1   **Assessment of Functional Connectome Construction Strategies in**

2   **Neurodegeneration**

3   **Vanhoecke J[1], McColgan P[1], Razi A[2,3], Gregory S[1], Seunarine K[4], Durr A[5], Roos R[6],**

4   **Leavitt B[4], Scahill RI[1], Clark C[4], Tabrizi SJ[1,7], Rees G[2] and Track On-HD Investigators**

5   1 Department of Neurodegenerative Disease, UCL Institute of Neurology, London, WC1N

6   3BG, UK

7   2 Wellcome Trust Centre for Neuroimaging, UCL Institute of Neurology, London, WC1N

8   3BG, UK

9   3 Department of Electronic Engineering, NED University of Engineering and Technology,

10   Karachi, Pakistan

11   4 Developmental Imaging and Biophysics Section, UCL Institute of Child Health, London,

12   WC1N 1EH, UK

13   5 Institut du Cerveau et de la Moelle épinière (ICM), Genetic department, AP-HP, Sorbonne

14   Université, Inserm U 1127, CNRS UMR 7225, Pitié-Salpêtrière University Hospital, Paris,

15   France

16   6 Department of Neurology, Leiden University Medical Centre, 2300RC Leiden, The

17   Netherlands

18   7 National Hospital for Neurology and Neurosurgery, Queen Square, London, WC1N 3BG,

19   UK

20   Correspondence to: Dr Peter McColgan,

21   Huntington's Disease Centre, Department of Neurodegenerative Disease,

22   UCL Institute of Neurology,

23   Queen Square,

24   London WC1N 3BG, UK

25   E-mail: p.mccolgan@ucl.ac.uk

1   **Keywords**

2   Functional connectomics, modularity, machine learning, Huntington's disease, graph theory

3   **Abbreviations**

4   AUC: area under the curve; BCT: Brain Connectivity Toolbox; BW: Brainwaver; CM:

5   connectivity matrix/matrices; CompCor: principal component-based noise correction; CSF:

6   cerebrospinal fluid; FC: functional connectivity; FDR: false discovery rate; HC: healthy

7   controls; HD: Huntington's disease; MODWT: maximal overlap discrete wavelet transform;

8   MST: minimum spanning tree; PC: participation coefficient; PRoNTo: Pattern Recognition for

9   Neuroimaging Toolbox; ROI: region of interest; rs-fMRI: resting state functional Magnetic

10   Resonance Imaging; SD: standard deviation; SPM statistical parametric mapping; T1WI: T1-

11   weighted imaging; WMW: Wilcoxon-Mann-Whitney test

12

1 Abstract

2   Connectomics can be used to investigate functional brain networks in neurodegenerative

3   diseases including Huntington's disease (HD). In this developing field, different connectome

4   construction strategies have emerged in parallel. However, there is a need to understand the

5   influences of different strategies on subsequent analyses when constructing a connectome. This

6   study systematically compares connectome construction strategies based on their biological

7   relevance to functional networks in neurodegeneration.

8   We asked which functional connectome construction strategy was best able to

9   discriminate HD gene carriers from healthy controls, and how such a strategy affected modular

10 organization of the network. The major factors compared were principal component-based

11 correction versus wavelet decomposition for physiological noise correction, the type of

12 parcellation atlas (functional, structural and multi-modal), weighted versus binarized networks,

13 and unthresholded versus proportionally thresholded networks. We found that principal

14 component-based correction generated the most discriminatory connectomes, while

15 binarization and proportional thresholding did not increase discrimination between HD gene

16 carriers and healthy controls. When a functional parcellation atlas was used, the highest

17 discrimination rates were obtained. We observed that the group differences in modular

18 organization of the functional connectome were greatly affected by binarization and

19 thresholding, showing no consistent pattern of modularity.

20   This study suggests that functional connectome construction strategies using principal

21 component-based correction and weighted unthresholded connectivity matrices may

22 outperform other strategies.

23

Introduction

Functional and structural brain networks can be constructed using resting state fMRI and diffusion tractography respectively. For functional connectomics, different methodologies have been used to construct and compare brain networks (Sporns, 2011). Amongst these factors some of the important choices are the type of physiological noise correction (Marchitelli *et al.*, 2016), the size and nature of the parcellation atlas (Arslan *et al.*, 2017), and the application of thresholding and binarization of connectivity matrices (Garrison *et al.*, 2015). However, comparative studies of the interactions of these factors when constructing the connectome have not been performed. Understanding the effects of connectome construction strategies is important to provide insights when developing a standard for the field.

One difficulty in evaluating connectome construction strategies is the lack of a gold standard. To overcome this, we set out to investigate how different strategies affect the ability of using a constructed connectome to distinguish between Huntington's disease (HD) gene carriers and healthy controls. HD is an autosomal dominant neurodegenerative disease, and is fully penetrant in those with greater than 39 cytosine-adenine-guanine (CAG) repeat expansions in the HD gene on chromosome 4, see McColgan 2018 for a review (McColgan and Tabrizi, 2018). Thus HD gene testing allows us to identify with certainty those who will develop the disease. In this context, genetic testing can be used as a diagnostic gold standard. This then allowed us to evaluate connectome construction strategies in their ability to discriminate between HD and healthy controls.

Options for physiological noise correction include band pass filtering and regression of white matter, CSF and global signal. More recently, due to concerns regarding global signal regression (Carbonell, Bellec and Shmuel, 2014) more conservative approaches use only the principal components of white matter and CSF signals. This approach was developed as the anatomical CompCor method (Behzadi *et al.*, 2007) and is implemented in the freely available

1   Conn functional connectivity toolbox (Whitfield-Gabrieli and Nieto-Castanon, 2012). Wavelet

2   decomposition can also remove signal attributable to physiological noise. Here, the signal is

3   broken down into its underlying constituent frequencies by scaling and shifting of a brief

4   oscillation, a so-called wavelet. By applying the maximal overlap discrete wavelet transform

5   (MODWT) with Daubechies wavelets (Daubechies, 1988) to the raw rs-fMRI time series,

6   correlation matrices are formed. Brainwaver is a freely available R-based package that can be

7   used to perform such a wavelet decomposition (Achard *et al.*, 2006).

8   Following physiological noise correction, the brain is parcellated into discrete regions,

9   such that each brain region acts as a node in the network and the temporal correlations of the

10   fMRI time series between regions act as functional connections or edges. Atlases can be defined

11   anatomically, such as the Desikan-Killiany atlas where labeling is inferred from anatomic

12   curvature (Desikan *et al.*, 2006). Functional atlases can also be defined based on resting state

13   connectivity (Yeo *et al.*, 2011). More recently multimodal atlases have been developed, such

14   as the Glasser atlas (Glasser *et al.*, 2016), which is based on task fMRI, rs-fMRI, and

15   cytoarchitectonic features. However, currently there is no consensus in the literature as to the

16   optimal strategy for brain parcellation (Arslan *et al.*, 2017).

17   Following the parcellation of the brain into various nodes or brain regions, thresholding

18   is often performed in order to remove spurious connections. Binarization can also be carried

19   out, such that the connection is either present or absent with the magnitude of the temporal

20   correlation ignored. As there is no consensus about what threshold value to use, often a range

21   of different thresholds are applied (Bullmore and Sporns, 2009). A variety of different threshold

22   approaches have been described in the literature including absolute, proportional, consistency

23   and consensus thresholding, but no optimal method has been identified (Qi *et al.*, 2015). When

24   thresholding is applied, it is often combined with a minimum spanning tree to retrieve the

25   backbone of the graph (Alexander-Bloch *et al.*, 2010), which is a subgraph with minimal

5

1   connection weights, while ensuring all nodes are connected to the network. Alternatively,

2   thresholding can be avoided by using weighted connectivity matrices, using the raw correlations

3   (Bullmore and Sporns, 2009).

4   Machine learning tools such as support vector machines can determine a discrimination

5   rate between two groups, as implemented in for example the Pattern Recognition for

6   Neuroimaging Toolbox, PRoNTo (Schrouff *et al.*, 2013). This has been used previously to

7   classify structural networks from different Brain-derived Neurotrophic Factor (BDNF)

8   genotypes from healthy subjects (Ziegler *et al.*, 2013).

9   Modularity is a measure of functional segregation and represents how well the network

10   can be divided into distinct cooperating units, also called modules or communities (Rubinov

11   and Sporns, 2010). In neurodegeneration, changes occur in the modular organization of the

12   brain (McColgan *et al.*, 2015) and therefore this is a complementary way to assess the effect of

13   pathology on the brain network.

