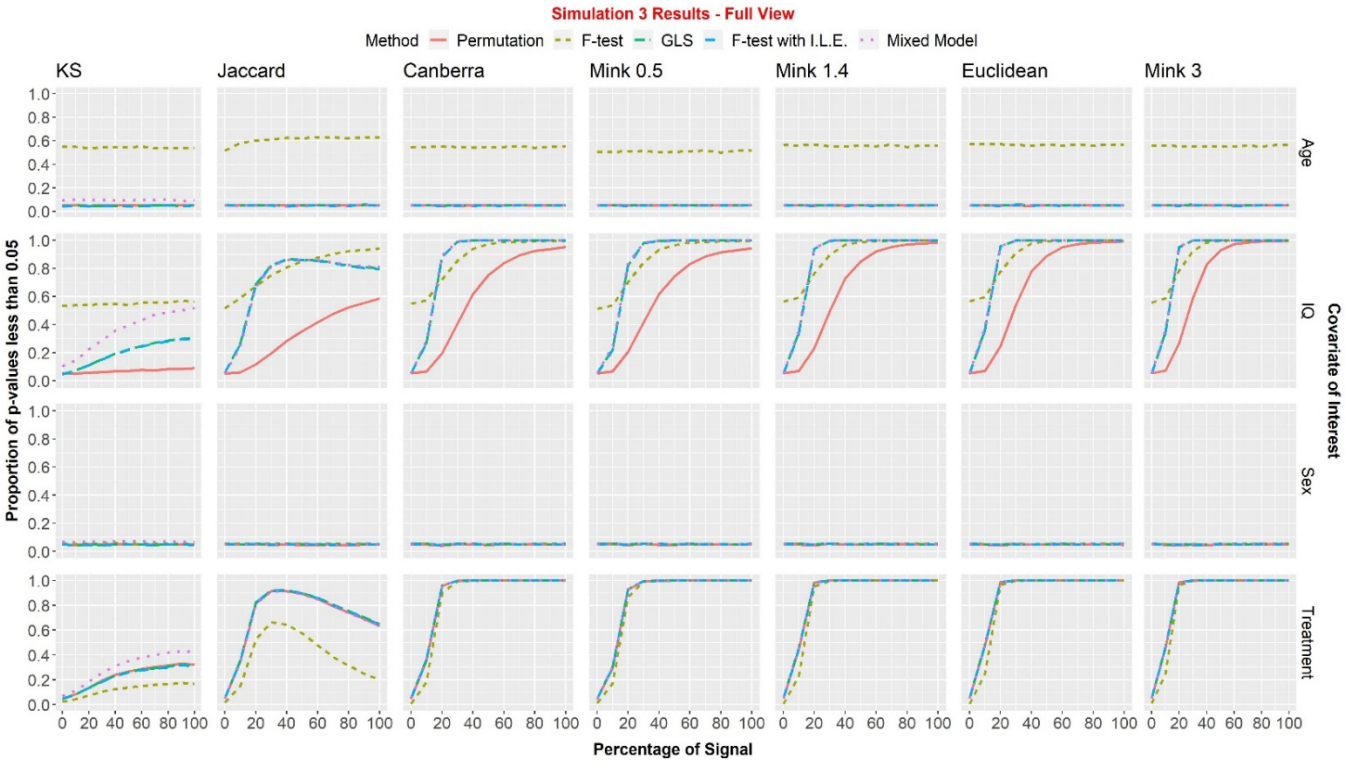

Figure 11: Simulation 3 Results - Full View

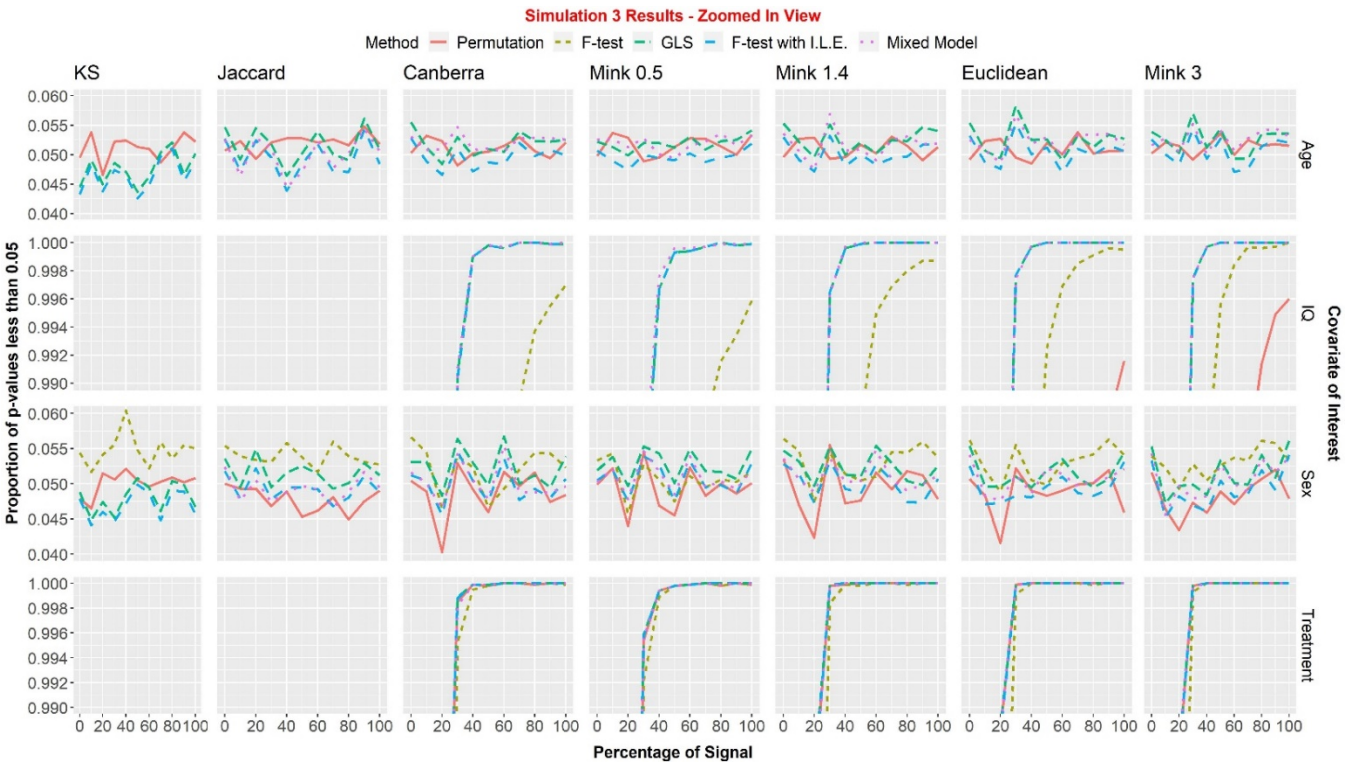

Figure 12: Simulation 3 Results - Zoomed In View

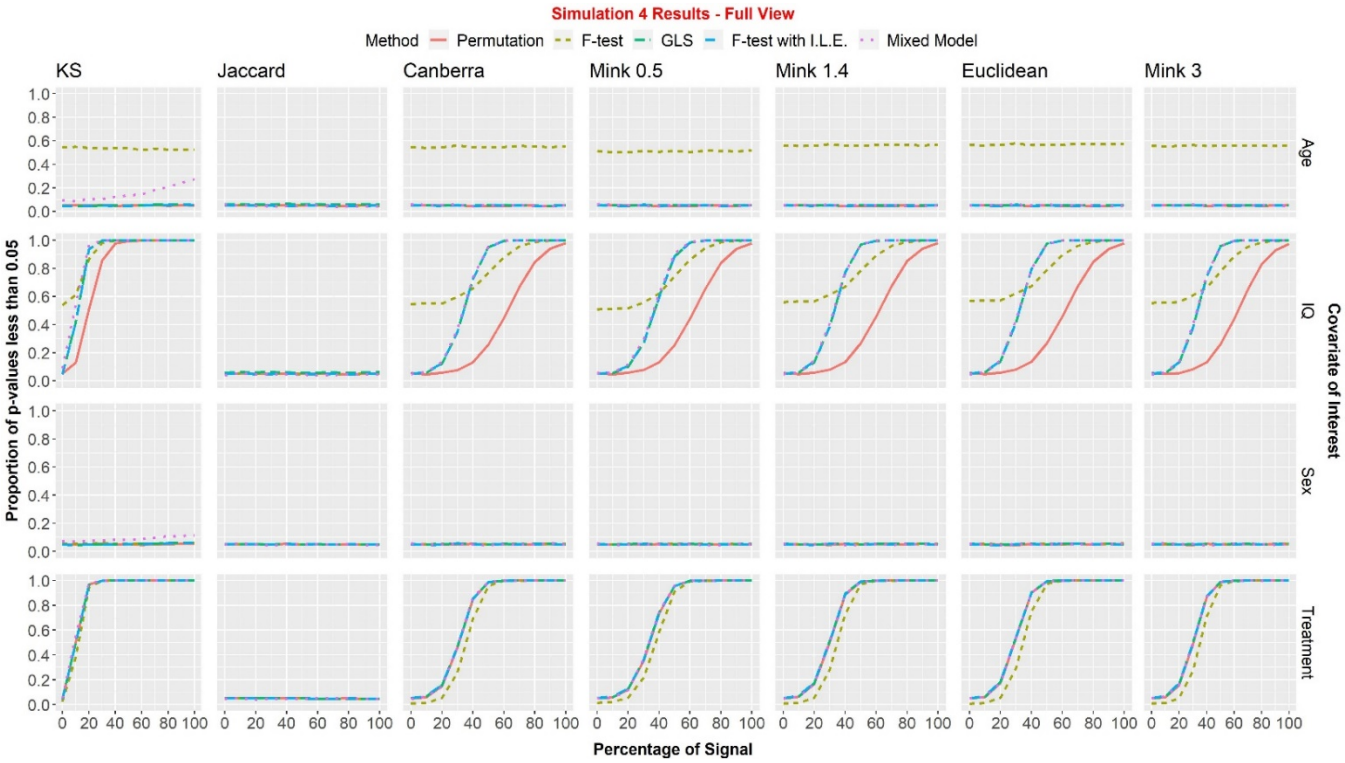

Figure 13: Simulation 4 Results - Full View

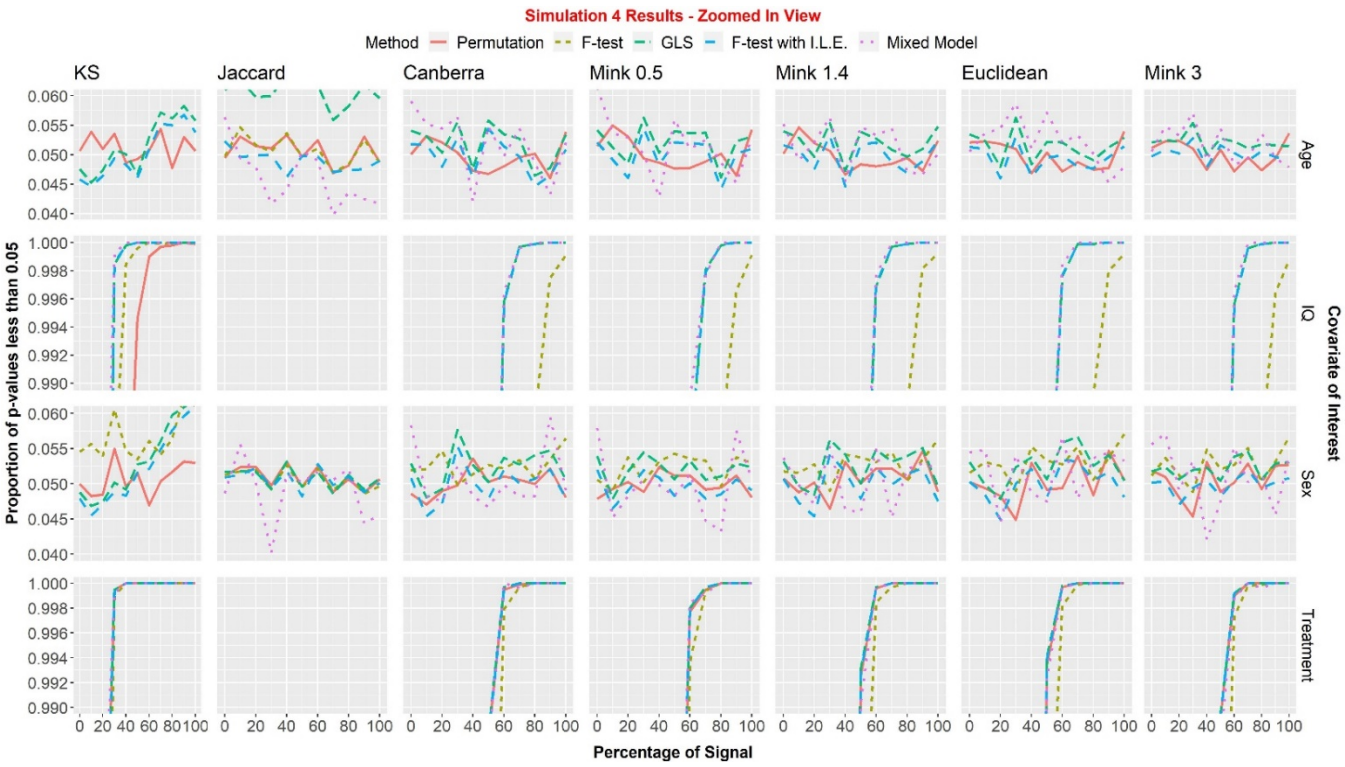

Figure 14: Simulation 4 Results - Zoomed In View

4. Experimental Studies

4.1. Data

The HCP data released to date include 1,200 individuals (Van Essen et al., 2013). Of those, 1,113 (606 females; 283 minority) have complete MRI images, cognitive testing, and detailed demographic information. The 397 subjects used are what remained after quality control assessment of head motion and global signal changes for both scan types, removal of those with missing data, and random selection of one individual from each family to ensure between-subject independence. Participants in the HCP completed two resting-state scans and two working memory scans. The two scans were collected with different phase encoding (right to left vs. left to right). The resting-state scans were collected back-to-back while participants quietly viewed a fixation point. The 