# 1.2. Appendix A: Methods

## 1.2.1. Inclusion and Exclusion

Our research acquired data from the original University of California Los Angles Consortium for Neuropsychiatric Phenomics LA5c Study (Bilder et al., 2017). Bilder et al. (2017) applied numerous selection criteria for the participant to be included in the study: The selection criteria were that the participants: (a) were right-handed, (b) had a minimum of 8 years of formal education, (c) used English or Spanish as their primary language, (d) had no contraindications to undergo an MRI scan (e.g., pregnancy, claustrophobia, metal implants), (e) had no significant medical illness, (f) did not have visual acuity poorer than 20/60, and (g) no positive urinalysis of cocaine, methamphetamine, morphine, tetrahydrocannabinoids, or benzodiazepines (Poldrack et al., 2016). All participants were assessed using the Structured Clinical Interview for DSM-IV (First & Gibbon, 2004). Bilder et al. (2017) excluded HC participants if they had: (a) a lifetime diagnosis of any psychotic, bipolar, major depressive, anxiety (panic, generalized anxiety disorder), post-traumatic stress, obsessive compulsive, or attention deficit hyperactivity disorder (b) suicidality, or (c) substance abuse or dependence, excluding caffeine and nicotine dependence. Bilder et al. (2017) excluded PDS participants if they had a comorbid diagnosis of attention deficit hyperactivity disorder or bipolar disorder. According to Poldrack et al. (2016) stable medications were permitted for the patients, however PDS did not have to be medicated with antipsychotics to participate in the study. Following the selection criteria, Bilder et al. (2017) recruited 50 PDS and 130 HC.

## 1.2.2. Participant Characteristics

Ninety participants remained following the initial screening and case-control methods (described in the sampling procedure section of this paper): 45 PDS and 45 HC. The case-control method allowed us to minimize differences in sociodemographic and clinical variables between PDS and HC. The descriptive statistics for the demographics of the case-control sample are found in Table 1 and Table 2. The descriptive statistics suggest PDS were older than HC, mostly male, had spent less time in education, and were mostly current smokers as opposed to never smoking cigarettes daily.

As in Table 2, Bilder et al. (2017) administered several clinical assessments to the PDS treatment group. On average PDS had mean scores between no symptoms (a score of 0) and questionable symptoms (a score of 1) on the scale for the assessment of positive symptoms (Andreasen, 1984), between no symptoms (a score of zero) and mild symptoms (a score of 2) on the scale for the assessment of negative symptoms (Andreasen, 1989), and very mild symptoms (a score of 1) to mild symptoms (a score of 2) on the brief psychotic rating scale (Overall & Gorham, 1988). PDS presented with very mild to mild depression and anxiety symptoms. The results suggest that this group of PDS presented with mild severity, although there was considerable heterogeneity within the PDS treatment group.

## 1.2.3. Sampling Procedures

Bilder et al. (2017) recruited HC and PDS using convenience sampling. Community advertisements for HC were presented in the Los Angeles area and participants could nominate themselves to participate. For PDS, Bilder et al. (2017) used a patient outreach method. Participants could enlist through their community clinics or online portals (Bilder et al., 2017; Poldrack et al., 2016).

Our study only included PDS participants from the original study who completed the paired associate memory task (45 PDS). Further, the Bilder et al. (2017) study included a larger sample of HC than that of PDS. Because the sample size imbalance can affect the complexity and bias of an inferred dynamic Bayesian network (Spirtes & Meek, 1995), matching the number of HC to PDS on key variables was undertaken. The matching procedure involved selecting both confounding variables we identified (as in the Measures and Covariates section), and all demographic variables collected by Bilder et al. (2017). Demographic variables were included in case-control methods because discrete variables were important to discriminate between PDS and HC.

A logistic regression of treatment group with forwards stepwise selection, using AIC as the criterion, was performed to select variables and estimate weights in the case-control method. These weights were derived from the coefficients in a logistic regression. The case-control method used in this study was the greedy nearest neighbour algorithm (Hansen, 2004) which finds a greedy match of HC to PDS. The list of all variables can be found in the phenotype files uploaded to Openneuro (see Supplementary Information: Notes section).