14   Here we set out to address two specific questions: first, which functional connectome

15   construction is most able to discriminate HD gene carriers from healthy controls? Second, how

16   does the method of functional connectome construction affect determination of modular

17   organization? For both research questions, we undertook a systematic comparison considering

18   different factors and their interactions. These factors were principal component-based

19   correction versus wavelet decomposition for physiological noise correction, the type of

20   parcellation atlas (functional, structural and multi-modal), weighted versus binarized networks,

21   and unthresholded versus proportional thresholded networks, resulting in a systematic

22   comparison of 66 connectome construction strategies (Figure 1).

23   To our knowledge, there is no systematic comparison of the effect of all factors

24   mentioned above on functional brain networks, especially not applied in a clinical population.

25   While optimization of connectome construction protocol has been done before, it mainly

6

1   focused on one single factor, for instance the effect of thresholding, ignoring other influences

2   (Garrison *et al.*, 2015). As these factors could have critical interactions in determining an

3   optimal strategy, a systematic comparison can be particularly valuable.

4

1   Materials and Methods

2   Cohort

3   We assembled a cohort that comprised HD gene carriers and healthy controls from the final

4   visit (2014) of the prospective Track-On HD study (Klöppel *et al.*, 2015). From a total of 243

5   eligible participants, 78 participants were excluded due to poor quality rs-fMRI data (Klöppel

6   *et al.*, 2015). If not further specified, the remaining cohort of 165 participants was used in full.

7   For the classification analysis, a subdivision was taken, so that the number of participants in

8   each group was equal. Participants were pairwise matched for age and gender, so that each

9   match consisted of one HD gene carrier and one healthy control participant from the same

10   gender and with an age difference spanning less than 1 year. This strategy yielded 49 HD gene

11   carriers and 49 HC (Table 1), without significant differences in age (2 tail t-test, p=0.93), gender

12   (*a priori* equal), education (WMW, p=0.38) and study site (Chi-square, p=0.91). This

13   subdivision is further referred to as a pairwise matched cohort. An alternative and less stringent

14   cohort subdivision was also performed, in which cohorts were selected by median age, see

15   Supplementary Methods.

16

17   MRI acquisition

18   T1-weighted and rs-fMRI data were acquired on two different 3T MRI scanner systems (Philips

19   Achieva at Leiden and Vancouver and Siemens TIM Trio at London and Paris). Scanning time

20   was approximately 12 min for the T1-weighted acquisitions and 8 min for the rs-fMRI. For the

21   rs-fMRI T2*-weighted echo planar imaging sequence with a repetition time of 3 s was used,

22   resulting in 165 whole-brain time series. For both Siemens and Philips scans the voxel size was

23   3.3 x 3.3 x 3.3 mm^3. Detailed information on parameters and control procedures are provided

24   in the Track-On HD study (Klöppel *et al.*, 2015; McColgan *et al.*, 2017).

Data Pre-processing

Pre-processing of the structural and functional MRI data was performed with MATLAB R2016a, making use of SPM8 (Friston *et al.*, 2011) and Conn (Whitfield-Gabrieli and Nieto-Castanon, 2012). Each stage of the pre-processing was subjected to the quality control procedures of Track-On HD (Klöppel *et al.*, 2015). T1-weighted images were segmented using three different parcellation atlases: a combined functional atlas containing cortical (Yeo *et al.*, 2011) and sub-cortical regions (Choi, Yeo and Buckner, 2012), a structural atlas (Romero-Garcia *et al.*, 2012) and a multi-modal atlas (Glasser *et al.*, 2016). Atlases were registered to the anatomical image using NiftiReg (Modat *et al.*, 2010). For the Romero atlas the globus pallidus, amygdala, nucleus accumbens and cerebellum were excluded. The functional, structural and multi-modal parcellation strategies entailed 121 ROIs, 316 ROIs and 360 ROIs respectively.

Functional images were realigned, then corrected for estimated head movement, and coregistered to the new anatomical image. For each participant, a regional mean time series of the resting-state were obtained by averaging the fMRI time series over each voxel of a ROI. Physiological noise correction was then performed as detailed below.

Physiological noise correction

To perform a wavelet decomposition, six motion parameters were first linearly regressed out from the time series (Achard *et al.*, 2006). Residuals were decomposed with Brainwaver (http://cran.r-project.org/web/packages/brainwaver/index.html) in RStudio 0.98.1078, using Daubechies wavelets, and the correlation matrix generated in the 4th wavelet scale (0.03-0.06 Hz) was retrieved for further analysis, in keeping with the literature (Achard *et al.*, 2006). The use of higher frequencies from lower wavelet scales has been reported (Richiardi *et al.*, 2011; Vértes *et al.*, 2016). However, frequencies of 0.1-0.5 Hz are thought to originate from

1   respiration (Cordes *et al.*, 2001) and thus, to make an unbiased comparison with bandpass-

2   filtering used in CompCor, the lowest wavelet scale corresponding to a frequency below 0.1 Hz

3   was used.

4   An alternative physiological noise correction was performed using Conn

5   (http://www.conn-toolbox.org). Using a CompCor method (Behzadi *et al.*, 2007), regression of

6   the first 5 principal components of the white matter and cerebrospinal fluid (CSF) signal was

7   performed along with 6 movement parameters (Whitfield-Gabrieli and Nieto-Castanon, 2012).

8   Global signal regression was not performed. Subsequently, a bandpass filter of 0.01-0.10 Hz

9   was applied. Bivariate correlations were retrieved after a Fisher transform was performed to

10 improve the normality assumption of the distribution of the correlation coefficients (Whitfield-

11 Gabrieli and Nieto-Castanon, 2012).

12

13 **Connectome construction strategies**

14 First, a minimum spanning tree (MST) algorithm was applied to the raw weighted and

15 unthresholded connectivity matrices (CM) to retrieve a maximally sparse matrix while ensuring

16 all nodes are connected to the network. This means that graph subset is the matrix backbone

17 that has a minimum possible edge weight and does not include cycles. Subsequently,

18 connections were added back to the network in decreasing strength to obtain weighted

19 proportionally thresholded CM with a density of 5%, 10%, 15%, 20% and 25%. The second

20 series of 5 CM was obtained by binarization of the weighted and proportionally thresholded

21 CM (Rubinov and Sporns, 2010). This resulted in 1 weighted and unthresholded; 5 weighted

22 and proportionally thresholded; and 5 thresholded and binarized CM in total. Because these 11

23 CM were used for each of three parcellation atlases and two physiological noise correction

24 methods resulting in 66 connectomes per participant in total (see also Figure 1).

10

1 **Classification with support vector machine**

2 For each of the connectome as described above, the connectomes of pairwise matched HD gene

3 carriers and healthy controls were classified using Pattern Recognition for Neuroimaging

4 Toolbox v2.01 (PRoNTo), http://www.mlnl.cs.ucl.ac.uk/pronto (Schrouff *et al.*, 2013). Leave-

5 one-out cross-validation (100 permutations) was applied. A post-hoc analysis was performed

6 to assess effects of overall functional connectivity and motion (see Supplementary Methods).

7

8 **Modularity**

9 For each connectome construction strategy, the community affiliation vector was computed for

10 the group averaged CM of HD gene carriers and for the group averaged CM of HC, using the

11 Louvain algorithm for community detection (Blondel *et al.*, 2008). 250 iterations were

12 performed with a resolution parameter $\gamma = 1$ (default), followed by a consensus clustering

13 (Lancichinetti and Fortunato, 2012). The modules for brain networks were visualized with

14 BrainNet Viewer (Xia, Wang and He, 2013). Two related measures of modularity are the

15 participation coefficient (PC); a measure that quantifies how evenly the node's connections are

16 distributed across modules (Sporns and Betzel, 2016) and within module Z-score; a way to

17 express intra-modular connectivity (Fornito, Zalesky and Breakspear, 2015). The rationale for

18 using these measures was that it allows a quantitative comparison of the hubness of networks

19 in addition to a qualitative description of a network using modularity, which are both important

20 in functional imaging (Power and Schlaggar, 2013). PC and within module Z-score were

21 computed as implemented in the Brain Connectivity Toolbox version 2017-01

22 (https://sites.google.com/site/bctnet/). For these measures, differences between HD gene

23 carriers and HC were assessed using permutation testing (10 000 iterations). Age, gender, site

24 of acquisition, education and overall functional connectivity (FC) were regressed out as

25 covariates and false discovery rate (FDR) correction was applied. The nodes that were identified

11

1   as significant (p-value<0.05, FDR corrected) for the PC and within module Z-score separately

2   were contrasted in both the number as location of these nodes across the different connectome

3   construction strategies. Nodes were visualized with BrainNet Viewer (Xia, Wang and He,

4   2013).

