◆ Human Brain Mapping 36:4880–4896 (2015) ◆

Identification of Infants at High-Risk for Autism Spectrum Disorder Using Multiparameter Multiscale White Matter Connectivity Networks

Yan Jin,[1] Chong-Yaw Wee,[1] Feng Shi,[1] Kim-Han Thung,[1] Dong Ni,[2]* Pew-Thian Yap,[1] and Dinggang Shen[1,3]*

[1]*Biomedical Research Imaging Center, Department of Radiology, School of Medicine, University of North Carolina at Chapel Hill, North Carolina*
[2]*The Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, Department of Biomedical Engineering, Shenzhen University, China*
[3]*Department of Brain and Cognitive Engineering, Korea University, Seoul, Republic of Korea*

Abstract: Autism spectrum disorder (ASD) is a wide range of disabilities that cause life-long cognitive impairment and social, communication, and behavioral challenges. Early diagnosis and medical intervention are important for improving the life quality of autistic patients. However, in the current practice, diagnosis often has to be delayed until the behavioral symptoms become evident during childhood. In this study, we demonstrate the feasibility of using machine learning techniques for identifying high-risk ASD infants at as early as six months after birth. This is based on the observation that ASD-induced abnormalities in white matter (WM) tracts and whole-brain connectivity have already started to appear within 24 months after birth. In particular, we propose a novel multikernel support vector machine classification framework by using the connectivity features gathered from WM connectivity networks, which are generated via multiscale regions of interest (ROIs) and multiple diffusion statistics such as fractional anisotropy, mean diffusivity, and average fiber length. Our proposed framework achieves an accuracy of 76% and an area of 0.80 under the receiver operating characteristic curve (AUC), in comparison to the accuracy of 70% and the AUC of 70% provided by the best single-parameter single-scale network. The improvement in accuracy is mainly due to the complementary information provided by multiparameter multiscale networks. In addition, our framework also provides the potential imaging connectomic markers and an objective means for early ASD diagnosis. *Hum Brain Mapp 36:4880–4896, 2015.* © 2015 Wiley Periodicals, Inc.

Key words: Autism spectrum disorder; classification; connectivity networks; diffusion weighted imaging; infant

Contract grant sponsor: National Institute of Health grants; Contract grant numbers: EB006733, EB008374, EB009634, AG041721, MH100217, and AG042599; Contract grant sponsor: National Natural Science Foundation of China grant; Contract grant number: 61571304.

Correction added on 06 October 2015, after first online publication.

*Correspondence to: Dinggang Shen, Ph.D., Biomedical Research Imaging Center, 130 Mason Farm Road, Suite 3117, Chapel Hill, NC 27599. E-mail: dinggang_shen@med.unc.edu or Dong Ni, Ph.D., Department of Biomedical Engineering, Shenzhen University, Shenzhen 518060, China. E-mail: nidong@szu.edu.cn

Received for publication 10 May 2015; Revised 27 July 2015; Accepted 20 August 2015.

DOI: 10.1002/hbm.22957

Published online 14 September 2015 in Wiley Online Library (wileyonlinelibrary.com).

INTRODUCTION

Autism spectrum disorder (ASD) is a group of developmental disabilities that can cause difficulties in cognitive wellness, social communication, and social interaction. About 1 percent of world population is affected by ASD. It is usually first diagnosed in childhood. In the US, 1 out of 68 children has ASD. Prevalence of ASD in children in the US has increased by 120% since 2000. Now ASD becomes one of the fastest-growing developmental disabilities, which results in approximately annual $250 billion healthcare expenses in the US [Centers for Disease Control and Prevention, 2014].

Although symptom severity differs among individuals, the common characteristics of ASD include: (1) repetitive body movements or behaviors, (2) difficulty in using or understanding language, and (3) trouble in making friends and interacting with people [Gillberg, 1993; Wing, 1997]. Unfortunately, there is no single medical test that can accurately diagnose ASD. The diagnosis is mostly confirmed by the specially trained physicians based on an individual's abnormal behaviors compared with other children of the same age [Lord and Jones, 2012]. In most cases, ASD is a life-long disorder. Early diagnosis and treatment are therefore tremendously important for improving the life quality of patients and their families, as well as for reducing the financial burden borne by the society. Although evidence suggests that a number of risk behavioral markers can be observed at as early as 12 months after birth [Landa et al., 2007; Ozonoff et al., 2010; Zwaigenbaum et al., 2005], little is known prior to this, especially in the first 6 months of life.