## 1.2.4. Measures and Covariates

In total, there were 2025 unique phenotype variables collected in the original study, excluding meta data. A systematic method was used to select variables to be included in the dynamic Bayesian networks. This method was obtained following a review by an expert panel of the authors. The following selection criteria were used to select variables: (a) Assessment variables were measured on both HC and PDS, (b) variables were continuous, (c) assessment variables were administered in English, (d) raw data, standard deviations, and medians for an individual’s responses on the items of the test were excluded and the means or composite scores were used instead, if they were available, (e) variables that had constant values across all HC or PDS were removed, (f) variables that assess cognition were removed as assessing differences between PDS and HC in memory encoding and retrieval is the aim of this study, and (g) administrative variables were removed. We excluded discrete variables in the dynamic Bayesian networks as they would have violated the normal distribution assumptions of the dynamic Bayesian network algorithm.

After applying these selection criteria, 94 variables remained and a logistic regression analysis with stepwise variable selection was performed (in both directions) to distinguish PDS from HC, relying on the AIC criterion. This left nine variables as confounding factors: Years of education, number of marriages, body mass index, total months smoking, the exploratory excitability dimension of the temperament and character inventory, the anxiety dimension from the Hopkin’s symptom checklist, Golden and Meehl’s schizoid taxon dimensions from the MMPI, time taken to rise after waking from the Munich chronotype questionnaire, and number of brothers the participant had. The confounding and scanner related variables (scan number and number of scans since last trial) were added as nodes, to control for their effect, along with the 116 brain regions in each dynamic Bayesian network.

## 1.2.5. Data Diagnostics

There were artifacts in the MRI images in the study by Bilder et al. (2017), however, these were dismissed, as they did not affect the fMRI images. We imputed missing data in the demographic and clinical variables, and this is discussed in the results section. In terms of the fMRI data, a considerable number of voxels represented non-brain matter or air after parcellation. Areas in the brain not associated with a region in the automated anatomical labelling (AAL) atlas were discarded based on the method by Chuang (2020).

## 1.2.6. Analytic Strategy

### 1.2.6.1. Preprocessing

The fMRI data was preprocessed through a preprocessing pipeline and information regarding this can be found in the section Supplementary Information: Notes. The data preprocessing steps conducted for our research were parcellation and smoothing. These were performed using R Studio 4.1.1, and packages Rnifti (Clayden et al., 2021) and label4MRI(Chuang, 2020). Parcellation involved dividing the brain into distinct brain regions. Voxels, with three dimensional coordinates, from the fMRI scans were assigned labels for the brain region they belong to. The AAL atlas, as derived by Chuang (2020) and Tzourio-Mazoyer et al. (2002), was used to segment the fMRI brain images into 116 regions, and label each voxel with the corresponding brain region. Smoothing was performed by averaging BOLD responses across voxels in a region. A mean BOLD response was then obtained for each region, time period, experimental condition, and participant. The amount of smoothing therefore depended on how many voxels were present within a region, and consequently there may be more variability in some brain regions compared to others.

### 1.2.6.2. Hybrid Algorithm and Statistical Criteria Selection

Following a review of the literature on the different algorithms and statistical criteria, the hybrid algorithm RSMAX2 was selected, as the accuracy of hybrid-based algorithms is comparable to constraint-based algorithms and RSMAX2 is one of the fastest algorithms (Scutari et al., 2019). In terms of statistical criteria, for the restrict component of the algorithm, which involves reducing candidate parent nodes from the whole node set to a smaller set, the mutual information test was selected. This may have reduced the effect of a departure from normality in the data on the results, as the mutual information test is non-parametric. For the maximize component, which involves maximizing the statistical criterion of the network through adding, deleting, or changing the direction of edges, the *BIC* criterion was selected to avoid having to apply prior distributions for the parameters. Adding a prior, without expert knowledge, would likely result in minimal improvement, or worsening of the model, at the expense of larger computational time.