2-back task was a block design that interleaved the 2-back condition with a 0-back condition and a rest period. The working memory task utilized photos and different blocks had different photo types (faces, body parts, houses, tools). Participants were alerted prior to each block to indicate the task type. For the 2-back they were instructed to respond anytime the current stimulus being presented matched the stimulus two trials back. The HCP performed extensive testing and development to ensure comparable imaging at the two sites (Van Essen et al., 2012). The blood oxygenation level dependent (BOLD)-weighted images were collected using the following parameters: TR = 720 ms, TE = 33.1 ms, voxel size 2 mm^3, 72 slices, 1,200 images.

The main covariate of interest for this analysis was Fluid Intelligence. Other covariates included in model formulation were Age, BMI, Education, Handedness, Income, Alcohol Abuse, Alcohol Dependence, Ethnicity, Gender, Race, and Smoking Status. For a summarization and explanation of the variables, see Tables *** and ***.

	Mean (S.D.)	Notes
Age	28.7 (3.7)	In Years.
BMI	26.5 (5.1)	Body Mass Index.
Education	15.0 (1.7)	Integer Values 11 to 17 (years pf education completed).
Fluid Intelligence	17.3 (4.7)	Integer valued from 4 to 24.
Handedness	65.1 (43.6)	Values range from -100 to 100 by 5 (-100, -95, …, 95, 100)
Income	5.0 (2.1)	SSAGA income score - Total household income: <$10,000 = 1, 10K-19,999 = 2, 20K-29,999 = 3, 30K-39,999 = 4, 40K-49,999 = 5, 50K-74,999 = 6, 75K-99,999 = 7, >=100,000 = 8.

Table 1: Summarization and explanation of HCP covariates treated as continuous (within the regression framework).

Alcohol Abuse	**54** met the DSM4 criteria for alcohol abuse, **343** did not
Alcohol Dep.	**28** met the DSM4 criteria for alcohol dependence, **369** did not
Ethnicity	**41** Hispanic/Latino, **351** Not Hispanic/Latino, **5** Unknown or Not Reported
Gender	**207** Female, **190** Male
Race	**1** Am. Indian/Alaskan Nat., **31** Asian/Nat. Hawaiian/Other Pacific Is., **48** Black or African Am., **13** More than one, **10** Unknown or Not Reported, **294** White
Smoking Status	**69** reported as still smoking, **328** did not

Table 2: Summarization and explanation of HCP covariates treated as categorical (within the regression framework).