Results

Classification of HD gene carriers and healthy controls

The first aim of this study was to assess which functional connectome construction method was most discriminative for a machine learning classification of HD gene carriers and healthy controls (HC). Regardless of thresholding or binarization, the classification for the pairwise matched cohort yielded a higher area under the curve (AUC) for physiological noise correction performed using Conn connectomes, ranging from 0.58 to 0.78, compared to Brainwaver connectomes, ranging from 0.32 to 0.57 (Figure 2). The AUC using the functional, multi-modal or structural parcellation atlas were 0.78, 0.70 and 0.65 respectively. For weighted matrices the functional atlas consistently gave for each threshold a higher AUC compared to the other atlases. The AUC for the functional atlas using Conn ranged from 0.69 to 0.78, compared to a range of 0.60 to 0.70 for the multi-modal and 0.59 to 0.65 for the structural atlas. For binarized matrices using Conn the AUC for the functional atlas ranged from 0.64 to 0.71, while the AUC ranges for the multi-modal and the structural atlas were 0.58 to 0.68 and 0.58 to 0.63 respectively. Therefore, thresholding or binarization had a minimal effect on the AUC.

There were no group differences in age, gender, acquisition site or education (Table 1). A replication was also performed using an alternative subdivision strategy in which a selection of cohorts was made based on median age instead of pairwise matching for gender and age. This replication showed similar patterns, with Conn connectomes showing an AUC ranging from 0.56 to 0.70, compared to Brainwaver connectomes yielding 0.32 to 0.52 (Supplementary Figure 1). Overall functional connectivity (van den Heuvel *et al.*, 2017) was corrected by rescaling all elements of a matrix with respect to their contribution to overall functional connectivity on group level (Supplementary Figures 2-4). After application of this correction, the same patterns in the results were obtained for classification, with Conn connectomes yielding an AUC from 0.57 to 0.76 and Brainwaver connectomes 0.32 to 0.57 (Supplementary

1   Figure 5). Exclusion of outliers in terms of mean framewise displacement did not have a

2   considerable effect on the classification, with the AUC for Conn connectomes ranging from

3   0.56 to 0.73 compared to Brainwaver connectomes ranging from 0.30 to 0.51 (Supplementary

4   Figures 6-7).

5

6   **Effect on modular organization**

7   We then investigated how modular organization was affected by different connectome

8   construction methods. The assignment of each node to a module was qualitatively compared

9   and quantitative group differences in participation coefficient (PC) and within module-Z score

10   were also investigated (see Materials and Methods).

11   Across all atlases, the nodes of modules for the Brainwaver connectomes showed less

12   spatial proximity compared to Conn connectomes. For Brainwaver connectomes, nodes of the

13   same module were more dispersed for both HD and healthy controls when compared to Conn

14   connectomes (Figure 3). The Supplementary Materials include all lists of the module

15   assignments for the three atlases (see also corresponding Supplementary Figures 9-11). The

16   most important effect of physiological noise correction on modularity was observed in the

17   functional atlas: for healthy controls, Brainwaver connectomes provided dispersed modules that

18   were highly different in number and location across thresholds, while Conn connectomes

19   showed nodes in spatial proximity that had a consistent modularity across thresholds (Figure

20   4). In addition, the functional atlas also gave rise to modules that were less demarcated due to

21   its lower resolution (121 ROIs), compared to the structural and multi-modal atlases (316 and

22   360 ROIs, see Figure 5).

23

24   For quantitative assessment of modularity, nodes that were significantly different

25   ($p<0.05$, FDR corrected) between HD and healthy controls were summed with respect to

14

1   participation coefficient (PC) and within module-Z score. Table 2 shows that the type of

2   physiological noise correction, the parcellation atlas, thresholding and binarization all had an

3   important effect on modularity group differences.

4   First, the effect of thresholding and binarization was investigated for Conn connectomes

5   using the multi-modal atlas. The weighted connectomes showed 46 significant nodes for PC at

6   a threshold of 5% and only 1 node at threshold 10% and zero at thresholds of 15%, 20% or 25%

7   (Figure 6a). After binarization an irregular pattern across the thresholds in both the number and

8   location of significant nodes was observed, which showed that binarization had a substantial

9   effect on group differences in modularity (Figure 6b).

10   In addition to the influence of binarization and thresholding, the effect of physiological

11   noise correction was also examined. For the structural atlas, Brainwaver connectomes showed

12   the following number of PC significant differences: 107 nodes at 5%, 4 nodes at 10%, 155

13   nodes at 15%, 215 at 20% and no nodes were significant at a threshold of 25%, or when no

14   threshold was applied. Conn showed 10 nodes at a threshold of 5% and 10%, zero nodes at 15%

15   and 20%, 26 nodes at 25% and no nodes were found significant when no threshold was applied

16   (Figure 6c-d).

17   In order to distinguish between the number of significant nodes and their location, the

18   within module Z-score using the functional atlas was further investigated because both types of

19   physiological noise correction yielded a similar number of nodes. Brainwaver showed 3 nodes

20   at 5%, 4 nodes at 10%, 3 nodes at 15%, 15 nodes at 20% and 25%, and 10 nodes were found

21   significant when no threshold was applied. Conn showed 9 nodes at 5%, 10 nodes at 10%, 11

22   nodes at 15%, 21 nodes at 20% and 25%, and 9 nodes when no threshold was applied.

23   Regarding the location of significant nodes, there was no overlap when comparing Brainwaver

24   and Conn connectomes. When no threshold was applied, the 10 nodes detected in the

25   Brainwaver connectome did not overlap with the 9 nodes detected in the Conn connectome.

1   There was also no overlap in the nodes at the thresholds 5%, 10%, 15% or 25%. Only at a

2   threshold of 20% one single node (showed in yellow) was in common from 15 and 21 nodes

3   detected using Brainwaver and Conn connectomes respectively (Figure 6e-f).

4

5   **Discussion**

6   The aim of this work was to perform a systematic comparison of connectome construction

7   strategies to provide recommendations for representing functional networks in

8   neurodegeneration. The following aspects of connectome construction strategies were

9   investigated: principal component-based physiological noise correction versus wavelet

10   decomposition, the type of parcellation atlas (functional, structural and multi-modal), weighted

11   versus binarized networks, and unthresholded versus proportionally thresholded networks.

12   For the classification of HD gene carriers and healthy controls, the connectomes making

13   use of principal component-based physiological noise correction yielded higher discrimination

14   rates compared to a physiological noise correction with wavelet decomposition. When a

15   functional atlas was used in combination with principal component-based correction, the

16   highest discrimination rates were obtained. The operation of thresholding or binarization did

17   not enhance classification performance. The classification was not driven by group differences

18   in gender, age, education, site of acquisition, overall functional connectivity or motion. The

19   findings were consistent across two strategies of cohort subdivision ensuring equal sample size.

20   The modular organization was most anatomically coherent and consistent across

21   thresholds when using principal component-based physiological noise correction. Connectomes

22   obtained after wavelet decomposition for physiological noise correction resulted in dispersed

23   modules, and these modules were different again across thresholds when a functional atlas was

24   used. While this could represent complete modular breakdown in HD, the observation of

16

1   dispersed modules in healthy controls does not support this (Sporns, 2011), suggesting this is

2   methodological rather than biological.

3   Group differences in modular organization in terms of participation coefficient and

4   within module Z-score were highly affected by the type of physiological noise correction and

5   by the threshold applied, whether or not combined with binarization. There was no relationship

6   between the percentage of thresholding and the number of nodes that were significantly

7   different between HD gene carriers and healthy controls. Moreover, steep changes in both the

8   number of nodes and their location were observed when incrementally increasing the threshold.