Various constraint and score-based algorithms can be applied when using RSMAX2 for a dynamic Bayesian network. To select the constraint and score-based algorithm for the dynamic Bayesian network, a random sample of 13 PDS participants and 13 HC participants was obtained. Both encoding and retrieval trials were included. A random sample was applied to reduce potential bias that the algorithm may have had towards PDS or HC, encoding or retrieval. Demographic, clinical, and scan related factors were not allowed (blacklisted) to be child nodes of brain regions. There are nine constraint-based algorithms and two maximize algorithms available in the RSMAX2 function in the R package dbnR (Quesada, 2021). We tested every combination of constraint-based and score-based algorithm in individual dynamic Bayesian networks. Hence, 18 dynamic Bayesian networks were reconstructed. For each of the 18 dynamic Bayesian networks *BIC* scores and completion times were obtained. The constraint-based and score algorithm pair that produced a dynamic Bayesian network in a 72-hour period with the best *BIC* score was selected. As a result, the constraint-based and score-based algorithms selected to evaluate the hypotheses in this study were Si Hiton PC and Tabu, respectively.

### 1.2.6.3. Statistical Methods

For the dynamic Bayesian networks to assess the aims of the study, 100 bootstrap samples were obtained for both PDS and HC, and the encoding and retrieval tasks. We reconstructed a Gaussian dynamic Bayesian network for each of these bootstrapped samples to build 100 averaged networks for each group. One hundred bootstrap samples were used to obtain edge strengths (Thresholds) out of 100 and to reduce computational time. The dynamic Bayesian networks used in this study were models of two-time instances. This means the Gaussian dynamic Bayesian networks assumed a fixed—two time period—model that was homogeneous over all images acquired during the paired associate memory task.

When constructing averaged dynamic Bayesian networks, a threshold must be defined to decide whether to include an edge. This threshold refers to the proportion of occurrences of an edge between two brain regions from the 100 networks. Further, our study identified shared and non-shared edges. For each threshold, shared edges refer to an edge that is present in two comparison groups (Intersection) and non-shared edges refers to an edge that is present in one group but not the other (Symmetric difference). Some edges were bidirectional in a relationship between two brain regions. We excluded the edge directions that occurred in less than 50% of the total edges between the two brain regions (Nagarajan et al., 2013). The dynamic Bayesian networks used throughout the study were found using the dmmhc function in the dbnR package (Quesada, 2021).

The significance thresholds for all four groups were found using the bnlearn package (Scutari, 2009). A significance threshold refers to the strength of each edge needed to be included in an averaged network. The edge strengths were calculated as the proportion of occurrences the edge was present over the 100 bootstrapped reconstructed networks. Additionally, the edge strength for an edge between two brain regions, for example, from the left amygdala to the right amygdala, needed to be equal to or greater than the edge strength for the edge in the opposite direction, for example, from the right amygdala to the left amygdala, to avoid feedback loops in the averaged networks. The vector of edge strengths was then optimized by applying the Euclidean norm to obtain the significance threshold. More information regarding optimizing averaged networks to obtain a significance threshold can be found in (Scutari & Nagarajan, 2013) and it implementation can be found in (Scutari & Denis, 2014).

# Appendix B: Notes

The present research is an archival study using data collected from the University of Los Angeles California Consortium for Neuropsychiatric Phenomics LA5c Study accession number ds000030 (Bilder et al., 2017). The data is available from https://openneuro.org/datasets/ds000030/versions/1.0.0. The Bilder et al. (2017) study included six tasks that assess different cognitive functions. These tasks were presented during distinct fMRI image acquisitions. The tasks that apply to our study is memory encoding and retrieval trials in the paired associate memory task. Information for the physical properties of the scanner, the conditions in the paired associate memory tasks, and the preprocessing method by the authors of the data can be retrieved from Gorgolewski et al. (2017). Our research included only PDS and HC with images acquired during experimental trials, I.e., when the paired associate memory trial was presented. Hence, we excluded images during non-experimental conditions. Links to the interactive networks’ charts, which vary by threshold can be found in this web address: https://github.com/KhanBuchwald/KhanBuchwald-Schizophrenia\_Dymanic\_Network\_Plot.git