ASD has long been thought of as a disorder with disrupted brain connectivity patterns [Shi et al., 2013], first disclosed by functional magnetic resonance imaging (fMRI) [Belmonte et al., 2004; Di Martino et al., 2011]. In the last decades, diffusion-weighted MRI (DWI) [Basser et al., 1994] has emerged as a powerful noninvasive neuroimaging technique for providing biologically and clinically relevant information on white matter (WM) integrity and connectivity that is not available from other imaging modalities. It has been increasingly used to study connectivity of WM pathways in a variety of psychiatric or neurological disorders such as Alzheimer's disease [Jin et al., 2015], bipolar disorder [Torgerson et al., 2013], schizophrenia [Lee et al., 2013], traumatic brain injury [Dennis et al., 2015a,c], or even genetics [Jin et al., 2012, 2013, 2014]. For ASD, various studies have shown widespread abnormalities in many WM tracts [Alexander et al., 2007; Barnea-Goraly et al., 2010; Cheon et al., 2011]. Specifically, Wolf et al. [2012] suggested that the aberrant development of WM tracts was ahead of the manifestation of ASD symptoms in the first year of life. For example, the failure of flexibly orienting gaze and visual attention to the most salient or biologically relevant information in the environment implied abnormal functional specialization of posterior cortical circuits in 7-month-old ASD infants [Elison et al., 2013].

The term connectome, that is, a comprehensive map of neural connections in the brain, was first introduced in 2005 [Hagmann, 2005; Sporns et al., 2005]. DWI can be used to map structural connectome of the human brain and has been used to study normal subjects [Klein et al., 2007; Tomassini et al., 2007], neurological diseases such as Alzheimer's disease [Daianu et al., 2013, 2015; Li et al., 2013; Zhu et al., 2013, 2014], and neuropsychiatric diseases such as schizophrenia [Shi et al., 2012a; van den Heuvel et al., 2010]. Using the DWI data of ASD infants acquired at the 24th month, Lewis et al. [2014] reported both local and global network deficits in the regions of temporal, parietal, occipital, as well as frontal lobes. Those results are in line with the findings in a number of studies involving infants and older children affected by ASD [Courchesne et al., 2005; Elsabbagh et al., 2013]. Since abnormalities are spread over different brain regions, a network-based approach for early identification of ASD infants seems reasonable.

Machine learning and pattern recognition technologies have been widely used to aid the diagnosis of neurodisorders, such as Alzheimer's disease [Davatzikos et al., 2008; Guo et al., 2014; Zhang et al., 2011; Zhu et al., 2014], schizophrenia [Ardekani et al., 2011; Calhoun et al., 2008], and Parkinson's disease [Haller et al., 2012; Pan et al., 2012]. In particular, different imaging modalities have been used for ASD diagnosis, including anatomical MRI [Wee et al., 2014], resting-state fMRI [Price et al., 2014], electroencephalogram and magnetoencephalography [Duffy and Als, 2012; Tsiaras et al., 2011], and DWI [Ingalhalikar et al., 2012]. However, the subjects involved in these studies were at least 2 years old with a typical age range of 7–10 years, during which ASD has progressed considerably. It is desirable to identify infants with ASD at a much earlier stage, preferably even before the first trace of symptomatic behaviors. However, identifying ASD infants is challenging and not well studied, because of the difficulty in infant image acquisition and the lack of obvious symptoms at this stage.

In this study, we propose a novel multichannel machine learning based classification framework to identify infants at high-risk for ASD when they are as young as 6 months old. Infants with autistic siblings are considered as high-risk. High-risk infants were reported up to 20 times higher probability to develop ASD than low-risk infants [Constantino et al., 2010; Lauritsen et al., 2005; Ritvo et al., 1989; Spencer et al., 2011] since they presumably carry some of the genes associated with ASD [Abrahams and Geschwind, 2008; Bailey et al., 1995]. Investigating them could help understand the mechanism of the ASD and prevent or mitigate the full onset of ASD through effective treatment if necessary. The major contribution of this work is as follows: First, we develop a new brain parcellation strategy to partition a publicly available atlas "infant AAL" (Automatic Anatomical Labeling) [Shi et al., 2011] into anatomical meaningful regions of interest (ROIs) of various sizes; Second, unlike Ingalhalikar et al. [2012], we propose to use the features from multiscale whole-brain WM connectivity networks, instead of the conventional region-based