9   Whereas thresholding and binarization did not increase the ability to discriminate

10 between groups when classifying, these operations showed a strong effect on identification of

11 modular group differences. This was not keeping in with former findings of the reliability of

12 modular organization in thresholded matrices, be it with a larger sparsity (He *et al.*, 2009;

13 Meunier, 2009), but was in agreement with the caution around thresholding (Scheinost *et al.*,

14 2012; Garrison *et al.*, 2015), and the relevance of weaker connections (Gallos, Makse and

15 Sigman, 2012; Santarnecchi *et al.*, 2014; Goulas, Schaefer and Margulies, 2015). These

16 findings together suggest that thresholding, whether or not combined with binarization, has a

17 risk to introduce artificial group differences, especially when considering that there is no

18 standard to choose the threshold. One way to overcome the arbitrariness of choosing a threshold

19 and to enhance the reliable detection of network group differences, may be the use of

20 unthresholded weighted matrices.

21 While machine learning can have high diagnostic accuracy in schizophrenia

22 (Davatzikos *et al.*, 2005) or Alzheimer's disease (Klöppel *et al.*, 2008), it is of limited utility

23 because the diagnosis can usually be made clinically in such cases. However, if the diagnosis

24 is known *a priori*, machine learning can be useful to indicate which imaging methodology

25 yields the highest discrimination rate. In the present study, the diagnosis of Huntington's

17

1   disease was made with genetic testing so could be used as a gold standard, and so machine

2   learning can indicate the most reliable connectome construction strategy. However, the

3   robustness of the use of machine learning in connectomics is under further investigation (Brown

4   and Hamarneh, 2016). The use of proposed null models for brain networks (Rubinov and

5   Sporns, 2011; Hosseini and Kesler, 2013), test-retest analysis (Du *et al.*, 2015; Marchitelli *et*

6   *al.*, 2016), exploring different confidence measures of classification prediction (Gammerman

7   and Vovk, 2007; Nouretdinov *et al.*, 2011), different types of datasets and the use of different

8   types of machine learning algorithms such as Scikit-learn or its recent neuroimaging equivalent

9   Nilearn (Abraham *et al.*, 2014), are all but a few methods to assess the generality of application

10 of machine learning in connectomics.

11   The assessment of modular organization is complimentary to machine learning

12 classification, as it enables direct mapping of modules on the brain. For both of these

13 methodologies, the functional atlas gave rise to the highest classification performances, and it

14 was also the most stable atlas across thresholds in terms of modularity for Conn based

15 connectomes. Potentially, this is due to the size of the atlas. This is important as currently there

16 is no consensus on the most reliable type of parcellation atlas (Arslan *et al.*, 2017). Another

17 explanation for the performance of the functional atlas could be that the multi-modal and

18 structural atlas did not include sub-cortical regions, as for instance is the case for the multi-

19 modal Brainnetome atlas (Fan *et al.*, 2016). However, the specific way how an atlas is

20 composed, remains an issue for its application. For example, Arslan and colleagues cautioned

21 that the Glasser multi-modal atlas may be positively biased toward the specific modalities (task

22 fMRI, myelin content) used for its composition (Arslan *et al.*, 2017), and other atlases may

23 suffer from the same biases. Recently, gradient-weighted Markov Random Field models have

24 been used to account for these biases by detecting changes in functional connectivity

18

1   similarities, showing that these multiresolution parcellations may be more homogenous than

2   other multimodal atlases, including the Glasser atlas (Schaefer *et al.*, 2017).

3   One limitation of our study is that it does not provide a standard to assess which

4   construction provides the best topological organization of the connectome. While replication of

5   our findings in other neurodegenerative diseases will be necessary to generalize our current

6   findings to a broader field, the strength of Huntington's disease as a model for

7   neurodegenerative diseases is that the diagnosis is technically robust and accurate. This implies

8   that cohort of HD gene carriers consisted of both participants with manifest symptoms and

9   participants at a pre-manifest stage, sometimes decades before the onset of disease. When

10 assessing group differences in functional connectomes of HD gene carriers and controls, genetic

11 testing thus served as a gold standard.

12   Our work suggests that principal component-based physiological noise correction

13 outperformed wavelet decomposition. An in depth analysis of these and other types of

14 physiological noise correction, such as Bayesian methods (Särkkä *et al.*, 2012) or independent

15 component analysis (Griffanti *et al.*, 2014) is beyond the scope of this study. The novelty in

16 this systematic comparison is that it combined the strengths of machine learning and

17 investigation of modular organization. Our work also considered a variety of factors from

18 different levels (type of parcellation atlas, physiological noise correction, choice of threshold

19 and binarization) in connectome construction to assess their mutual effects on the connectome,

20 instead of limiting the investigation to a single factor alone. A reproducibility analysis of our

21 methodology would be welcome, to verify how consistent the results are across different

22 datasets.

Conclusions

We performed a systematic comparison of various connectome construction strategies based on resting state fMRI data from TrackOn-HD project. Principal component-based physiological noise correction resulted in a higher group discrimination rate, and the use of a functional atlas outperformed other atlases. We showed that the type of physiological noise correction, the parcellation atlas, thresholding and binarization all have a substantial effect on the detection of group differences based on modularity. This raises important methodological concerns for functional connectomics in general. The results of the present study support the use of unthresholded and weighted matrices in combination with principal component-based noise correction methods. While the use of a functional atlas may provide more consistent results, a meticulously made trade-off with higher resolution atlases may also be necessary.

Acknowledgements

Track-On HD investigators

The Track-On HD Investigators are A. Coleman, J. Decolongon, M. Fan, T. Petkau (University of British Columbia, Vancouver, Canada); C. Jauffret,D. Justo, S. Lehericy, K. Nigaud, R. Valabrègue (ICM and APHP, Pitié-Salpêtrière University Hospital, Paris, France). A. Schoonderbeek,E. P. 't Hart (Leiden University Medical Centre, Leiden, the Netherlands);D. J. Hensman Moss, R. Ghosh, H. Crawford, M. Papoutsi, C. Berna,D. Mahaleskshmi (University College London, London, United Kingdom). R. Reilmann, N. Weber (George Huntington Institute, Münster, Germany); I. Labuschagne, J. Stout (Monash University, Melbourne, Australia); B. Landwehrmeyer, M. Orth, I. Mayer (University of Ulm, Ulm, Germany); H. Johnson (University of Iowa, Iowa City, IA); D. Crawfurd (University of Manchester, Manchester, United Kingdom).

References

Abraham, A. *et al.* (2014) 'Machine learning for neuroimaging with scikit-learn.', *Frontiers in neuroinformatics*, 8(February), p. 14. doi: 10.3389/fninf.2014.00014.

Achard, S. *et al.* (2006) 'A Resilient, Low-Frequency, Small-World Human Brain Functional Network with Highly Connected Association Cortical Hubs', *Journal of Neuroscience*, 26(1), pp. 63–72. doi: 10.1523/JNEUROSCI.3874-05.2006.

Alexander-Bloch, A. F. *et al.* (2010) 'Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia.', *Frontiers in systems neuroscience*, 4(October), p. 147. doi: 10.3389/fnsys.2010.00147.

1   Arslan, S. *et al.* (2017) 'Human brain mapping: A systematic comparison of parcellation
2   methods for the human cerebral cortex', *NeuroImage*. Elsevier. doi:
3   10.1016/j.neuroimage.2017.04.014.

4   Behzadi, Y. *et al.* (2007) 'A component based noise correction method (CompCor) for
5   BOLD and perfusion based fMRI', *NeuroImage*, 37(1), pp. 90–101. doi:
6   10.1016/j.neuroimage.2007.04.042.

7   Blondel, V. D. *et al.* (2008) 'Fast unfolding of communities in large networks', *Journal*
8   *of Statistical Mechanics: Theory and Experiment*, 2008(10), p. P10008. doi: 10.1088/1742-
9   5468/2008/10/P10008.

10   Brown, C. J. and Hamarneh, G. (2016) 'Machine Learning on Human Connectome Data
11   from MRI', *Arxiv preprint*. Available at: http://arxiv.org/abs/1611.08699.

12   Bullmore, E. and Sporns, O. (2009) 'Complex brain networks: graph theoretical analysis
13   of structural and functional systems', *Nature Publishing Group*, 10(3), pp. 186–198. doi:
14   10.1038/nrn2575.

15   Carbonell, F., Bellec, P. and Shmuel, A. (2014) 'Quantification of the impact of a
16   confounding variable on functional connectivity confirms anti-correlated networks in the
17   resting-state', *NeuroImage*. Elsevier Inc., 86, pp. 343–353. doi:
18   10.1016/j.neuroimage.2013.10.013.