4.2. Data Processing and Network Generation

The current project used the minimally processed fMRI data provided by the HCP (Glasser et al., 2013) for resting state and working memory. Additional preprocessing steps included: motion correction using ICA-AROMA (Pruim et al., 2015), removal of the first 14 volumes from each scan, and band-pass filtering (0.009–0.08 Hz) to remove physiological noise and low-frequency drift. It was necessary to account for the block design of the working memory task. First, the block design was modeled in SPM12, yielding regressors for the 0-back and rest blocks as well as the cues. Additionally, since each scan was collected twice with different phase encoding, the scans were concatenated and a scan-specific regressor was added. All regressors, including the total gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) signals and realignment parameters were used in a single regression analysis. The residual signal following regression of the extraneous variables was retained only for periods aligned with the 2-back blocks. The blocks of data were than concatenated into a single time series containing 274 BOLD images. Resting state data followed a similar pipeline, but without the task-specific regressors. Additionally, the first 14 volumes of each resting state scan were removed to ensure signal stabilization. The resulting resting state data contained 2,372 bold images. After preprocessing, the brain was parcellated into 268 regions as defined in the Shen Atlas (Shen et al., 2013) by averaging the signal from all voxels within each region for each participant. A functional network was constructed for each

participant by computing the Pearson (full) correlation between the resultant time series for each region pair. Negative correlation values were set to zero for reasons detailed in (Cao et al., 2014).

4.3. Results

Nodal degree vectors (used for the KS, Minkowski, and Euclidean distances) were created by summing across rows of the connectivity matrices. Key nodes of interest (binary degree vectors used for the Jaccard index) based on node degree were identified, selecting the top 20% highest degree (hub) nodes. KS statistic, Minkowski distance of order 1.4, and Euclidean distance (Minkowski of order 2) were calculated for each pair of individuals using their nodal degree vectors. The Jaccard distance was calculated for each pair of individuals using their binary degree vectors.

Distance covariates for each pair of individuals were calculated. A continuous variable's distance (Age, for instance) was calculated as $|Age_i - Age_j|$ for the pair of individuals i and j. A binary or categorical variable's distance (Education, for instance) was calculated as $\mathbb{1}\{Edu_i \neq Edu_j\}$ for the pair of individuals i and j.

We evaluated differences between networks using the Standard F-test with ILE. Parameter and standard error estimates can be found in Supplemental Materials in Tables S1 and S2. Each parameter estimate represented the average amount the given brain distance metric (KS, Jaccard, etc.) changed based on a one-unit difference in the respective covariate, after controlling for other covariates. A complete list of p-values for both resting state and working memory can be seen in Table 3.

	Resting State				Working Memory			
	KS	Jaccard	Mink 1.4	EUC	KS	Jaccard	Mink 1.4	EUC
Fluid Intelligence	1.30E-14	5.96E-03	3.24E-10	1.12E-09	1.74E-01	3.47E-02	4.91E-02	2.80E-02
Age	6.32E-05	6.35E-02	2.11E-05	9.09E-05	1.26E-02	9.68E-04	6.57E-09	9.41E-09
Alcohol Abuse	6.29E-05	4.35E-01	2.03E-02	1.10E-02	7.79E-01	2.02E-01	4.04E-01	4.45E-01
Alc. Dependence	2.93E-01	2.26E-01	2.19E-01	2.73E-01	9.34E-01	7.84E-02	1.55E-01	1.86E-01
BMI	1.03E-01	4.97E-03	6.53E-06	1.40E-05	1.32E-02	1.62E-01	5.05E-01	3.95E-01
Education	8.67E-02	3.10E-01	3.15E-01	4.06E-01	2.91E-02	2.93E-04	1.91E-02	2.03E-02
Ethnicity	6.14E-01	4.25E-01	4.30E-01	4.66E-01	7.10E-01	1.44E-01	3.94E-01	4.44E-01
Gender	6.28E-02	6.24E-16	2.98E-17	4.08E-16	1.36E-01	8.70E-10	1.88E-08	1.03E-08
Handedness	8.03E-01	2.43E-02	1.30E-01	1.13E-01	3.31E-01	2.30E-01	7.44E-01	7.14E-01
Income	2.77E-01	2.06E-02	2.68E-01	1.82E-01	5.90E-01	6.00E-01	3.88E-01	3.70E-01
Race	4.20E-06	2.14E-04	3.74E-14	3.79E-13	4.67E-01	1.34E-07	3.28E-03	6.05E-03
Smoking Status	2.05E-01	7.16E-01	2.67E-01	3.00E-01	8.54E-01	6.37E-01	8.99E-01	8.63E-01

Legend: 0 — 0.05 — 1

Table 3: P-values for HCP resting state and working memory brain scans when modeled with our given regression framework and tested using the standard F-test with fixed individual level effects. Parameter estimates and standard errors can be found in Supplementary Materials.

After adjusting for the other confounding variables, the covariate of interest, Fluid Intelligence, was statistically significant with all distance metrics (KS, Jaccard, Minkowski 1.4, Euclidean) for Resting State fMRI, and non-KS distance metrics for Working Memory fMRI. Figure 15a shows that the spatial distribution of brain hubs are relatively

concentrated in the brain regions making up the default mode network (known to have high degree at rest at rest). However, there is a higher concentration of hubs localized to the default mode network in the bottom quartile of fluid intelligence compared to those in the top quartile. Figure 15b shows that hub locations shift from default mode areas to regions that make up the central executive attention network more so in the top quartile of fluid intelligence. Note that in those individuals in the bottom quartile of intelligence, the hubs are most prominent in the areas of the default mode network even during the working memory task, though the magnitude of the overlap is lower.

