Revisiting Abnormalities in Brain Network Architecture Underlying Autism Using Topology-Inspired Statistical Inference

Sourabh Palande[1,2], Vipin Jose[1,2], Brandon Zielinski[3], Jeffrey Anderson[4], P. Thomas Fletcher[1,2], and Bei Wang[1,2]

[1] Scientific Computing and Imaging Institute, University of Utah
[2] School of Computing, University of Utah
[3] Pediatrics and Neurology, University of Utah
[4] Radiology, University of Utah

Abstract. A large body of evidence relates autism with abnormal structural and functional brain connectivity. Structural covariance MRI (scMRI) is a technique that maps brain regions with covarying gray matter density across subjects. It provides a way to probe the anatomical structures underlying intrinsic connectivity networks (ICNs) through the analysis of the gray matter signal covariance. In this paper, we apply topological data analysis in conjunction with scMRI to explore network-specific differences in the gray matter structure in subjects with autism versus age-, gender- and IQ-matched controls. Specifically, we investigate topological differences in gray matter structures captured by structural covariance networks (SCNs) derived from three ICNs strongly implicated in autism, namely, the salience network (SN), the default mode network (DMN) and the executive control network (ECN). By combining topological data analysis with statistical inference, our results provide evidence of statistically significant network-specific structural abnormalities in autism, from SCNs derived from SN and ECN. These differences in brain architecture are consistent with direct structural analysis using scMRI (Zielinski et al. 2012).

1 Introduction

Autism is a complex developmental disorder characterized by impairment in social interactions, difficulty in verbal and nonverbal communications and repetitive behaviors. Although the exact mechanism of its development remains unclear, there is strong evidence relating autism to abnormal white matter and functional connectivity between brain regions. Structural abnormalities can be identified using voxel-based morphometry by comparing gray matter, white matter volumes, cortical thicknesses and their growth trajectories [11] across diagnostic groups. Although the gross brain differences have been well-documented [5], investigations into specific regional abnormalities in brain structure have reported divergent results [14]. These inconsistent findings, however, may reflect discrete abnormalities in the brain network. Research has revealed a finite

set of canonical domain-specific *resting state* or *intrinsic connectivity networks* (ICNs) that organize the brain function [8]. Many of the regions with reported abnormalities in autism lie within these ICNs. Network-specific differences could account for seemingly contradictory findings from previous studies.

Structural covariance MRI (scMRI) maps brain regions with covarying gray matter density across subjects, suggesting shared developmental or genetic influences. Seeley et al. [12] have used scMRI to demonstrate that specific brain disorders affect distinct ICNs and the corresponding gray matter regions. Using a similar technique, Zielinski et al. [16] have shown that there are network-specific structural differences between autism and control groups which are consistent with clinical aspects of the disease and that reported functional abnormalities in autism have a structural bias. Several recent studies have applied the scMRI technique to find network-specific structural abnormalities in other diseases such as Alzheimer's [10] and Huntington's [9].

scMRI identifies regions of gray matter that have a statistically significant correlation with a specific seed region of interest (ROI). We can model all pairwise correlations (across subjects) among the gray matter regions identified by the seed-based covariance map as a network. Comparing these networks across diagnostic groups may provide information not captured by direct comparisons between individual regions.

Several graph-theoretic measures have been proposed previously to compare networks [3]. However, a major drawback of these measures is their reliance on a fixed network topology. That is, these measures are typically based on a graph obtained by thresholding the connectivity matrix. The choice of threshold is crucial in such analyses. Different heuristics have been suggested to determine the threshold depending on which properties of the network are of interest. However, it is often not possible to determine a unique optimal threshold.

In this paper, we apply topological data analysis to structural covariance networks (SCNs) derived from three ICNs strongly implicated in autism; the default mode network (DMN), the salience network (SN) and the executive control network (ECN). Our method is based upon a core technique from topological data analysis known as *persistent homology* [6] where we extract topological features across all thresholds from a given network. We make use of topology-inspired statistical inference first reported by Chung et al. [4] to compare the extracted topological features. By combining topological data analysis with statistical inference, our results provide statistically significant evidence of structural abnormalities underlying SN and ECN in autism. Our results are consistent with the observations of Zielinski et al. [16] and may offer new insights towards interpreting fine-scale network-specific structural differences.