19   Choi, E. Y., Yeo, B. T. T. and Buckner, R. L. (2012) 'The organization of the human
20   striatum estimated by intrinsic functional connectivity.', *Journal of neurophysiology*, 108(8),
21   pp. 2242–63. doi: 10.1152/jn.00270.2012.

22   Cordes, D. *et al.* (2001) 'Frequencies contributing to functional connectivity in the
23   cerebral cortex in "resting-state" data', *American Journal of Neuroradiology*, 22(7), pp. 1326–
24   1333.

25   Daubechies, I. (1988) 'Orthonormal bases of compactly supported wavelets',
26   *Communications on Pure and Applied Mathematics*, 41(7), pp. 909–996. doi:
27   10.1002/cpa.3160410705.

28   Davatzikos, C. *et al.* (2005) 'Whole-Brain Morphometric Study of Schizophrenia
29   Revealing a Spatially Complex Set of Focal Abnormalities', *Archives of General Psychiatry*,
30   62(11), p. 1218. doi: 10.1001/archpsyc.62.11.1218.

31   Desikan, R. S. *et al.* (2006) 'An automated labeling system for subdividing the human
32   cerebral cortex on MRI scans into gyral based regions of interest', *NeuroImage*, 31(3), pp. 968–
33   980. doi: 10.1016/j.neuroimage.2006.01.021.

34   Du, H. X. *et al.* (2015) 'Test-Retest Reliability of Graph Metrics in High-resolution
35   Functional Connectomics: A Resting-State Functional MRI Study', *CNS Neuroscience and*
36   *Therapeutics*, 21(10), pp. 802–816. doi: 10.1111/cns.12431.

37   Fan, L. *et al.* (2016) 'The Human Brainnetome Atlas: A New Brain Atlas Based on
38   Connectional Architecture', *Cerebral Cortex*, 26(8), pp. 3508–3526. doi:
39   10.1093/cercor/bhw157.

40   Fornito, A., Zalesky, A. and Breakspear, M. (2015) 'The connectomics of brain
41   disorders', *Nature Reviews Neuroscience*. Nature Publishing Group, 16(3), pp. 159–172. doi:

10.1038/nrn3901.

Friston, K. J. *et al.* (2011) *Statistical parametric mapping: the analysis of functional brain images: the analysis of functional brain images*, *Functional neuroimaging: Technical*. London: Elsevier Academic Press.

Gallos, L. K., Makse, H. A. and Sigman, M. (2012) 'A small world of weak ties provides optimal global integration of self-similar modules in functional brain networks.', *Proceedings of the National Academy of Sciences of the United States of America*, 109(8), pp. 2825–30. doi: 10.1073/pnas.1106612109.

Gammerman, A. and Vovk, V. (2007) 'Hedging Predictions in Machine Learning The Second Computer Journal Lecture ', *The Computer Journal*, 50(2), pp. 151–163. Available at: http://dx.doi.org/10.1093/comjnl/bxl065.

Garrison, K. A. *et al.* (2015) 'The (in)stability of functional brain network measures across thresholds', *NeuroImage*, 118(203), pp. 651–661. doi: 10.1016/j.neuroimage.2015.05.046.

Glasser, M. F. *et al.* (2016) 'A Multi-Modal Parcellation of Human Cerebral Cortex', *Nature*, 536(7615), pp. 171–178. doi: 10.1038/nature18933.A.

Goulas, A., Schaefer, A. and Margulies, D. S. (2015) 'The strength of weak connections in the macaque cortico-cortical network', *Brain Structure and Function*. Springer Berlin Heidelberg, 220(5), pp. 2939–2951. doi: 10.1007/s00429-014-0836-3.

Griffanti, L. *et al.* (2014) 'ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging', *NeuroImage*, 95, pp. 232–247. doi: 10.1016/j.neuroimage.2014.03.034.

He, Y. *et al.* (2009) 'Uncovering intrinsic modular organization of spontaneous brain activity in humans', *PLoS ONE*, 4(4), pp. 23–25. doi: 10.1371/journal.pone.0005226.

van den Heuvel, M. P. *et al.* (2017) 'Proportional thresholding in resting-state fMRI functional connectivity networks and consequences for patient-control connectome studies: Issues and recommendations', *NeuroImage*. Elsevier, 152(February), pp. 437–449. doi: 10.1016/j.neuroimage.2017.02.005.

Hosseini, S. M. H. and Kesler, S. R. (2013) 'Influence of Choice of Null Network on Small-World Parameters of Structural Correlation Networks', *PLoS ONE*, 8(6). doi: 10.1371/journal.pone.0067354.

Klöppel, S. *et al.* (2008) 'Automatic classification of MR scans in Alzheimer's disease', *Brain*, 131(3), pp. 681–689. doi: 10.1093/brain/awm319.

Klöppel, S. *et al.* (2015) 'Compensation in Preclinical Huntington's Disease: Evidence From the Track-On HD Study', *EBioMedicine*. Elsevier B.V., 2(10), pp. 1420–1429. doi: 10.1016/j.ebiom.2015.08.002.

Lancichinetti, A. and Fortunato, S. (2012) 'Consensus clustering in complex networks.', *Scientific reports*, 2, p. 336. doi: 10.1038/srep00336.

Marchitelli, R. *et al.* (2016) 'Test-retest reliability of the default mode network in a multi-centric fMRI study of healthy elderly: Effects of data-driven physiological noise

correction techniques', *Human Brain Mapping*, 37(6), pp. 2114–2132. doi: 10.1002/hbm.23157.

McColgan, P. *et al.* (2015) 'Selective vulnerability of Rich Club brain regions is an organizational principle of structural connectivity loss in Huntington's disease', *Brain*, 138(11), pp. 3327–3344. doi: 10.1093/brain/awv259.

McColgan, P. *et al.* (2017) 'Structural and functional brain network correlates of depressive symptoms in premanifest Huntington's disease', *Human Brain Mapping*, 2829(March), pp. 2819–2829. doi: 10.1002/hbm.23527.

McColgan, P. and Tabrizi, S. J. (2018) 'Huntington's disease: a clinical review', *European Journal of Neurology*, 25(1), pp. 24–34. doi: 10.1111/ene.13413.

Meunier, D. (2009) 'Hierarchical modularity in human brain functional networks', *Frontiers in Neuroinformatics*, 3(October), pp. 1–12. doi: 10.3389/neuro.11.037.2009.

Modat, M. *et al.* (2010) 'Fast free-form deformation using graphics processing units', *Computer Methods and Programs in Biomedicine*, 98(3), pp. 278–284. doi: http://dx.doi.org/10.1016/j.cmpb.2009.09.002.

Nouretdinov, I. *et al.* (2011) 'Machine learning classification with confidence: Application of transductive conformal predictors to MRI-based diagnostic and prognostic markers in depression', *NeuroImage*. Elsevier Inc., 56(2), pp. 809–813. doi: 10.1016/j.neuroimage.2010.05.023.

Power, J. and Schlaggar, B. (2013) 'Evidence for hubs in human functional brain networks', *Neuron*, 79(4), pp. 1–29. doi: 10.1016/j.neuron.2013.07.035.Evidence.

Qi, S. *et al.* (2015) 'The influence of construction methodology on structural brain network measures: A review', *Journal of Neuroscience Methods*. Elsevier B.V., 253(500), pp. 170–182. doi: 10.1016/j.jneumeth.2015.06.016.

Richiardi, J. *et al.* (2011) 'Decoding brain states from fMRI connectivity graphs', *NeuroImage*. Elsevier Inc., 56(2), pp. 616–626. doi: 10.1016/j.neuroimage.2010.05.081.

Romero-Garcia, R. *et al.* (2012) 'Effects of network resolution on topological properties of human neocortex', *NeuroImage*. Elsevier Inc., 59(4), pp. 3522–3532. doi: 10.1016/j.neuroimage.2011.10.086.

Rubinov, M. and Sporns, O. (2010) 'Complex network measures of brain connectivity: Uses and interpretations', *NeuroImage*. Elsevier Inc., 52(3), pp. 1059–1069. doi: 10.1016/j.neuroimage.2009.10.003.

Rubinov, M. and Sporns, O. (2011) 'Weight-conserving characterization of complex functional brain networks', *NeuroImage*. Elsevier Inc., 56(4), pp. 2068–2079. doi: 10.1016/j.neuroimage.2011.03.069.