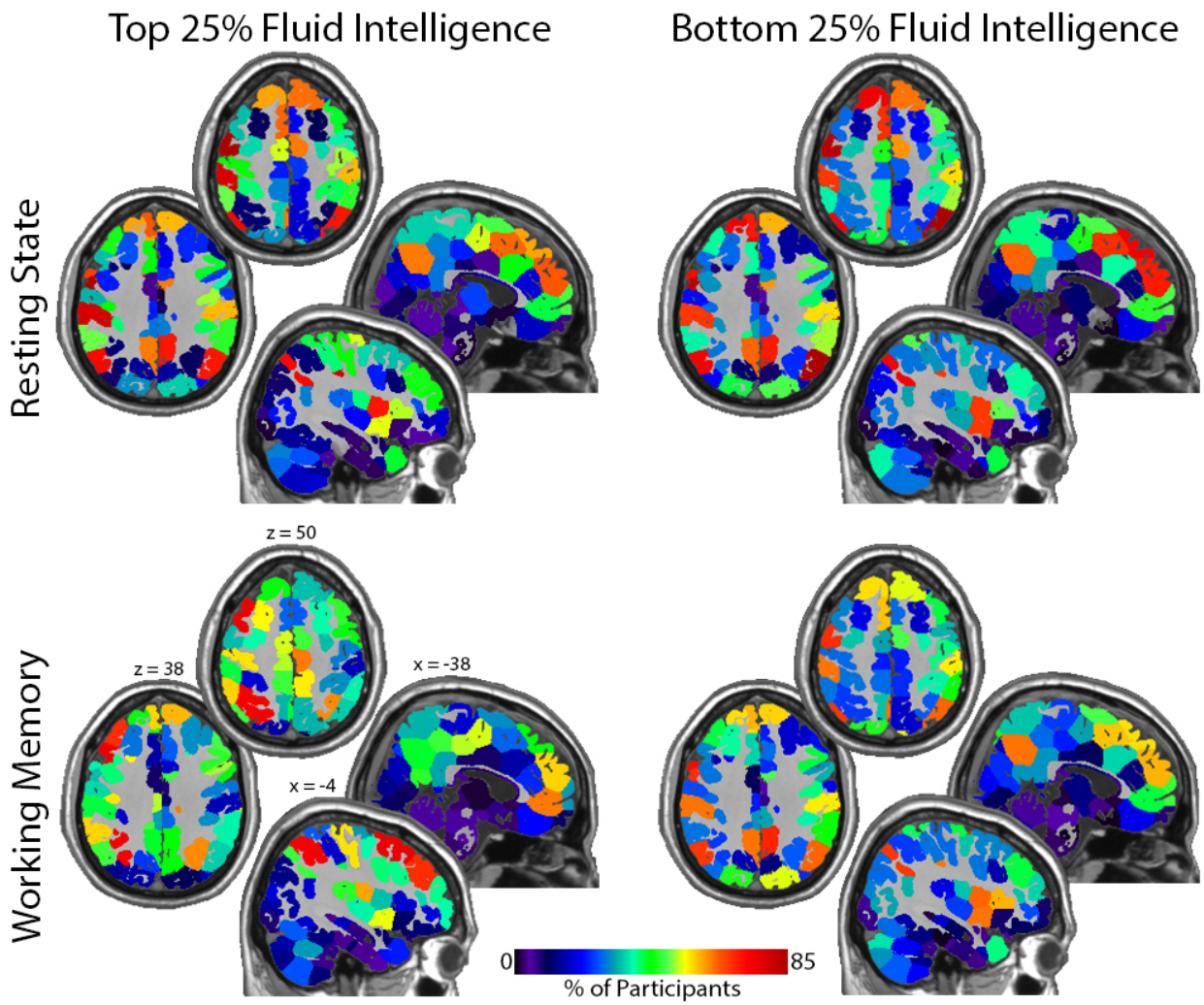

Figure 15: Maps showing the location of network hubs (top 20% degree) for the top and bottom intelligence quartiles. Resting State: note the high incidence of hubs in the medial and lateral parietal cortex and the medial frontal lobe in both groups. These regions are central components of the default mode network. The significant association between degree location and intelligence is demonstrated by the reduction in default mode network hubs with increases in fluid intelligence. Working Memory: Hubs are more concentrated in the central executive network as intelligence increases. This shift is best appreciated in the top quartile group where hubs are concentrated in lateral rather than medial frontal regions and in more superior parietal areas. Note that the hubs remain fairly localized to the default mode network for the bottom quartile group. Each quadrant shows 2 axial and 2 sagittal images. The Montreal neurological Institute (MNI) coordinates shown in the bottom left quadrant apply to all quadrants. Calibration bar applies to all images.

Among the confounding variables, there were a couple that had significant relationships with all distance metrics except for the Jaccard Index (Age and Alcohol Abuse within Resting State). Ideally, we would expect Jaccard to be statistically

significant if Minkowski norms were heavily significant. This could have been due to the Jaccard-related issues that showed up in our simulations. There were also certain variables that were significant for one fMRI task but not the other (Alcohol Abuse, BMI and Education). Some covariates were significant for all location-specific metrics but not for the distance between distribution comparator KS statistic (BMI – Rest; Gender – Rest and WM; Fluid Intelligence and Race–WM). This was indicative of a detectable difference in location-specific degree, but not in distribution.

The difference between Minkowski of order 1.4 and order 2 (Euclidean) was very slight. This was more evidence of the choice between location-specific, degree-weighted distance metrics not being clear.