2 Technical Background

2.1 Structural Covariance Network

We use scMRI to extract a network-specific set of brain regions with covarying gray matter density across subjects. Given a seed ROI, separate condition-by-

covariate analysis is performed for each gray matter region, in which the mean seed gray matter density is the covariate of interest and disease status is the grouping variable. Total brain volume (TBV) is included as a covariate-of-no-interest. This design enables us to determine the whole-brain patterns of seed-based structural covariance in each group. One-sample t-tests are performed to identify regions with significant groupwise gray matter density covariance with the seed ROI across subjects.

All pairwise correlations between gray matter densities across subjects, for pairs of identified regions, are modeled as a network. In what follows, we refer to such a network as the structural covariance network (SCN). The SCN, therefore, is a weighted, undirected graph $G(V, E, W)$, with gray matter regions as vertices and absolute values of pairwise correlations as edge weights. In particular, we compare SCNs generated with seed ROIs anchoring the three ICNs strongly implicated in autism, the SN, the ECN and the DMN. In the context of this paper, for simplicity (unless otherwise specified), we describe these SCNs by the name of their corresponding ICNs, namely, SN-SCN, ECN-SCN and DMN-SCN.

2.2 Graph Filtration

We extract topological features at multiple scales from a structural covariance network G by applying topological data analysis to a nested sequence of graphs constructed from G, referred to as the graph filtration.

Let $V = \{v_i \mid i = 1, \ldots, n\}$ be the vertex set with n vertices. Let E denote the edge set and W denote the set of edge weights. The edge between vertices v_i, v_j is denoted by e_{ij} and its weight is denoted by w_{ij}. $|E|$ denotes the number of edges. For a given threshold λ, we obtain a binary graph G_λ by removing edges with weight $w_{ij} \leq \lambda$. The adjacency matrix $A_\lambda = (a_{ij}(\lambda))$ is given by:

$$a_{ij}(\lambda) = \begin{cases} 0 & w_{ij} \leq \lambda \\ 1 & o.w. \end{cases}$$

As λ increases, more and more edges are removed from the graph. We can generate a sequence of thresholds, $\lambda_0 = 0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_q$, where $q \leq |E|$ by setting λ_i's equal to edge weights arranged in ascending order.

Corresponding to the sequence of thresholds we get a nested sequence of binary graphs, referred to as a *graph filtration* $\mathbf{G}$:

$$G_{\lambda_0} \supseteq G_{\lambda_1} \supseteq G_{\lambda_2} \supseteq \cdots \supseteq G_{\lambda_q}.$$

We can measure the connectivity of a graph by its 0-th Betti number, β_0, which is the number of connected components in the graph. As the threshold λ increases, $\beta_0(G_\lambda)$ of the corresponding graph also increases. The $\beta_0(G_{\lambda_i})$ of the graphs in filtration $\mathbf{G}$ form a monotonic sequence of integers,

$$\beta_0(G_{\lambda_0}) \leq \beta_0(G_{\lambda_1}) \leq \beta_0(G_{\lambda_2}) \leq \cdots \leq \beta_0(G_{\lambda_q}).$$

Suppose we start with a connected graph $G = G_{\lambda_0}$. We have $\beta_0(G_{\lambda_0}) = 1$ and $\beta_0(G_{\lambda_q}) = |V| = n$ by construction. Given n nodes, there are at most $\binom{2n}{n}$

unique edge weights; therefore $q \leq \binom{2n}{n}$. The number of all possible monotonic integer sequences of length q, starting with 1 and ending with n, is finite.