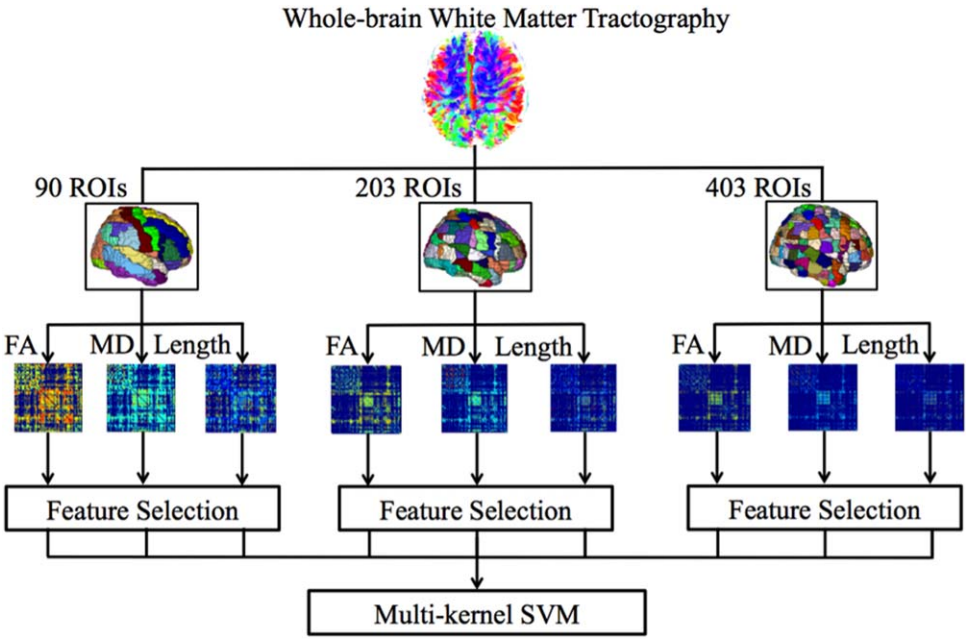

Figure 1.
The proposed classification framework based on multiparameter multiscale WM connectivity networks. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

features, to identify high-risk ASD infants; Finally, we use an effective two-stage feature selection scheme and a multikernel support vector machine (SVM) classifier that can incorporate the complementary information from multiple sources to improve the classification accuracy.

METHODS

Overview

This study involved infants who were at genetic risk for ASD (since their older siblings have ASD). For comparison, age-matched low-risk ASD infants (no first-degree relatives with ASD) were also included in this study. The study is aimed to distinguish high-risk from low-risk 6-month-old infants. Since WM abnormalities seem to be one of the first signs that appears in ASD infants, we employ features derived from the whole-brain connectivity networks that are constructed from pairwise connections between ROIs. More specifically, we start with a publicly available infant AAL atlas and parcellate its 90 cerebral ROIs into 203 and 403 sub-ROIs with our proposed parcellation algorithm, respectively. By using these ROIs, the connectivity networks of each subject are constructed using whole-brain tractography. Specifically, the WM networks of three different scales are constructed based on the average values of the DWI-derived parameters such as fractional anisotropy (FA), mean diffusivity (MD), and fiber length, all computed over the WM tracts connecting a pair of ROIs. Those values are considered features. Due to the large number of ROIs, the number of features is also large. For example, for the connectivity network at the 403-ROI scale, the maximum number of the possible features can be 81,003. Hence, a feature selection step, combining *t*-test and the Least Absolute Shrinkage and Selection Operator (LASSO) logistic regression [Tibshirani, 1996], is necessary to remove irrelevant or redundant features. Finally, the selected relevant features are fed into a multikernel SVM classifier for classification. The diagram of the complete workflow for our proposed method is shown in Figure 1.

Subjects and Image Acquisition

The participants in this study were chosen from the Lab Study 19 of National Database for Autism Research (NDAR), an on-going quadro-site study of brain and behavioral development in infants funded by the National Institute of Health (NIH). The study sites included Children's Hospital of Philadelphia, University of North Carolina at Chapel Hill (UNC-CH), University of Washington, and Washington University in St. Louis. The infants of interest were considered carrying high familial risk for ASD if at least one of their older siblings had ASD. The neuroimaging data were first collected at the 6-month visit stage. Parents or legal guardians signed on the informed consent forms. In the meantime, the infants were considered to have low familial risk if they had no first-degree relatives with ASD. They were also enrolled for comparison. All the subjects were screened to exclude other medical conditions that could affect the study. Institutional review boards at all sites approved the study. 40 high-risk infants (29 males

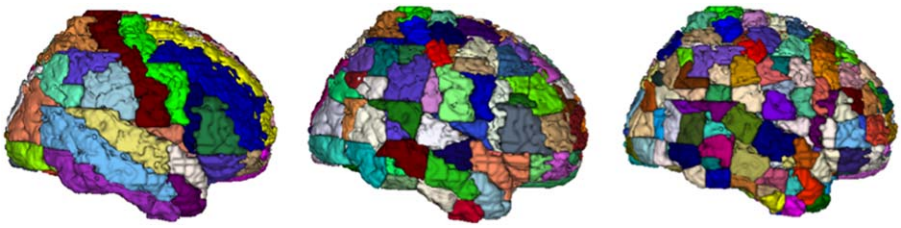

Figure 2.
The 3D views of ROIs of three different scales - 90 ROIs, 203 ROIs, and 403 ROIs, respectively (from left to right). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

and 11 females) and 40 low-risk infants (27 males and 13 females) were used in this study.

All the MRI images were scanned using a 3T Siemens Tim Trio scanner at each site. The scans were acquired during the natural sleeping of infants. Three types of MRI images were acquired: (a) T1-weighted (T1w) images with a 3D MPRAGE sequence (voxel dimensions: $1.0 \times 1.0 \times 1.0$ mm^3; image dimensions: $160 \times 224 \times 256$, TE = 3.16 ms, TR = 2400 ms); (b) T2-weighted (T2w) images with a 3D FSE sequence (voxel dimensions: $1.0 \times 1.0 \times 1.0$ mm^3; image dimensions: $160 \times 256 \times 256$, TE = 499 ms, TR = 3200 ms); (c) Diffusion-weighted images with 3D EPI sequence (voxel dimensions: $2.0 \times 2.0 \times 2.0$ mm^3; image dimensions: $96 \times 96 \times 75$, TE = 102 ms, TR = 12,800 ms). The diffusion-weighted images consisted of one b_0 volume with no diffusion sensitization, and 25 volumes with a dynamic range of diffusion gradient b from 0 to 1000 s/mm^2.