5. Discussion

It is of great interest to compare differences in fMRI connection matrices between individuals by covariates. Our previous work developed a novel permutation testing framework to detect differences between two groups (Simpson et al., 2013b). Here we advanced that work to allow both assessing differences by continuous phenotypes and controlling for confounding variables. We proposed an innovative regression framework to relate distances between brain network features to functions of absolute differences in continuous covariates and indicators of difference for categorical variables, and explored several similarity metrics for comparing distances between connection matrices. The KS statistic measures how different distributions of topological properties vary between two individuals. Key-node metrics (like the Jaccard Index) quantify how much the spatial location of key brain network regions differs between two networks. The Minkowski norm (along with Canberra, Weighted Jaccard, etc.) measures whether the spatial location of degree-weighted brain network regions differ. Several additional similarity/dissimilarity metrics were also assessed, but left out for the sake of brevity. The binary network-based metrics left out included the Russel-Rao, Kulczynski, and Baroni-Urbani and Buser. Little difference or improvement was found (in simulations 1-3) for any of these metrics when compared with the commonly used Jaccard Index. The weighted network-based metrics left out included the Weighted Jaccard, Manhattan, and Maximum distances, with none different or better than the already presented Canberra or Minkowski distances (when tested in simulations 1-3).

Several standard methods for estimation and inference were adapted to fit into our regression framework: Standard F-test, F-test with individual level effects (ILE), Generalized Least Squares (GLS), Mixed Model, and Permutation (inference only). All combinations of these approaches and the distance metrics were assessed via four simulation scenarios. The KS statistic was found to have low comparable power when testing location-specific differences. However, if interested in comparing distributions, the KS statistic was preferred. The Jaccard Index did not have consistent or predictable power, and further work should investigate the reasons underlying this. An argument could be made for several different distance metrics when detecting location-specific differences between degree vectors, as they had very similar results. We prefer to use Euclidean distance, as it is commonly recognized and had the best overall (slightly) type I control and power. Future work could include a further investigation into how and when to choose specific distance metrics. Regarding the comparison of estimation and testing methods (Standard F-test, F-test with ILE, GLS, Mixed

Model, Permutation), we recommend the F-test with ILE as it was among the best at controlling type I error and providing sufficient power while also being computationally inexpensive.

A novel modification of the standard GLS approach (with the novelty being the estimation approach taken for the variance-covariance matrix) for estimation and inference was proposed in this framework. Although it is slower (computationally) than the Standard F-test with ILE, the GLS approach performed just as well as the recommended Standard F-test with ILE (in terms of type I and type II error).

An analysis of the HCP data was completed using the Standard F-test with ILE and several distance metrics. After adjusting for the other confounding variables, the covariate of interest, Fluid Intelligence, was significantly related with all distance metrics (KS, Jaccard, Minkowski 1.4, Euclidean) for both fMRI tasks (Resting State, Working Memory). This analysis suggested the hubs become more strongly concentrated in default mode network regions at rest with decreasing fluid intelligence. As intelligence increases, there are greater shifts in the spatial locations of hubs from the default mode network at rest to the central executive network during a working memory task.

We have developed a testing framework that detects whether the spatial location of key brain network regions and distributions of topological properties differ by phenotype (continuous and discrete) after controlling for confounding variables in static networks. More generally, this framework allows relating distances between networks (e.g., Jaccard, K-S distance) to covariates of interest. Future work extend this approach by creating a dynamic network analog that uncovers whether within- and across-task time-varying changes in these spatial and distributional patterns differ by phenotype.

Appendix. Supplementary materials

Tables S1 and S2 contain the parameter estimates and standard errors from the HCP analysis described in the main text of the paper.

	Resting State							
	KS		Jaccard		Mink 1.4		Euclidean	
	Est.	Std. Err.	Est.	Std. Err.	Est.	Std. Err.	Est.	Std. Err.
Fluid Intelligence	6.60E-04	8.56E-05	-9.72E-05	3.53E-05	2.83E-01	4.51E-02	9.09E-02	1.49E-02
Age	4.24E-04	1.06E-04	-8.13E-05	4.38E-05	2.38E-01	5.59E-02	7.24E-02	1.85E-02
Alcohol Abuse	4.71E-03	1.18E-03	3.79E-04	4.86E-04	1.44E+00	6.20E-01	5.22E-01	2.05E-01
Alc. Dependence	-2.23E-03	2.12E-03	1.06E-03	8.77E-04	-1.37E+00	1.12E+00	-4.06E-01	3.71E-01
BMI	1.40E-04	8.56E-05	-9.93E-05	3.53E-05	2.03E-01	4.51E-02	6.49E-02	1.49E-02
Education	-3.76E-04	2.20E-04	-9.22E-05	9.07E-05	-1.16E-01	1.16E-01	-3.19E-02	3.83E-02
Ethnicity	-7.36E-04	1.46E-03	-4.80E-04	6.02E-04	6.06E-01	7.68E-01	1.85E-01	2.54E-01
Gender	1.03E-03	5.51E-04	-1.84E-03	2.28E-04	2.45E+00	2.90E-01	7.82E-01	9.61E-02
Handedness	-2.87E-06	1.15E-05	-1.07E-05	4.73E-06	9.14E-03	6.04E-03	3.17E-03	2.00E-03
Income	-1.97E-04	1.81E-04	-1.73E-04	7.49E-05	1.06E-01	9.55E-02	4.23E-02	3.16E-02
Race	4.52E-03	9.82E-04	-1.50E-03	4.06E-04	3.92E+00	5.17E-01	1.24E+00	1.71E-01
Smoking Status	1.22E-03	9.64E-04	-1.45E-04	3.98E-04	5.63E-01	5.08E-01	1.74E-01	1.68E-01
Participant 1	4.82E-02	3.88E-03	4.04E-01	1.60E-03	1.36E+02	2.04E+00	4.67E+01	6.77E-01
Participant 2	6.62E-02	3.88E-03	4.09E-01	1.60E-03	1.28E+02	2.05E+00	4.40E+01	6.77E-01
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Participant 397	7.57E-02	3.96E-03	3.59E-01	1.64E-03	1.43E+02	2.09E+00	4.86E+01	6.91E-01

Table S1: Parameters and Standard Errors for HCP resting state brain scans (nodal degree vectors) when modeled with our given regression framework and tested using the standard F-test with fixed individual level effects.