Following the formulation of Chung et al. [4], the distance between two given graph filtrations $\mathbf{G}$ and $\mathbf{H}$ can be defined as:

$$D_q(\mathbf{G}, \mathbf{H}) = \sup_{0 \leq i \leq q} |\beta_0(G_{\lambda_i}) - \beta_0(H_{\lambda_i})|. \quad (1)$$

Intuitively, if we plot the two sequences of Betti numbers as a function of λ (the graph of such a function is referred to as the β_0 curve), this distance D_q measures the largest gap between the two curves. Given that the number of possible sequences is finite, D_q can take only a finite number of discrete integer values. Computing the β_0 curve for a given graph filtration could follow the standard algorithm for persistent homology [6]; in practice, a simpler algorithm can be used to capture the λ values when the number of components (clusters) decreases during the filtration.

2.3 Statistical Inference

We model the structural covariance networks for autism and control groups as weighted graphs G and H, respectively, with the corresponding graph filtrations $\mathbf{G}$ (autism) and $\mathbf{H}$ (control). We would like to test the equivalence of the two filtrations. In particular, we would like to test the null hypothesis H_0 against the alternative hypothesis H_1, where

$$H_0 : \beta_0(G_{\lambda_i}) = \beta_0(H_{\lambda_i}) \quad \text{for all } \lambda_i;$$

$$H_1 : \beta_0(G_{\lambda_i}) \neq \beta_0(H_{\lambda_i}) \quad \text{for some } \lambda_i.$$

By taking the *supremum* over all λ_i, D_q takes care of multiple comparisons implied in the hypotheses. Chung et al. [4] have provided a combinatorial derivation of the exact probability distribution of D_q. The proof is based on the Kolmogorov-Smirnov test [2]. This eliminates the need for numerically permuting samples for the test of hypothesis. The asymptotic probability distribution of D_q is given by :

$$\lim_{q \to \infty} P(D_q/\sqrt{2q} \geq d) = 2 \sum_{i=1}^{\infty} (-1)^{i-1} e^{-2i^2 d^2},$$

and the p-value under the null hypothesis can be computed as :

$$p = 2e^{-d_0^2} - 2e^{-8d_0^2} + 2e^{-18d_0^2} + \cdots \approx 2e^{-d_0^2} - 2e^{-8d_0^2} + 2e^{-18d_0^2},$$

where d_0 is the smallest integer greater than $D_q/\sqrt{2q}$.

3 Methods

3.1 Data Preprocessing

We derive our SCNs from the ICNs previously reported by Zielinski et al. [16,17]. Here, we review the preprocessing pipeline. 49 male subjects with autism, aged 3-22 years, are compared to 49 age-, gender- and IQ-matched typically developing control subjects. Images are acquired using a Siemens 3.0 Tesla Trio MRI scanner. Whole brain isotropic MPRAGE image volumes are acquired in the sagittal plane using an 8-channel receive-only RF head coil, employing standard techniques (TR = 2300 ms, TE median = 3 ms, matrix median = $256 \times 256 \times 160$, flip angle = 12°, voxel resolution = 1 mm^3, acquisition time = 9 min 12 sec).

Customized image analysis templates are created by normalizing, segmenting and averaging T1 images using SPM5 according to the processing pipeline proposed in [1,15]. First, images are transformed into standard space using a 12-parameter affine-only linear transformation and segmented into three tissue classes representing gray matter, white matter and cerebrospinal fluid. Then smoothly varying intensity changes as well as artifactual intensity alterations as a result of the normalization step are corrected for using a standard modulation algorithm within SPM5. Finally, the resulting segmented maps are smoothed using a 12-mm full-width at half-maximum Gaussian kernel.

In performing the scMRI analysis, a two-pass procedure is utilized, wherein study-specific templates are first created by segmenting our sample using a canonical pediatric template. Then tissue-specific prior probability maps are created from our sample. The tissue compartments are then re-segmented using this sample-specific template, so that the age range of our sample precisely matches that of the template(s) upon which the ultimate segmentations are based.