# Appendix C: Figure 4

**Figure 4**

*Clustering Coefficient, Mean Centrality, Node Degree, Mean Weighted Shortest Path Length as a Function of Significance Threshold*

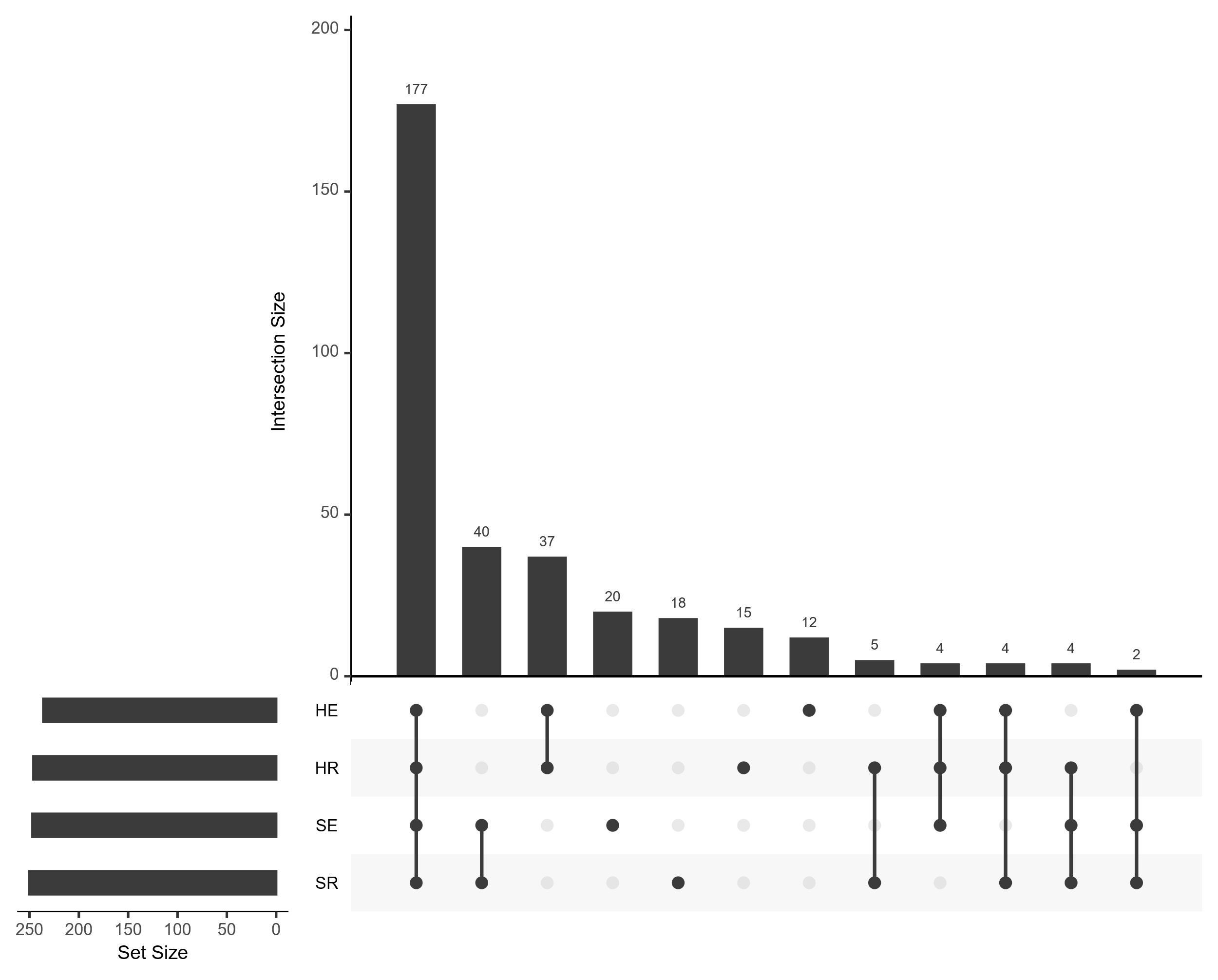
|  |  |
| --- | --- |
|  |  |
|  |  |

Note. The solid dark grey line refers to encoding in people diagnosed with schizophrenia; the dotted dark grey line refers to retrieval in people diagnosed with schizophrenia; the solid light grey line refers to encoding in healthy case-controls; the dotted light grey line refers to retrieval in healthy case-controls; Weights in the mean weighted shortest path length was calculated by subtracting 101 from the number of occurrences of that edge was present in all bootstrap samples. Average path lengths were calculated by averaging the weighted length of each node to all other nodes in the network.