Santarnecchi, E. *et al.* (2014) 'Efficiency of weak brain connections support general cognitive functioning', *Human Brain Mapping*, 35(9), pp. 4566–4582. doi: 10.1002/hbm.22495.

Särkkä, S. *et al.* (2012) 'Dynamic retrospective filtering of physiological noise in BOLD fMRI: DRIFTER', *NeuroImage*. Elsevier Inc., 60(2), pp. 1517–1527. doi:

1 10.1016/j.neuroimage.2012.01.067.

2 Schaefer, A. *et al.* (2017) 'Local-Global Parcellation of the Human Cerebral Cortex
3 from Intrinsic Functional Connectivity MRI', *Cerebral Cortex*, pp. 1–20. doi:
4 10.1093/cercor/bhx179.

5 Scheinost, D. *et al.* (2012) 'The intrinsic connectivity distribution: A novel contrast
6 measure reflecting voxel level functional connectivity', *NeuroImage*. Elsevier Inc., 62(3), pp.
7 1510–1519. doi: 10.1016/j.neuroimage.2012.05.073.

8 Schrouff, J. *et al.* (2013) 'PRoNTo: Pattern recognition for neuroimaging toolbox',
9 *Neuroinformatics*, 11(3), pp. 319–337. doi: 10.1007/s12021-013-9178-1.

10 Sporns, O. (2011) 'The human connectome: A complex network', *Annals of the New*
11 *York Academy of Sciences*, 1224(1), pp. 109–125. doi: 10.1111/j.1749-6632.2010.05888.x.

12 Sporns, O. and Betzel, R. F. (2016) 'Modular brain networks.', *Annual Review of*
13 *Psychology*, 67(1), pp. 613–640. doi: 10.1146/annurev-psych-122414-033634.

14 Vértes, P. E. *et al.* (2016) 'Gene Transcription Profiles Associated with Inter-modular
15 Hubs and Connection Distance in Human Functional Magnetic Resonance Imaging Networks.',
16 *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*,
17 371(1705), pp. 735–769. doi: 10.1098/rstb.2015.0362.

18 Whitfield-Gabrieli, S. and Nieto-Castanon, A. (2012) '*Conn*: A Functional
19 Connectivity Toolbox for Correlated and Anticorrelated Brain Networks', *Brain Connectivity*,
20 2(3), pp. 125–141. doi: 10.1089/brain.2012.0073.

21 Xia, M., Wang, J. and He, Y. (2013) 'BrainNet Viewer: A Network Visualization Tool
22 for Human Brain Connectomics', *PLoS ONE*, 8(7). doi: 10.1371/journal.pone.0068910.

23 Yeo, B. T. T. *et al.* (2011) 'The organization of the human cerebral cortex estimated by
24 intrinsic functional connectivity', *Journal of neurophysiology*, 106, pp. 1125–1165. doi:
25 10.1152/jn.00338.2011.

26 Ziegler, E. *et al.* (2013) 'Altered White Matter Architecture in BDNF Met Carriers',
27 *PLoS ONE*, 8(7). doi: 10.1371/journal.pone.0069290.

28

Tables

Table 1: Cohort demographics. Values are mean (standard deviation) or n/n. See text for composition strategy.

	HD gene carrier	Control	Statistical test	*P*-value
Total cohort after quality control				
N	86	79	—	—
Age (SD)	45.8 (9.1)	51.7 (9.9)	2 sample *t*-test, 2 tailed	<0.01***
Gender (F/M)	41/45	47/32	Chi-square	0.12
Education (2/3/4/5/6)	6/22/21/36/1	10/16/28/23/2	Wilcoxon-Mann-Whitney test	0.30
Study site (N) (Leiden/London/Paris/Vancouver)	21/22/25/18	24/16/24/15	Chi-square	0.77
Pairwise matched cohort				
N	49	49	—	—
Age (SD)	48.2 (7.9)	48.4 (8.0)	2 sample *t*-test, 2 tailed	0.93
Gender (F/M)	28/21	28/21	—	—
Education (2/3/4/5/6)	3/10/17/19/0	7/9/18/13/2	Wilcoxon-Mann-Whitney test	0.38
Study site (N) (Leiden/London/Paris/Vancouver)	15/11/14/9	12/13/14/10	Chi-square	0.91

Table 2: Number of nodes with a different modularity property for Brainwaver and Conn connectomes. For each connectome construction strategy, the connectomes were assessed for differences in participation coefficient (PC) or within module Z-score (Z-score) between HD gene carriers and healthy controls with permutation testing (10 000 iterations). The sums of all significant nodes (p-value < 0.05, FDR corrected) are reported, regardless of the direction of the effect.

			Binarized connectomes					Weighted connectomes					
		Threshold (%)	5	10	15	20	25	5	10	15	20	25	none
Brainwaver	**Structural atlas**	Z-score	8	9	36	42	32	16	1	6	37	1	85
		PC	32	18	182	243	25	107	4	155	215	0	0
	Functional atlas	Z-score	2	5	1	5	8	3	4	3	15	15	10
		PC	2	0	0	1	1	0	0	16	0	0	0
	Multi-modal atlas	Z-score	0	0	2	0	0	0	0	1	0	0	3
		PC	0	0	0	0	0	0	0	0	0	0	0
	Total differences		44	32	221	291	66	126	9	181	267	16	98
Conn	**Structural atlas**	Z-score	12	20	3	0	0	23	3	27	17	10	3
		PC	13	0	0	0	14	10	10	0	0	26	0
	Functional atlas	Z-score	6	10	18	11	5	9	10	11	21	21	9
		PC	20	12	16	29	18	21	10	14	10	10	0
	Multi-modal atlas	Z-score	35	30	30	33	24	35	46	33	27	32	19
		PC	45	11	0	159	80	46	1	0	0	0	0
	Total differences		131	83	67	232	141	144	80	85	75	99	31