3.2 Structural Covariance Networks and Statistical Inference

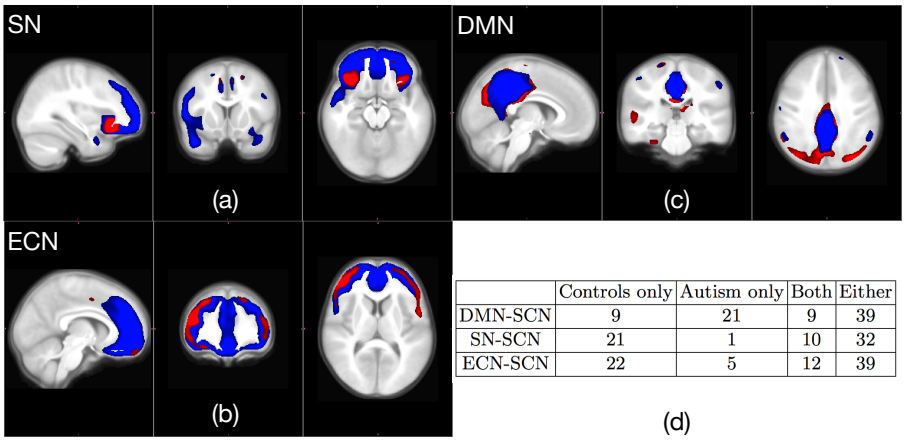
The preprocessed images contain 7266 gray matter voxels. For each diagnostic group, a *whole-brain* SCN is constructed by computing pairwise correlations among all voxels.

To study network-specific structural covariance, 4-mm-radius spherical seed ROIs are selected within the right frontoinsular cortex (R FI) [12], the right dorsolateral prefrontal cortex (R DLPC) [13] and the right posterior cingulate cortex (R PCC) [7]. These regions anchor the salience network (SN), the executive control network (ECN) and the default-mode network (DMN), respectively [12,7].

For each diagnostic group and each seed ROI, we generate SCNs following the process described in Section 2. The structural covariance maps corresponding to the seed ROI are shown in Fig. 1(a)-(c). The SCNs are composed of 4-mm-radius spherical regions identified by these maps. Further comparisons in Fig. 2 show that the two maps overlap in very few regions. Some regions present in the map for the control group are absent in the map for the autism group. Conversely, some regions are present only in the map for the autism group but not in the map for the control group. Fig. 1(d) lists the number of regions present in controls

but not in autism, in autism but not in controls and in both as well as in either autism or control.

We then construct and compare the SCNs derived from corresponding subsets of regions for each seed ROI. For each comparison, SCNs are derived for both diagnostic groups (autism and controls) separately. Graph filtrations are constructed for both networks. The distance D_q between the two resulting β_0 curves and the corresponding p-value for the test hypotheses is obtained accordingly.



	Controls only	Autism only	Both	Either
DMN-SCN	9	21	9	39
SN-SCN	21	1	10	32
ECN-SCN	22	5	12	39

(d)

Fig. 1: (a)-(c) Structural covariance maps with seed in R FI, R DLPC and R PCC, anchoring SN, ECN and DMN, respectively. Red represents the autism group map, blue represents the control group map. (d) Number of regions identified from scMRI map for a given seed region.

4 Results

We apply statistical inference and compare SCNs across groups of subjects with autism and typically developing control subjects. We begin by comparing the global SCNs composed of all 7266 gray matter voxels in the preprocessed images. Applying the statistical inference detailed in Section 3, we obtain a p-value of $6.6250179 \times 10^{-19}$. The differences in whole-brain gray matter composition between the autism and control groups have been well established in previous studies [5]. The near-zero p-value shows that such differences can also be captured by the topological features extracted from the global SCNs.