# Appendix D: Figure 5

**Figure 5**

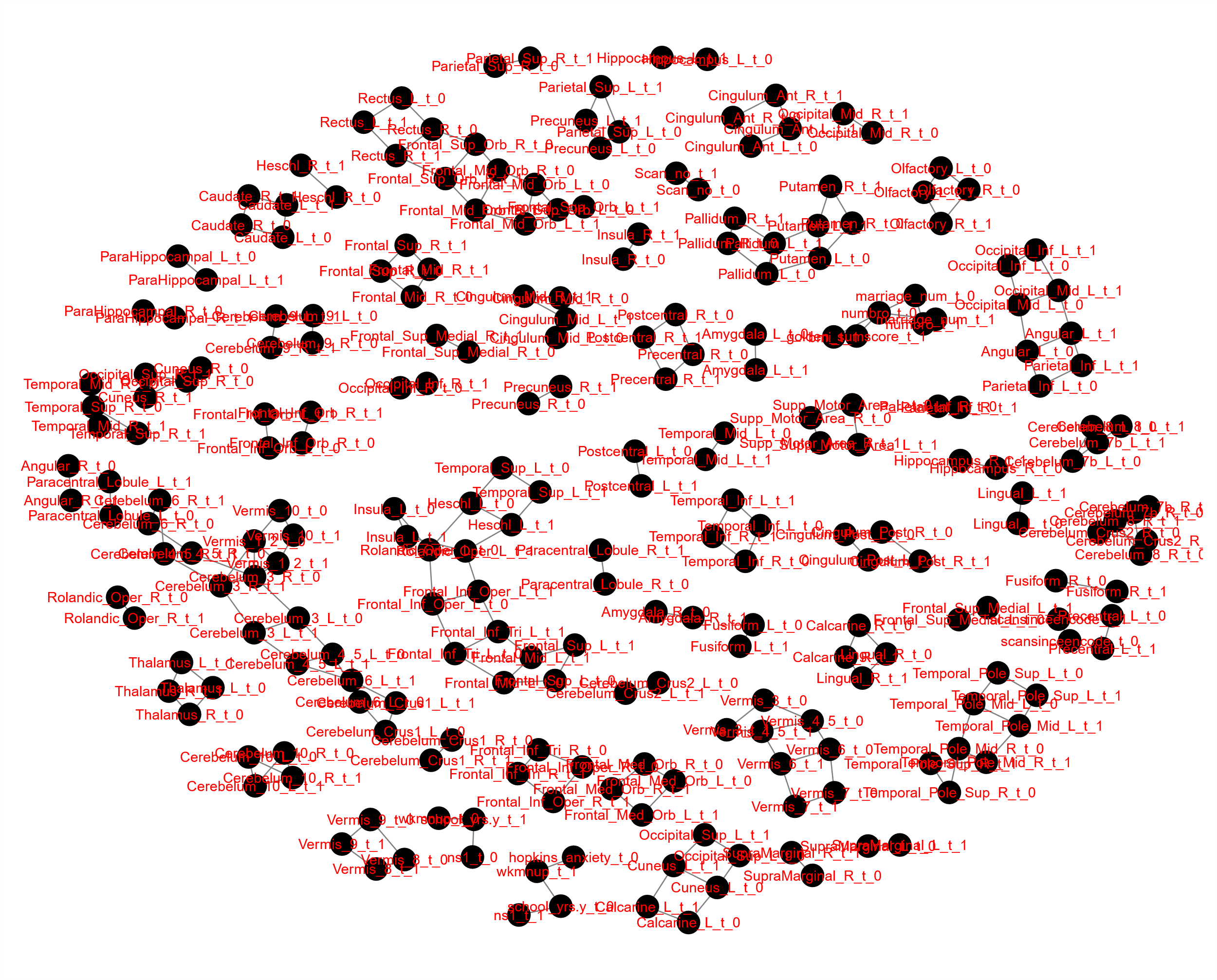
*Upset plot of edges for averaged networks at a threshold of zero*