	Working Memory							
	KS		Jaccard		Mink 1.4		Euclidean	
	Est.	Std. Err.	Est.	Std. Err.	Est.	Std. Err.	Est.	Std. Err.
Fluid Intelligence	1.09E-04	8.05E-05	-5.93E-05	2.81E-05	7.08E-02	3.60E-02	2.69E-02	1.22E-02
Age	2.49E-04	9.97E-05	-1.15E-04	3.48E-05	2.59E-01	4.46E-02	8.72E-02	1.52E-02
Alcohol Abuse	-3.10E-04	1.11E-03	4.92E-04	3.86E-04	-4.12E-01	4.94E-01	-1.29E-01	1.68E-01
Alc. Dependence	-1.65E-04	2.00E-03	1.23E-03	6.97E-04	-1.27E+00	8.93E-01	-4.02E-01	3.04E-01
BMI	2.00E-04	8.05E-05	3.93E-05	2.81E-05	2.40E-02	3.60E-02	1.04E-02	1.23E-02
Education	4.51E-04	2.07E-04	-2.61E-04	7.21E-05	2.16E-01	9.23E-02	7.30E-02	3.15E-02
Ethnicity	5.10E-04	1.37E-03	6.98E-04	4.78E-04	-5.22E-01	6.12E-01	-1.60E-01	2.09E-01
Gender	7.72E-04	5.18E-04	-1.11E-03	1.81E-04	1.30E+00	2.32E-01	4.52E-01	7.89E-02
Handedness	-1.05E-05	1.08E-05	-4.51E-06	3.76E-06	1.58E-03	4.82E-03	6.00E-04	1.64E-03
Income	9.19E-05	1.71E-04	-3.12E-05	5.95E-05	6.58E-02	7.62E-02	2.33E-02	2.60E-02
Race	6.72E-04	9.24E-04	-1.70E-03	3.22E-04	1.21E+00	4.13E-01	3.86E-01	1.41E-01
Smoking Status	1.67E-04	9.06E-04	1.49E-04	3.16E-04	5.12E-02	4.05E-01	-2.39E-02	1.38E-01
Participant 1	7.66E-02	3.65E-03	3.74E-01	1.27E-03	1.57E+02	1.63E+00	5.34E+01	5.55E-01
Participant 2	6.58E-02	3.65E-03	3.69E-01	1.27E-03	2.03E+02	1.63E+00	7.01E+01	5.56E-01
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Participant 397	2.21E-01	3.73E-03	3.78E-01	1.30E-03	4.14E+02	1.67E+00	1.43E+02	5.67E-01

Table S2: Parameters and Standard Errors for HCP working memory brain scans (nodal degree vectors) when modeled with our given regression framework and tested using the standard F-test with fixed individual level effects.

Acknowledgements

This work was supported by National Institute of Biomedical Imaging and Bioengineering (NIBIB) R01EB024559, and Wake Forest Clinical and Translational Science Institute (WF CTSI) NCATS UL1TR001420.

References

Aitken, A.C., 1936. On Least-squares and Linear Combinations of Observations. Proc. R. Soc. Edinburgh 55, 42–48.

Bassett, D.S., Bullmore, E.T., 2009. Human brain networks in health and disease. Curr. Opin. Neurol. https://doi.org/10.1097/WCO.0b013e32832d93dd

Burdette, J.H., Laurienti, P.J., Espeland, M.A., Morgan, A., Telesford, Q., Vechlekar, C.D., Hayasaka, S., Jennings, J.M., Katula, J.A., Kraft, R.A., Rejeski, W.J., 2010. Using network science to evaluate exercise-associated brain changes in older adults. Front. Aging Neurosci. 2. https://doi.org/10.3389/fnagi.2010.00023

Cao, M., Wang, J.H., Dai, Z.J., Cao, X.Y., Jiang, L.L., Fan, F.M., Song, X.W., Xia, M.R., Shu, N., Dong, Q., Milham, M.P., Castellanos, F.X., Zuo, X.N., He, Y., 2014. Topological organization of the human brain functional connectome across the lifespan. Dev. Cogn. Neurosci. 7, 76–93. https://doi.org/10.1016/j.dcn.2013.11.004

Deuker, L., Bullmore, E.T., Smith, M., Christensen, S., Nathan, P.J., Rockstroh, B., Bassett, D.S., 2009. Reproducibility of graph metrics of human brain functional networks. Neuroimage 47, 1460–1468. https://doi.org/10.1016/j.neuroimage.2009.05.035

Fornito, A., Zalesky, A., Pantelis, C., Bullmore, E.T., 2012. Schizophrenia, neuroimaging and connectomics. Neuroimage. https://doi.org/10.1016/j.neuroimage.2011.12.090

Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., Van Essen, D.C., Jenkinson, M., 2013. The minimal preprocessing pipelines for the Human Connectome Project. Neuroimage 80, 105–124. https://doi.org/10.1016/j.neuroimage.2013.04.127

Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C.J., Van Wedeen, J., Sporns, O., 2008. Mapping the structural core of human cerebral cortex. PLoS Biol. 6, 1479–1493. https://doi.org/10.1371/journal.pbio.0060159

Hayasaka, S., Laurienti, P.J., 2010. Comparison of characteristics between region-and voxel-based network analyses in resting-state fMRI data. Neuroimage 50, 499–508. https://doi.org/10.1016/j.neuroimage.2009.12.051

Joyce, K.E., Laurienti, P.J., Burdette, J.H., Hayasaka, S., 2010. A new measure of centrality for brain networks. PLoS One 5. https://doi.org/10.1371/journal.pone.0012200

Kherad-Pajouh, S., Renaud, O., 2014. A general permutation approach for analyzing repeated measures ANOVA and mixed-model designs. Stat. Pap. 56, 947–967. https://doi.org/10.1007/s00362-014-0617-3

Kherad-Pajouh, S., Renaud, O., 2010. An exact permutation method for testing any effect in balanced and unbalanced fixed effect ANOVA. Comput. Stat. Data Anal. 54, 1881–1893. https://doi.org/10.1016/j.csda.2010.02.015

Kolmogorov, A., 1933. Sulla determinazione empirica di una lgge di distribuzione. Inst. Ital. Attuari, Giorn. 4, 83–91.