For a closer analysis, we compare the SCNs generated with seed ROIs anchoring the three ICNs (SN, ECN and DMN), referred to as SN-SCN, ECN-SCN and DMN-SCN, respectively. Recall that the structural covariance maps for the autism and the control groups overlap in very few regions. We construct and compare SCNs derived from subsets of regions that are present in controls but

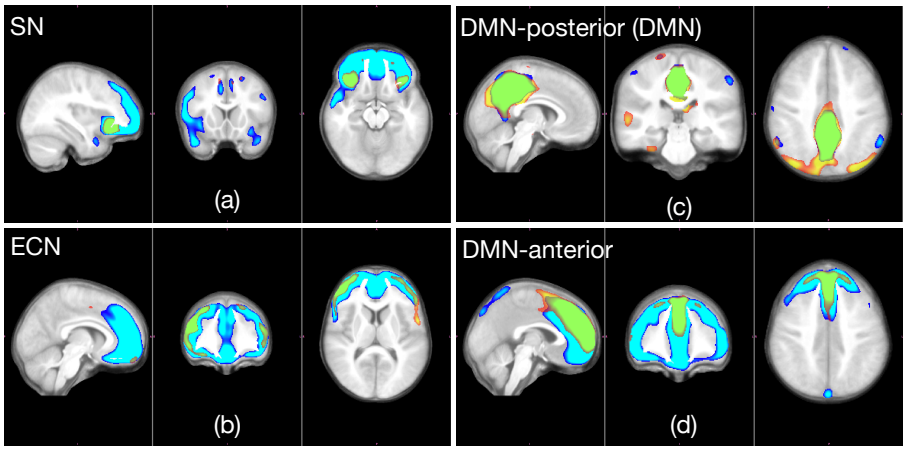


Fig. 2: scMRI maps are further illustrated here with red to yellow (autism) and dark blue to light blue (control) color look up tables. The color gradation indicates increasing statistical significance. The overlapping regions among the autism and control groups are highlighted in green. Note for (c) and (d): Our data consists of subjects with an average age of about 13 years. The underlying structure of the DMN is not fully developed at this age. We include two DMN maps with different seeds to show that the posterior part (c) is not yet integrated with the anterior part (d). In our analysis, we use the posterior covariance map (c) which corresponds to the most common seed for DMN in adults (R PCC).

not in autism, present in autism but not in controls and present in both as well as present in either.

The β_0 curves corresponding to comparisons among global SCNs, and seed-specific SCNs generated from regions present either in autism or controls, are shown in Fig. 3. Table 1 lists the p-values obtained by applying the statistical inference procedure to the corresponding SCNs. By combining topological data analysis with statistical inference, our results provide statistically significant evidence of network-specific structural abnormalities in autism for both SN-SCNs and ECN-SCNs.

	Controls only	Autism only	Both	Either
DMN-SCN	0.6271670	0.0815188	0.9538228	0.2369032
SN-SCN	**0.0014932**	NA	**0.0366311**	$\mathbf{1.3269078 \times 10^{-6}}$
ECN-SCN	**0.0422562**	0.9960098	**0.0059460**	$\mathbf{1.7996732 \times 10^{-6}}$

Table 1: p-values for statistical inference on SCNs derived from ICNs; DMN-SCNs, SN-SCNs and ECN-SCNs. Only one region in SN is present in autism but not in controls where the inference procedure is not applicable.