# Appendix E: Figure 6 and 7

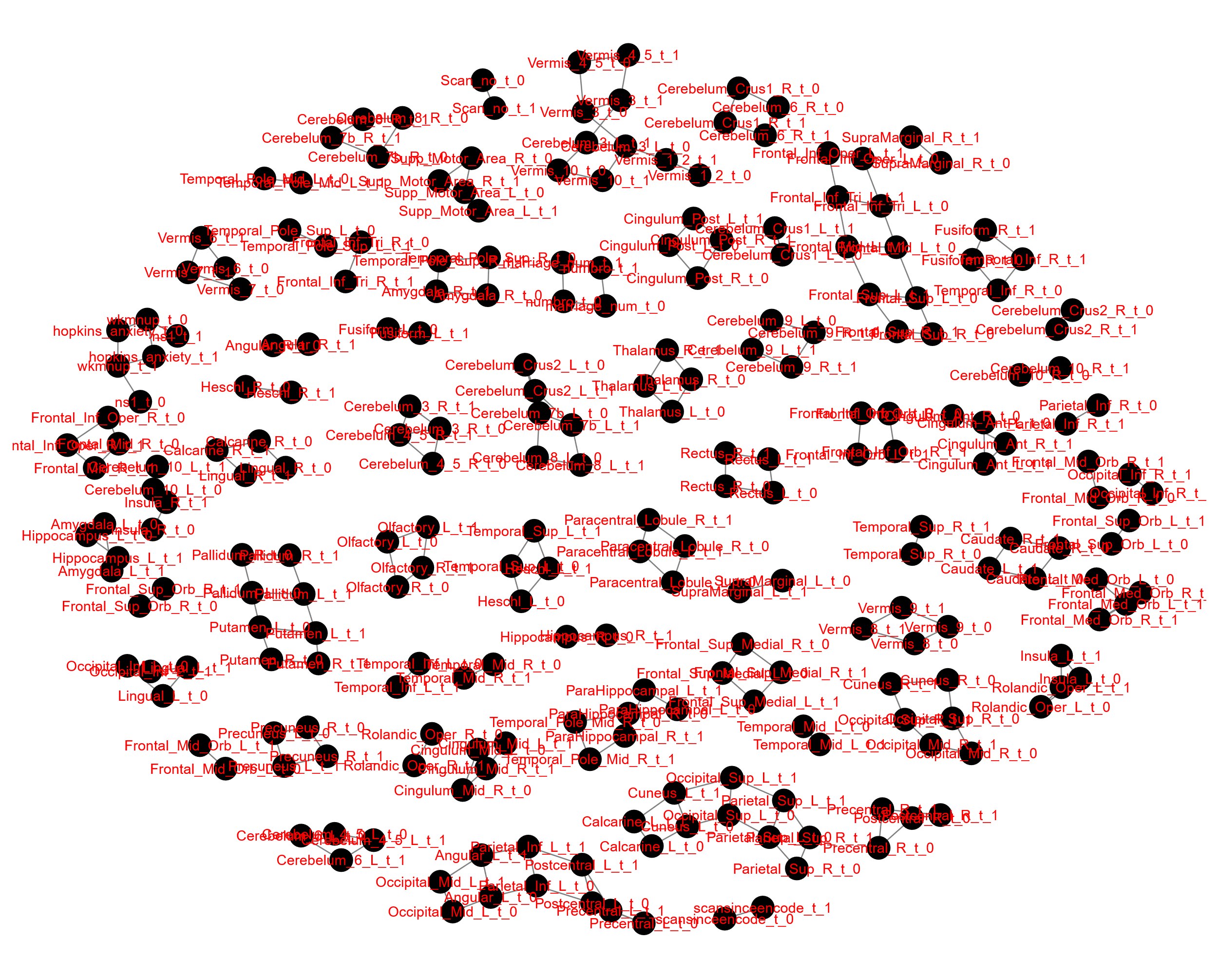
**Figure 6**

*Network plot of encoding in PDS at a threshold of .5.*



**Figure 7**

*Network plot of encoding in HC at a threshold of .5.*



# References

[Dataset] Bilder, R., Poldrack, R., Cannon, T., London, E., Freimer, N., Congdon, E., Karlsgodt, K., & Sabb, F. (2017). *UCLA Consortium for Neuropsychiatric Phenomics LA5c Study* Version 1.0.5). <https://doi.org/10.18112/openneuro.ds000030.v1.0.0>

Poldrack, R.A., Congdon, E., Triplett, W., Gorgolewski, K., Karlsgodt, K., Mumford, J., Sabb, F., Freimer, N., London, E., & Cannon, T., 2016. A phenome-wide examination of neural and cognitive function. *Sci. Data.* 3(1), 1-12. <https://doi.org/10.1038/sdata.2016.110>

First, M.B., & Gibbon, M. (2004). The structured clinical interview for DSM-IV axis I disorders (SCID-I) and the structured clinical interview for DSM-IV axis II disorders (SCID-II). In *Comprehensive handbook of psychological assessment, Vol. 2: Personality assessment.* (pp. 134-143). John Wiley & Sons.

Andreasen, N.C., 1984. Scale for the assessment of positive symptoms (SAPS). The University of Iowa.

Andreasen, N.C., 1989. The scale for the assessment of negative symptoms (SANS): conceptual and theoretical foundations. *Br. J. Psychiatr.* 155(S7), 49-52. <https://doi.org/https://doi.org/10.1192/S0007125000291496>

Overall, J.E., & Gorham, D.R., 1988. The brief psychiatric rating scale (BPRS): Recent developments in ascertainment and scaling. *Psychopharmacol. Bull.*, 97-99.

Spirtes, P., & Meek, C. (1995). Learning Bayesian networks with discrete variables from data. KDD.