Image Preprocessing

T1w and T2w images were first skull-stripped using an open source Skull Stripping Toolkit [Shi et al., 2012b]. DWI images were skull-stripped using Brain Extraction Tool (BET) from FSL (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/) [Smith, 2002]. The skull-stripped T1w and T2w images then underwent field bias correction using nonuniform intensity normalization (N3) algorithm [Sled et al., 1998]. The DWI images were corrected for eddy-current induced distortion using FSL. The diffusion gradient tables were adjusted accordingly. FA and MD images were then extracted from the DWI data after diffusion tensor fitting.

Multiscale ROI Parcellation

The AAL atlas [Tzourio-Mazoyer et al., 2002] is a widely used high-resolution T1w brain parcellation based on a single adult subject, where the main sulci were used as landmarks to parcellate the entire brain into 90 ROIs (cerebellum excluded) based on anatomical definitions. However, directly applying adult atlas into infant brains may compromise its performance due to their large anatomical variations. Instead, we used a publicly available 1-year-old infant brain atlas [Shi et al., 2011], which adapted the original AAL atlas to infant space and has been proven effective in infant normalization and parcellation tasks. To promote the characterization of brain connectivity at different scales, we propose a novel strategy to further parcellate the infant AAL atlas into smaller sub-ROIs.

Similar to [Hagmann et al., 2008], we created normalized sub-ROIs by controlling them to have a similar size. Given a user-specified size a, we first initialized a group of sub-ROIs by dividing the entire volume of the infant AAL template into cubes of size a^3. Then, for each sub-ROI, its center was calculated by taking the mean coordinates of all the voxels involved. If the resulting sub-ROI was too small, for example, if it was near the boundary of the original ROI and had less than 30% of the standard size, we would attach it to a neighboring bigger divided region located in the same ROI whose center was the closest. The new center of the combined region was accordingly recalculated. After doing this for each of the 90 original ROIs, we had a series of center points inside each original ROI. Finally, we performed Voronoi partition [Aurenhammer, 1991], that is, for each original ROI, we assigned each voxel inside it a sub-ROI index based on its shortest distance to one of those sub-ROI centers. The steps described above are summarized in Algorithm 1. In our case, 203 and 403 ROIs were obtained when the target ROI size was chosen

Algorithm 1 Multiscale ROI Parcellation

Set cube size a and load the original infant AAL ROIs.
Divide the entire volume (including the background) evenly into cubes of size a^3.
for ROI $i = 1{:}90$ //compute the center of each new sub-ROI
  if any interior volume is at least $0.3{*}a^3$, e.g., $V_{cube} > 0.3{*}a^3$,
    calculate the center of the cube.
  else Find a neighboring cube whose center is the closest to its center and combine the two regions.
    Recompute the new center of the combined region.
  end if
end for
for ROI $i = 1{:}90$ //assign new sub-ROIs
  Perform Voronoi partition based on the new centers inside each ROI i.
end for

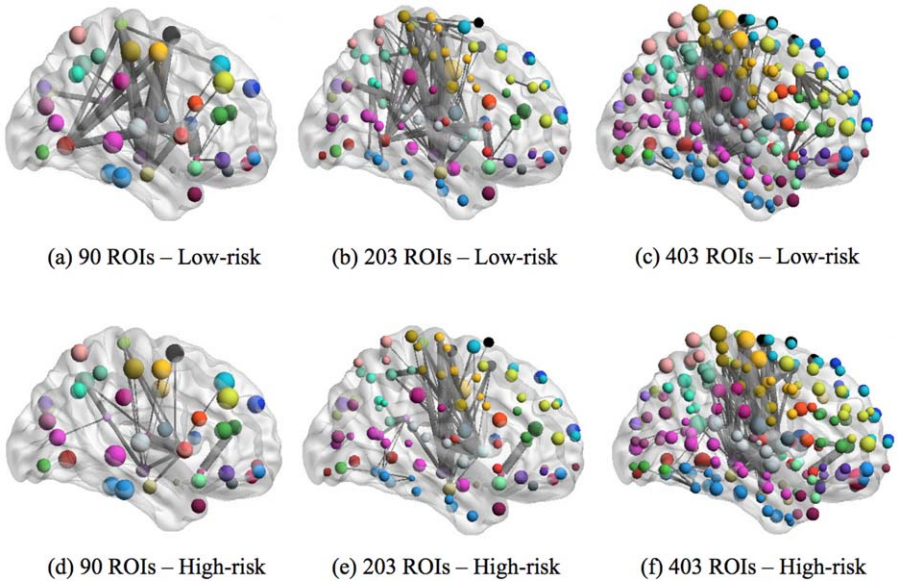

Figure 3.