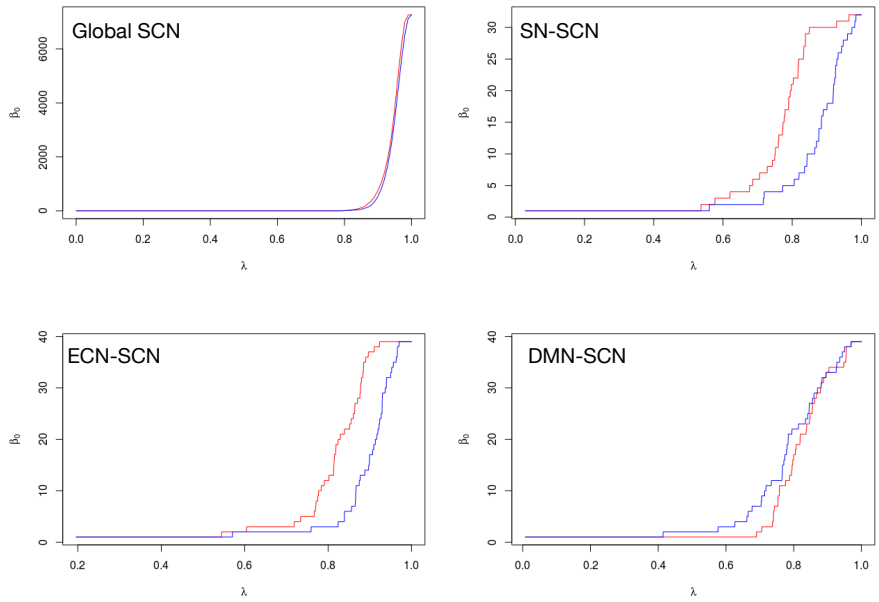


Fig. 3: β_0 curves from Global SCNs as well as SN-SCNs, ECN-SCNs and DMN-SCNs, generated from regions present in either autism (red) or controls (green) respectively.

5 Conclusion and Discussion

Using direct comparisons of structural covariance maps, Zielinski et al. have shown the structural differences in gray matter regions underlying *intrinsic connectivity networks* (ICNs) such as SN [16], DMN [16] and ECN (Brandon Zielinski, personal communication, May 2017) between the autism and the control groups. In contrast, our method compares the *structural covariance networks* (SCNs), which are composed of all possible pairwise correlations between gray matter regions and not just their covariance with a specific seed region.

Our inference procedure obtains statistically significant p-values among the SCNs derived from SN and ECN (SN-SCNs and ECN-SCNs) when comparing networks constructed from regions present in controls only, regions present in both autism and controls, as well as regions present in either autism or controls. Our results indicate statistically significant differences in the 0-dimensional topological features of these SCNs; this result is consistent with the findings of Zielinski et al. [16].

Our method, however, does not capture significant differences in the topology of SCNs derived from DMN (DMN-SCNs). It is possible that considering only pairwise interactions among gray matter regions (that is, 0-order topological features encoded by the β_0 curves, corresponding to the number of connected components) may not be sufficient to capture the complex topological differ-

ences within these SCNs. Analyzing three-way or four-way interactions, capturing higher-order topological features such as tunnels and voids and focusing on specific sites directly involved in merging components in the graph filtration may provide further insights into these SCNs.

Final Remarks A key insight from the work of Zielinski et al. [16] is that structure enables function and functional collaboration enables structure. Our work, in particular, helps summarize these multidimensional, structure-function relationships by conceptualizing them as higher order topological relationships. The SCNs are constructed using inputs from both the function, in form of the seed ROI anchoring specific ICNs and the structure, in the form of gray matter density signals.

The techniques in [16] compare covariance maps directly. Such a comparison helps to identify whether a particular region is present or absent in the autism vs the control group maps. The regions in these maps are assigned significance measures using their covariance with respect to a specific seed region. Our work, on the other hand, uses the SCNs to encode all pairwise associations among regions, where the extent of an association is measured by the correlations across subjects. Our results indicate that there are statistically significant differences in the way networks are connected, which implies differences in the patterns of pairwise association across diagnostic groups.

To illustrate the advantage, consider the regions present in the SN- or ECN-specific covariance maps of both the autism and the control groups. Direct comparison of the covariance maps does not provide any further insight into these regions. Our method on the other hand, shows that there are statistically significant differences in the topological features derived from SN-SCN and ECN-SCN composed of these regions (Table 1, p-values of 0.0366311 and 0.0059460 respectively). However, it should be noted that the SCNs are abstract networks and do not represent physical connectivity between the regions. This limits the interpretability of our results to some extent. Further analysis is needed in order to quantify and better interpret the differences in the SCNs suggested by the statistical inference.

Acknowledgements This work was supported by NIH grant R01EB022876 and NSF grant IIS-1513616.