Hansen, B.B., 2004. Full matching in an observational study of coaching for the SAT. *J. Am. Stat. Assoc.* 99(467), 609-618. <https://doi.org/10.1198/016214504000000647>

Chuang, Y.-S. (2020). *label4MRI: MRI-labeling* [R Package].

Clayden, J., Cox, B., & Jenkinson, M. (2021). *RNifti: Fast R and C++ access to NIfTI images.* In (Version R package version 1.3.1) <https://CRAN.R-project.org/package=RNifti>

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*.15(1), 273-289. <https://doi.org/1006/nimg.2001.0978>

Scutari, M., Graafland, C.E., & Gutierrez, J.M., 2019. Who learns better bayesian network structures: accuracy and speed of structure learning algorithms. *Int. J. Approx. Reason.* 115, 235-253. <https://doi.org/https://doi.org/10.1016/j.ijar.2019.10.003>

Quesada, D. (2021). *dbnR: dynamic bayesian network learning and inference.* In (Version R package version 0.7.3) <https://CRAN.R-project.org/package=dbnR>

Nagarajan, R., Scutari, M., & Lebre, S., 2013. Bayesian networks in R: With applications in systems biology. Springer New York, London.

Scutari, M., 2009. Learning Bayesian networks with the bnlearn R package. *arXiv preprint arXiv:0908.3817*.

Scutari, M., & Nagarajan, R., 2013. Identifying significant edges in graphical models of molecular networks. *Artificial Intel. Med.* 57(3), 207-217. <https://doi.org/https://doi.org/10.1016/j.artmed.2012.12.006>

Scutari, M., & Denis, J.B., 2014. Bayesian networks: with examples in R. CRC press.

Gorgolewski, K.J., Durnez, J., & Poldrack, R.A., 2017. Preprocessed consortium for neuropsychiatric phenomics dataset. *F1000Res.* 6. <https://doi.org/10.12688/f1000research.11964.2>