The 3D views of connectivity networks of three different scales, thresholded by mean FA = 0.4 (considered as strong connection), from a randomly selected pair of low-risk and high-risk ASD subjects. At each scale, each sphere represents the center of an ROI and its size indicates the normalized volume of the ROI. The colors of the sub-ROIs in (**b–c**) and (**e–f**) correspond to those of the original ROIs in (**a**) and (**d**), respectively. The thickness of edges represents the mean value of FA averaged over the fibers connecting two respective ROIs. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

at around 20 and 16 mm, respectively. Figure 2 provides the 3D views of the ROIs in three different scales.

Multiscale Connectivity Networks

Features derived from structural connectivity networks provide rich information for identifying ASD subjects due to its comprehensive characterization of connections between brain regions. Furthermore, a multiscale set of connectivity networks may provide a range of complementary coarse-to-fine information for the multilevel analysis of brain connections. A connectivity network can be described mathematically as a graph consisting of (1) a collection of nodes, representing the ROIs and (2) a set of edges between nodes, describing the connections between ROI pairs.

To transfer the ROIs to the DWI space of each subject for construction of the connectivity networks, we performed the following steps. First, the infant AAL template was nonlinearly registered to the T1w image of the subject. Then, to further align the infant AAL template to the space of the DWI data, the T1w image was linearly aligned to the T2w image and the T2w image was nonlinearly registered (considering EPI distortion) to the MD image, since a T2w image had the similar contrast as an MD image. Here, we used the T1w and T2w images as the intermediate targets to transfer the AAL ROIs to the DWI space of the subject smoothly. It actually achieved the better accuracy than directly registering the T1w AAL template to the FA image of the subject, since both inter-modality and inter-subject made the latter process much more difficult. The deformation fields estimated in these series of registration were concatenated and then used to warp the atlas ROIs to the individual DWI space. All the registration steps were performed using ANTs [Avants et al., 2008].

Whole-brain tractography was performed using an image resolution enhanced deterministic streamline tractography with peaks detected from the WM orientation distribution functions [Yap and Shen, 2012]. This tractography algorithm allowed for a closer investigation of fiber structures and more accurate assessment of brain connectivity [Yap et al., 2014]. Seed points were chosen as voxels with $FA > 0.3$. To eliminate excess false positives, the maximum turning angle was set to 45°, and tracking was stopped when $FA < 0.15$. The final tractography of all subjects was checked manually to ensure the anatomical correctness.

Two ROIs were considered anatomically connected if there were common fibers traversing them. Three different metrics, that is, FA, MD, and fiber length, were used to indicate the connection strength. For each pair of ROIs, connection strengths were computed based on the values of these metrics spatially averaged over the fibers connecting these ROIs. To exclude noisy connections, any node pairs that had <4 average fiber connections across the entire population of subjects were considered unconnected [Zalesky et al., 2011]. By using ROIs of three different scales, we constructed three

sets of multiscale connectivity networks (nine in total) for each subject. Figure 3 shows the FA connectivity networks of three different scales for a pair of low-risk and high-risk ASD subjects from our dataset. We notice that when the scale becomes finer, the connections between the coarse ROIs in Figure 3a,d are spread into more detailed connections between the embedded sub-ROIs in Figure 3b,c,e,f. Notice that the multiscale networks of a low-risk subject (Figure 3a–c) contain stronger connections (i.e., more and thicker edges) than those of a high-risk subject (Figure 3d–f).

Feature Selection

We considered each element of a connectivity network as a feature. Since the connectivity network was symmetric, we only took the elements from its lower triangular part (since the diagonal values were always set zero). There are several reasons why feature selection is necessary: First, the number of features increases with the square of the number of ROIs, from thousands in a 90-ROI network to tens of thousands in a 403-ROI network, and many of these features may be irrelevant to the classification task. Second, feature selection is needed in order to improve generalization performance. Finally, reducing the number of features will reduce the computation time significantly. Here, we first normalized the features by their z-scores, and then performed a two-stage feature selection procedure to select the optimal subset of features that would both improve the accuracy and ease the computation burden.