Lance, G.N., Williams, W.T., 1966. Computer Programs for Hierarchical Polythetic Classification (“Similarity Analyses”).

Comput. J. 9, 60–64. https://doi.org/10.1093/comjnl/9.1.60

Liu, Y., Liang, M., Zhou, Y., He, Y., Hao, Y., Song, M., Yu, C., Liu, H., Liu, Z., Jiang, T., 2008. Disrupted small-world networks in schizophrenia. Brain 131, 945–961. https://doi.org/10.1093/brain/awn018

Meunier, D., Achard, S., Morcom, A., Bullmore, E., 2009a. Age-related changes in modular organization of human brain functional networks. Neuroimage 44, 715–723. https://doi.org/10.1016/j.neuroimage.2008.09.062

Meunier, D., Lambiotte, R., Fornito, A., Ersche, K.D., Bullmore, E.T., 2009b. Hierarchical modularity in human brain functional networks. Front. Neuroinform. 3. https://doi.org/10.3389/neuro.11.037.2009

Moussa, M.N., Steen, M.R., Laurienti, P.J., Hayasaka, S., 2012. Consistency of Network Modules in Resting-State fMRI Connectome Data. PLoS One 7. https://doi.org/10.1371/journal.pone.0044428

Moussa, M.N., Vechlekar, C.D., Burdette, J.H., Steen, M.R., Hugenschmidt, C.E., Laurienti, P.J., 2011. Changes in cognitive state alter human functional brain networks. Front. Hum. Neurosci. https://doi.org/10.3389/fnhum.2011.00083

Pruim, R.H.R., Mennes, M., Van Rooij, D., Llera, A., Buitelaar, J.K., Beckmann, C.F., 2015. ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. https://doi.org/10.1016/j.neuroimage.2015.02.064

Rombouts, S.A.R.B., Barkhof, F., Goekoop, R., Stam, C.J., Scheltens, P., 2005. Altered resting state networks in mild cognitive impairment and mild Alzheimer's disease: An fMRI study. Hum. Brain Mapp. 26, 231–239. https://doi.org/10.1002/hbm.20160

Rzucidlo, J.K., Roseman, P.L., Laurienti, P.J., Dagenbach, D., 2013. Stability of Whole Brain and Regional Network Topology within and between Resting and Cognitive States. PLoS One 8. https://doi.org/10.1371/journal.pone.0070275

Shen, X., Tokoglu, F., Papademetris, X., Constable, R.T., 2013. Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. Neuroimage 82, 403–415. https://doi.org/10.1016/j.neuroimage.2013.05.081

Simpson, S.L., Bowman, F.D.B., Laurienti, P.J., 2013a. Analyzing complex functional brain networks: Fusing statistics and network science to understand the brain. Stat. Surv. 7, 1–36. https://doi.org/10.1214/13-SS103

Simpson, S.L., Lyday, R.G., Hayasaka, S., Marsh, A.P., Laurienti, P.J., 2013b. A permutation testing framework to compare groups of brain networks. Front. Comput. Neurosci. 7, 1–13. https://doi.org/10.3389/fncom.2013.00171

Smirnov, N., 1948. Table for Estimating the Goodness of Fit of Empirical Distributions. Ann. Math. Stat. 19, 279–281. https://doi.org/10.1214/aoms/1177730256

Stam, C.J., Jones, B.F., Nolte, G., Breakspear, M., Scheltens, P., 2007. Small-world networks and functional connectivity in Alzheimer's disease. Cereb. Cortex 17, 92–99. https://doi.org/10.1093/cercor/bhj127

van den Heuvel, M.P., Stam, C.J., Boersma, M., Hulshoff Pol, H.E., 2008. Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain. Neuroimage 43, 528–539. https://doi.org/10.1016/j.neuroimage.2008.08.010

Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E.J., Yacoub, E., Ugurbil, K., 2013. The WU-Minn Human Connectome Project: An overview. Neuroimage 80, 62–79. https://doi.org/10.1016/j.neuroimage.2013.05.041

Van Essen, D.C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T.E.J., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtiss, S.W., Della Penna, S., Feinberg, D., Glasser, M.F., Harel, N., Heath, A.C., Larson-Prior, L., Marcus, D., Michalareas, G., Moeller, S., Oostenveld, R., Petersen, S.E., Prior, F., Schlaggar, B.L., Smith, S.M., Snyder, A.Z., Xu, J., Yacoub, E., 2012. The Human Connectome Project: A data acquisition perspective. Neuroimage. https://doi.org/10.1016/j.neuroimage.2012.02.018

Yuan, K., Qin, W., Liu, J., Guo, Q., Dong, M., Sun, J., Zhang, Y., Liu, P., Wang, W., Wang, Y., Li, Q., Yang, W., von Deneen, K.M., Gold, M.S., Liu, Y., Tian, J., 2010. Altered small-world brain functional networks and duration of heroin use in male abstinent heroin-dependent individuals. Neurosci. Lett. 477, 37–42. https://doi.org/10.1016/j.neulet.2010.04.032