References

1. M. Altaye, S. K. Holland, M. Wilke, and C. Gaser. Infant brain probability templates for {MRI} segmentation and normalization. *NeuroImage*, 43(4):721 – 730, 2008.
2. W. Böhm and K. Hornik. A kolmogorov-smirnov test for r samples. *Fundam. Inf.*, 117(1-4):103–125, Jan. 2012.
3. E. Bullmore and O. Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3):186–198, 2009.

4. M. K. Chung, V. Villalta-Gil, H. Lee, P. J. Rathouz, B. B. Lahey, and D. H. Zald. Exact topological inference for paired brain networks via persistent homology. *bioRxiv:140533*, 2017.
5. E. Courchesne, K. Pierce, C. M. Schumann, E. Redcay, J. A. Buckwalter, D. P. Kennedy, and J. Morgan. Mapping early brain development in autism. *Neuron*, 56(2):399 – 413, 2007.
6. H. Edelsbrunner, D. Letscher, and A. J. Zomorodian. Topological persistence and simplification. *Discrete and Computational Geometry*, 28:511–533, 2002.
7. D. A. Fair, A. L. Cohen, N. U. F. Dosenbach, J. A. Church, F. M. Miezin, D. M. Barch, M. E. Raichle, S. E. Petersen, and B. L. Schlaggar. The maturing architecture of the brain's default network. *Proceedings of the National Academy of Sciences*, 105(10):4028–4032, 2008.
8. M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen, and M. E. Raichle. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*, 102(27):9673–9678, 2005.
9. L. Minkova, S. B. Eickhoff, A. Abdulkadir, C. P. Kaller, J. Peter, E. Scheller, J. Lahr, R. A. Roos, A. Durr, B. R. Leavitt, S. J. Tabrizi, S. Klöppel, and T.-H. Investigators. Large-scale brain network abnormalities in huntington's disease revealed by structural covariance. *Human Brain Mapping*, 37(1):67–80, 2016.
10. M. Montembeault, I. Rouleau, J.-S. Provost, and S. M. Brambati. Altered gray matter structural covariance networks in early stages of alzheimer's disease. *Cerebral Cortex*, 26(6):2650, 2016.
11. C. M. Schumann, C. S. Bloss, C. C. Barnes, G. M. Wideman, R. A. Carper, N. Akshoomoff, K. Pierce, D. Hagler, N. Schork, C. Lord, and E. Courchesne. Longitudinal magnetic resonance imaging study of cortical development through early childhood in autism. *Journal of Neuroscience*, 30(12):4419–4427, 2010.
12. W. W. Seeley, R. K. Crawford, J. Zhou, B. L. Miller, and M. D. Greicius. Neurodegenerative diseases target large-scale human brain networks. *Neuron*, 62:42–52, 04 2009.
13. W. W. Seeley, V. Menon, A. F. Schatzberg, J. Keller, G. H. Glover, H. Kenna, A. L. Reiss, and M. D. Greicius. Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience*, 27(9):2349–2356, 2007.
14. K. A. Stigler, B. C. McDonald, A. Anand, A. J. Saykin, and C. J. McDougle. Structural and functional magnetic resonance imaging of autism spectrum disorders. *Brain Research*, 1380:146–161, 03 2011.
15. M. Wilke, S. K. Holland, M. Altaye, and C. Gaser. Template-o-matic: A toolbox for creating customized pediatric templates. *NeuroImage*, 41(3):903 – 913, 2008.
16. B. A. Zielinski, J. S. Anderson, A. L. Froehlich, M. B. D. Prigge, J. A. Nielsen, J. R. Cooperrider, A. N. Cariello, P. T. Fletcher, A. L. Alexander, N. Lange, E. D. Bigler, and J. E. Lainhart. scmri reveals large-scale brain network abnormalities in autism. *PLOS ONE*, 7(11):1–14, 11 2012.
17. B. A. Zielinski, E. D. Gennatas, J. Zhou, and W. W. Seeley. Network-level structural covariance in the developing brain. *Proceedings of the National Academy of Sciences*, 107(42):18191–18196, 2010.