Specifically, we first performed a t-test and then followed by the LASSO logistic regression [Tibshirani, 1996]. A t-test was conducted between the high-risk and the low-risk groups for each feature. Only the features whose P-values were below an empirical threshold were selected for LASSO logistic regression. Here, the t-tests acted as a feature ranking technique to select an initial set of most relevant features based on their individual discriminative powers [Fan et al., 2007], which also greatly eased the computation cost for the following steps. By considering the correlation among features, the LASSO logistic regression model further optimized the selection. The LASSO logistic regression is a commonly used feature selection algorithm that can select a parsimonious set of features from a large collection of possible candidates to improve the classification accuracy. It only keeps the most discriminative features while discarding the redundant ones. Mathematically, the LASSO logistic regression model can be described as follows:

$$\text{Let } \boldsymbol{X}=\begin{bmatrix}\boldsymbol{x}_1^T\\ \vdots\\ \boldsymbol{x}_M^T\end{bmatrix}\in\mathbb{R}^{M\times N},\quad \boldsymbol{A}=\begin{bmatrix}a_1\\ \vdots\\ a_N\end{bmatrix}\in\mathbb{R}^{N\times 1},\quad \text{and}\quad \boldsymbol{Y}=\begin{bmatrix}y_1\\ \vdots\\ y_M\end{bmatrix}\in\mathbb{R}^{M\times 1}.$$

Here $\boldsymbol{x}_i\in\mathbb{R}^{N\times 1}$ is the feature vector of the i-th subject. a_i is the feature regression coefficient for the i-th feature. y_i is the label for the i-th subject (i.e., $+1$ for high-risk and -1 for low-risk). M is the total number of subjects and N is the total number of features. The LASSO logistic regression aims to solve $\boldsymbol{A}$ by minimizing the following cost function:

$$\begin{gathered}\min_{\boldsymbol{A}}\sum_{i=1}^{M} w_i\log\left(1+\exp\left(-y_i\left(\mathbf{A}^T\boldsymbol{x}_i+c\right)\right)\right)+\lambda||\boldsymbol{A}||_1+\frac{\rho}{2}||\boldsymbol{A}||_2^2\\ \text{s.t. }\sum_{i=1}^{M} w_i=1 \text{ and } w_i=w_j \text{ if } y_i=y_j\end{gathered}\tag{1}$$

where w_i is the weight for the i-th subject, $||\boldsymbol{A}||_1=\sum_{i=1}^{N}|a_i|$ is the ℓ_1 norm of $\boldsymbol{A}$, and $||\boldsymbol{A}||_2^2=\sum_{i=1}^{N}a_i^2$ is the squared ℓ_2 norm. Here, λ is the L_1 norm regularization parameter that controls the model sparsity. The larger the value of λ, the sparser the model. Meanwhile, ρ is the regularization parameter for the squared ℓ_2 norm that helps to achieve numerical stability and increase predictive performance. We performed the LASSO logistic regression on each individual network separately and retain features with nonzero coefficients in $\boldsymbol{X}$.

Multikernel Classification

Kernel methods such as SVM have been proven to be efficient for classification problems [Scholkopf and Smola, 2001]. It has been demonstrated that multikernel SVM is able to utilize the complementary information from multiple representations for achieving higher accuracy than single-kernel SVM [Liu et al., 2014; Liu et al., 2015; Wee et al., 2014; Zhang et al., 2011; Zhu et al., 2015a,b]. Here, we used it to incorporate the selected features from all nine networks for classification.

A multikernel SVM classification framework is described as below. Let $\left\{\left(\boldsymbol{x}_i^j,\ y_i\right),\ i=1,\ldots,M,\ j=1,\ldots,K\right\}$ be the set of training data, where $\boldsymbol{x}_i^j\in\mathbb{R}^{n_j\times 1}$ represents the feature vector of the i-th subject for network j and $y_i\in 1$ (high-risk) or -1 (low-risk) is the label. M is the number of training subjects, and K is the number of networks (i.e., $K=9$ in our study). The primal formulation of multikernel SVM is given as:

$$\begin{gathered}\min_{\boldsymbol{P}_j,b,\xi_i}\frac{1}{2}\sum_{j=1}^{K}\beta_j||\boldsymbol{P}_j||^2+C\sum_{i=1}^{M}\xi_i\\ \text{s.t. } y_i\left(\sum_{j=1}^{K}\beta_j\left(\boldsymbol{P}_j^T\boldsymbol{\phi}_j\left(\boldsymbol{x}_i^j\right)+b\right)\right)\geq 1-\xi_i\\ \text{and } \xi_i>0,\ \sum_{j=1}^{K}\beta_j=1,\ \beta_j>0\end{gathered}\tag{2}$$

where β_j, $\boldsymbol{P}_j$, and $\boldsymbol{\phi}_j$ denote the weight, the normal vector of the classification hyperplane, and the kernel-induced mapping function for the j-th network, respectively. b denotes the bias term, ξ_i denotes the slack variable (for misclassification), and C is a parameter that controls the degree of misclassification. Given a test feature vector $=\left\{\boldsymbol{x}^j,\ j=1,\ldots,K\right\}$, the predicted label $\hat{y}$ is

$$\hat{y}=\text{sign}\left(\sum_{j=1}^{K}\beta_j\left(\boldsymbol{P}_j^T\boldsymbol{\phi}_j\left(\boldsymbol{x}^j\right)+b\right)\right)\tag{3}$$

Here, we applied an open-source software package SimpleMKL [Rakotomamonjy et al., 2008] that can decide