26

1 **Figures**

2 **Figure 1: Overview systematic comparison of connectome construction strategies.** Rs-
3 fMRI was obtained from 86 HD gene carriers and 79 healthy controls. Functional images were
4 realigned and then coregistered to anatomical images. ROIs of three parcellation atlases were
5 coregistered to the new anatomical images. For the use of each atlas a wavelet decomposition
6 or component-based correction was applied for physiological noise correction. This resulted in
7 weighted unthresholded matrices, whereof the minimum spanning tree (MST) is retrieved and
8 subsequently, the strongest connections are added back to obtain a range of proportionally
9 thresholded matrices. Finally, each of these thresholded matrices have been binarized. The
10 following pictures are adapted from original publications: MRI acquisition (McColgan,
11 Gregory, et al. 2017); functional (Yeo et al. 2011), structural (Romero-Garcia et al. 2012) and
12 multi-modal (Glasser et al. 2016) parcellation atlas. HC: healthy controls. HD: Huntington's
13 disease. MST: minimum spanning tree algorithm. ROI: region of interest. rs-fMRI: resting-state
14 functional Magnetic Resonance Imaging. SPM8: statistical parametric mapping. T1WI: T1-
15 weighted imaging.
16
17 **Figure 2: AUC for pairwise matched cohort.** The connectivity matrices of 49 HD gene
18 carriers and 49 healthy controls pairwise, matched for gender and age, are classified with
19 PRoNTo in a leave-one-out cross-validation. The area under the curve (AUC) is obtained for
20 each of 66 connectome construction strategies. The structural atlas is based on anatomical
21 demarcations (Romero-Garcia et al. 2012), the functional atlas is based on intrinsic connectivity
22 of brain regions (Yeo et al. 2011), and the multi-modal atlas made use of both criteria (Glasser
23 et al. 2016). Physiological noise correction was either performed with Conn, which is a
24 component based correction (Whitfield-Gabrieli & Nieto-Castanon 2012), or with Brainwaver,
25 which makes use of a wavelet decomposition (Achard et al. 2006). A proportional threshold
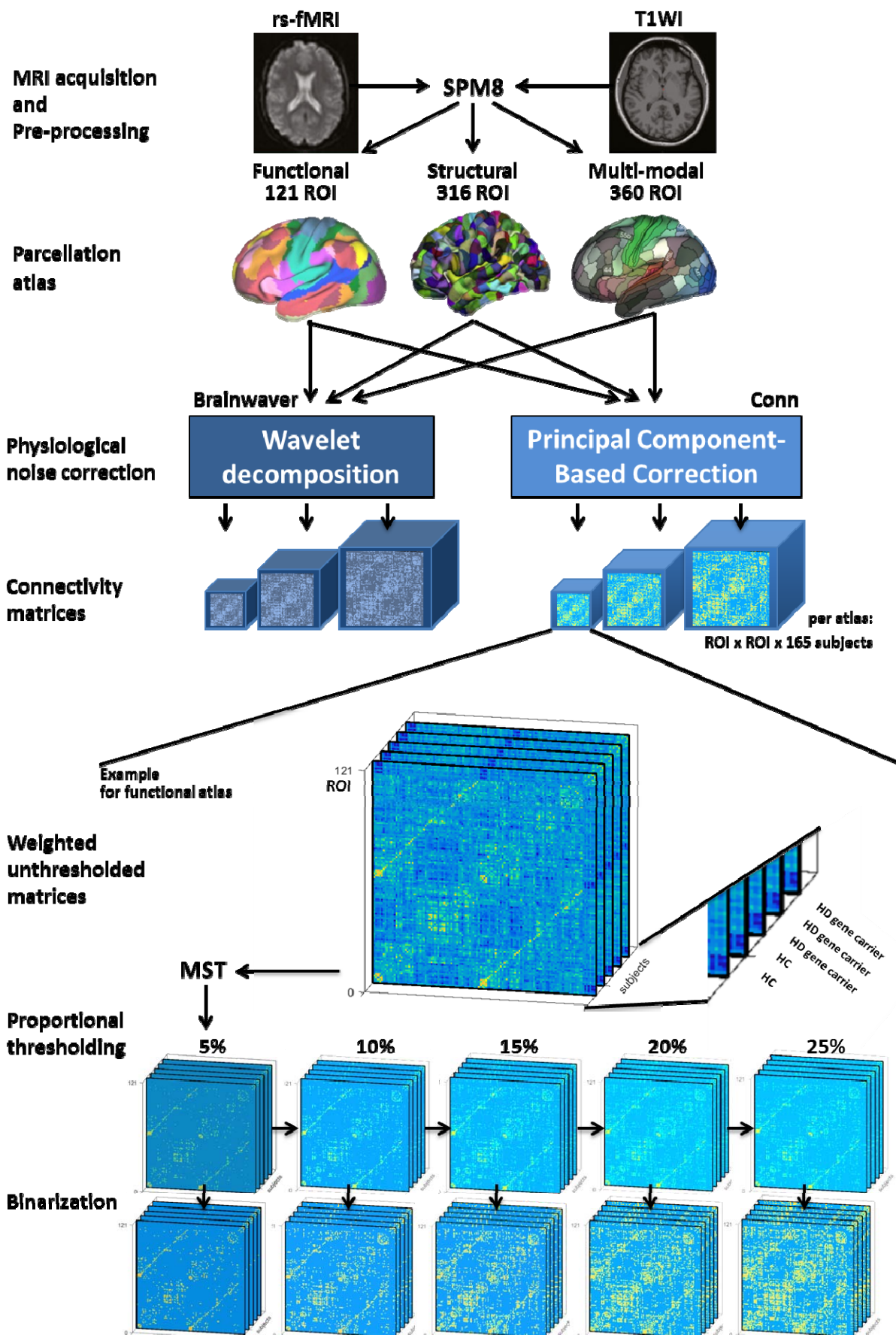26 was applied on weighted (A) or binarized (B) connectivity matrices.
27
28 **Figure 3: The nodes of the modules are more dispersed for Brainwaver connectomes,**
29 **while anatomically coherent for Conn connectomes.** The visualization of the group
30 community affiliation vector for the HD gene carriers and healthy controls is based on weighted
31 unthresholded matrices. In all instances, 3 modules were detected by the Louvain community
32 detection algorithm. For Conn connectomes these modules seem to correspond with a posterior
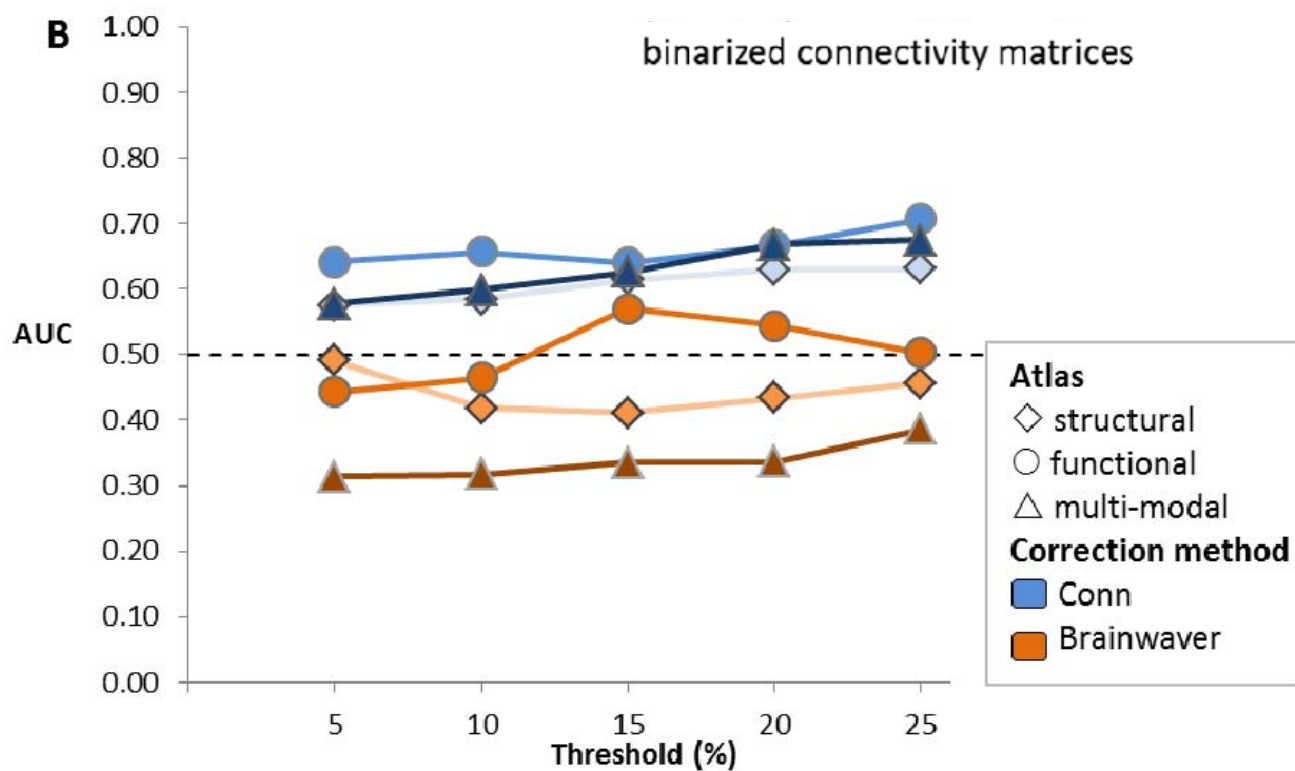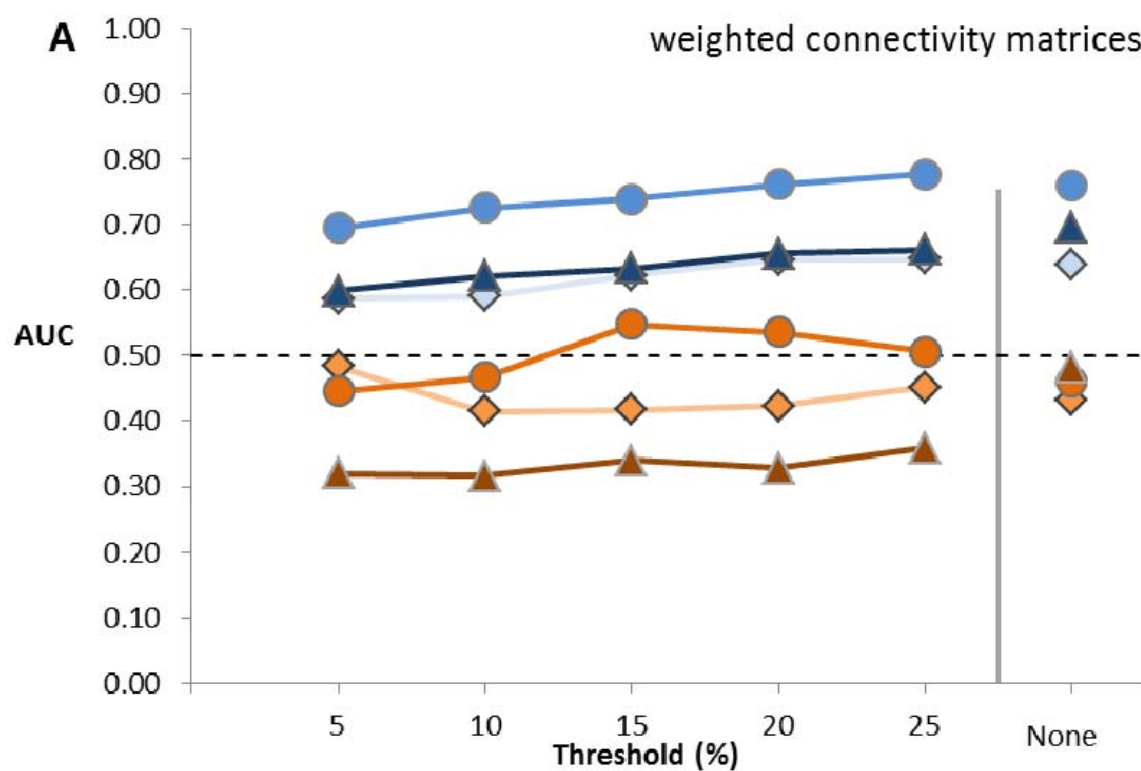33 (red), fronto-temporal (green) and fronto-parietal network (yellow).
34
35 **Figure 4: The modular organization is inconsistent in healthy controls across thresholds**
36 **for Brainwaver connectomes, while consistent for Conn connectomes.** The modules
37 detected in weighted connectomes are visualized. A functional atlas was used for both types of
38 physiological noise correction.
39
40 **Figure 5: Comparison of modularity on HD gene carrier group level across atlases.** The
41 visualization of the group community affiliation vector for the HD gene carriers and healthy
42 controls is based on weighted unthresholded matrices. There were 4 modules detected using the
43 functional atlas (121 ROI), and 3 in the structural and multi-modal atlas (316 and 360 ROI).
44 The distribution of the nodes of a module is less informative in a lower resolution atlas
45 compared to higher resolution.
46
47 **Figure 6: Number and location of nodes that differ modularity between HD and healthy**
48 **controls are substantially effected by thresholding, binarization and type of physiological**
49 **noise correction.** The number of nodes that had differences in participation coefficient (PC) or