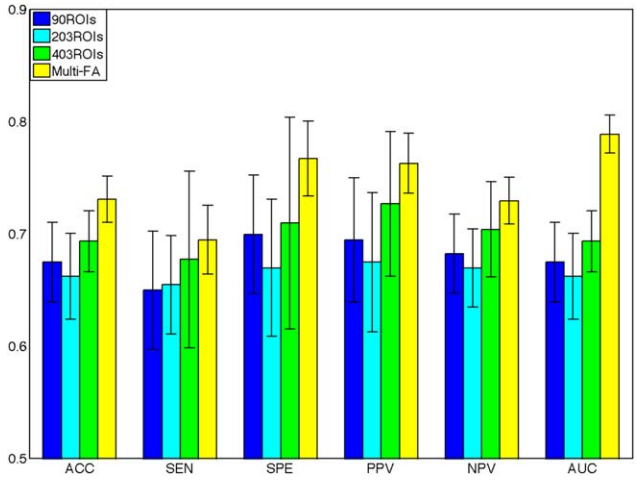

Figure 4.

The mean statistics of the 6-month-old infant ASD classification (high-risk vs. low-risk) with the 90-ROI (blue), 203-ROI (cyan), 403-ROI (green) network, respectively, and the integration (yellow) of all the three networks using mean **FA**. The error bars stand for the standard deviations of the measures after 10 repetitions. The multi-scale networks outperform any single-scale network in all statistics. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

the weights simultaneously during the optimization of the cost function.

Implementation and Evaluation

The classification performance was evaluated on a nested 5-fold cross validation. The 80 subjects were randomly divided into 5 folds. Each fold was comprised of 8 high-risk infants and 8 low-risk infants, respectively. Each fold was alternatively used as a testing fold, and the rest of the four folds were used as the training set.

For each network, during each of the five training sessions, a t-test was to select the features whose p-values < 0.001. Those features were fed into the LASSO logistic regression. In machine learning, feature ranking and selection can use a performance evaluation metric computed directly from the data and not based on the direct feedback from the classifier [Guyon et al., 2006]. In other words, we used the predicted results directly from the logistic regression to make feature selection without the classifier involved. The 64 subjects in the training set were divided into 5 folds again and an inner loop of 5-fold cross validation was performed. The binarized results of the regression (the result = 1 if positive and −1 otherwise) were compared to the ground truth. The set of parameters λ and ρ in Eq. (1) with the best average accuracy over the inner 5-fold cross validation were used to select the features for the corresponding testing fold.

The selected features were then fed into the SimpleMKL package with the Gaussian radial basic function (RBF) as the kernel. Note that another inner 5-fold cross validation was performed on the training set to find the optimal parameter C in Eq. (2) and the test fold was classified with that determined C.

The statistics we used to evaluate our classification algorithm performance are accuracy (ACC), sensitivity (SEN), specificity (SPE), positive predictive value (PPV), negative predictive value (NPV), and the area under the receiver operating characteristic curve (AUC). Letting TP, TN, FP, and FN denote, respectively, true positive, true negative, false positive, and false negative, the ACC, SEN, SPE, PPV, and NPV are defined as

$$\mathrm{ACC}=\frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}}$$

$$\mathrm{SEN}=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}$$

$$\mathrm{SPE}=\frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}}$$

$$\mathrm{PPV}=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}$$

$$\mathrm{NPV}=\frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FN}}$$

To avoid the biased result due to the fold selection, the entire process was repeated 10 times, each with different partitions of subjects. The average statistics of the 10 repetitions were finally reported.

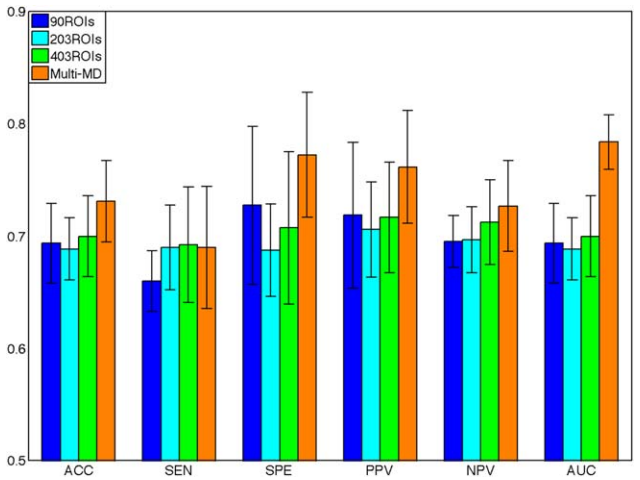

Figure 5.