1 module Z-score (Z-score) between HD gene carriers and healthy controls, are visualized in a
2 dorsal view (permutation testing, p-value < 0.05, FDR corrected). This number is for each
3 comparison given in the bottom right corner. Binarization and thresholding have a strong
4 influence on detection of group difference: example shown for weighted connectivity matrices
5 (A) and the binarized equivalent (B) using a multi-modal atlas. Connectomes after wavelet
6 decomposition (Brainwaver) versus component-based correction (Conn) for physiological
7 noise lead to substantially different outcomes (C and D): example shown for weighted matrices,
8 using a structural atlas. When number of detected nodes are comparable, connectomes corrected
9 with Brainwaver or Conn do not lead to overlap in location (E and F): example shown for a
10 functional atlas, using weighted matrices. From all pairwise comparisons between E and F, only
11 1 single node is in common between two methods of physiological noise correction (yellow
12 node, at threshold of 20%). For full results, see Table 2.

MRI acquisition and Pre-processing

rs-fMRI → SPM8 ← T1WI

Functional 121 ROI · Structural 316 ROI · Multi-modal 360 ROI

Parcellation atlas

Brainwaver · Conn

Physiological noise correction

Wavelet decomposition · Principal Component-Based Correction

Connectivity matrices

per atlas: ROI x ROI x 165 subjects

Weighted unthresholded matrices

Example for functional atlas

121 ROI

subjects

HD gene carrier
HD gene carrier
HD gene carrier
HC
HC

MST ←

Proportional thresholding

5% · 10% · 15% · 20% · 25%

Binarization

A weighted connectivity matrices

B binarized connectivity matrices

Atlas
◇ structural
○ functional
△ multi-modal
Correction method
■ Conn
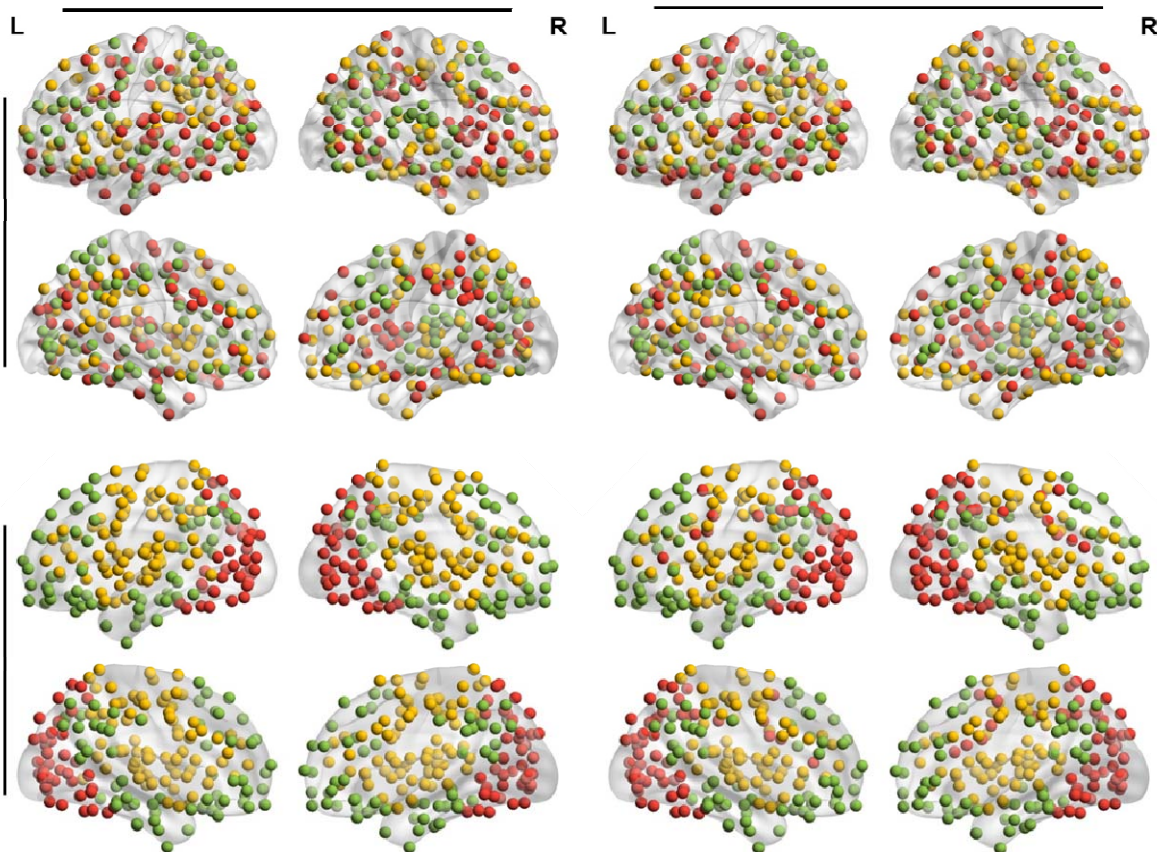■ Brainwaver

Atlas: multi-modal

HD gene carriers

healthy controls

Physiological noise correction:

Brainwaver

Conn

L   R   L   R

Atlas: functional

Brainwaver   Conn

L   R L   R

Threshold:

5%

10%

15%

20%

25%

None

Physiological noise correction: Conn

Atlas: Functional

HD gene carriers

L   R

Structural

Multi-modal

Threshold:

| | 5% | 10% | 15% | 20% | 25% | None |

A Weighted

Modularity: PC
Atlas: multi-modal
Physiological noise
correction: Conn

46   1   0   0   0   0

B Binarized

45   11   0   159   80

C Brainwaver

Modularity: PC
Atlas: structural
Matrices: weighted

107   4   155   215   0   0

D Conn

10   10   0   0   26   0

E Brainwaver

Modularity: Z-score
Atlas: functional
Matrices: weighted

3   4   3   15   15   10

F Conn

9   10   11   21   21   9