The mean statistics of the 6-month-old infant ASD classification (high-risk vs. low-risk) with the 90-ROI (blue), 203-ROI (cyan), 403-ROI (green) network, respectively, and the integration (orange) of all the three networks using mean **MD**. The error bars stand for the standard deviations of the measures after 10 repetitions. The multi-scale networks outperform any single-scale network in all statistics. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

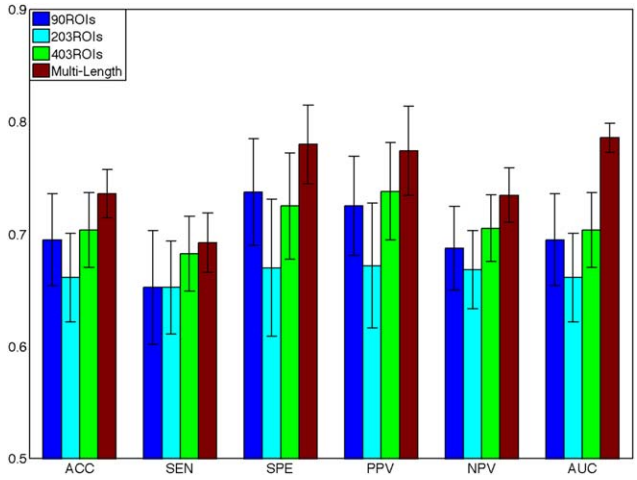

Figure 6.
The mean statistics of the 6-month-old infant ASD classification (high-risk vs. low-risk) with the 90-ROI (blue), 203-ROI (cyan), 403-ROI (green) network, respectively, and the integration (dark red) of all the three networks using mean **fiber length**. The error bars stand for the standard deviations of the measures after 10 repetitions. The multiscale networks outperform any single-scale network in all statistics. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

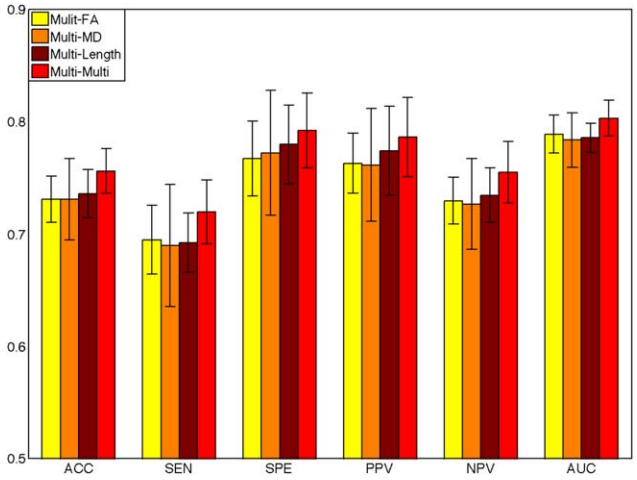

Figure 7.
The mean statistics of the 6-month-old infant ASD classification (high-risk vs. low-risk) with the single-parameter multiscale network framework using FA (yellow), MD (orange), and fiber length (dark red), vs. our method with multiparameter multiscale networks (red). The error bars stand for the standard deviations of the measures after 10 repetitions. It can be seen that our multiparameter multiscale networks outperform each of the single-parameter networks in all statistics. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

RESULTS

Multiscale Connectivity Networks

To evaluate the benefits of using multiscale connectivity networks on the ASD classification, we compared its performance with single-scale networks. Figures 4–6 compare the classification performance of the single-scale networks with that of the multiscale networks (using all three scales) for FA (Multi-FA), MD (Multi-MD), and fiber length (Multilength). Note that the results of the single-scale network were obtained by applying the single RBF kernel SVM on the selected features, while the results of the multiscale networks were obtained using multikernel SVM. The error bars denote the standard deviations obtained from 10 repetitions. Not surprisingly, the multiscale networks outperformed any single-scale network for each imaging parameter. In particular, the accuracies of Multi-FA, Multi-MD, and Multilength were 73.1, 73.1, and 73.6%, respectively. They achieved relative gains of 5.4, 4.5, and 4.6%, respectively, compared to the best single-scale networks (the 403-ROI networks). The AUCs were 0.789, 0.784, and 0.786 for Multi-FA, Multi-MD, and Multilength, respectively. They had the most significant improvement among all the statistics, by relative gains of 14, 12, and 12% against the best single networks (the 403-ROI networks) for FA, MD, and fiber length, respectively. We also noticed that the standard deviations of all statistics were smaller in the multiscale networks, especially for AUC, suggesting a more stable performance. To test whether the improvement was statistically significant, we also conducted pair-wise *t*-tests between the accuracy results of both the single-scale and the multiscale networks over the 50-fold tests in our experiment, with the results shown in Table I. The fact that all p-values < 0.05 implied that the improvements using the multiscale network approach were indeed statistically significant.

TABLE I. The *p*-values of the *t*-tests on the mean accuracies with the single-scale networks versus the multiscale networks using FA, MD, and fiber length

Method	403 ROIs	203 ROIs	90 ROIs
Multi-FA	0.016	<0.001	<0.001
Multi-MD	0.003	0.002	0.019
Multilength	0.040	<0.001	0.005

The t-tests were the pair-wise t-tests on the classification results from the total 50 testing folds, that is, over 10 iterations of 5-fold cross validation.

Multiparameter Multiscale Connectivity Networks

In this section, we demonstrate that our proposed method using all the 9 networks (multiscale networks with FA, MD, and fiber length) achieved the best result for our classification

TABLE II. The *p*-values of the *t*-tests on the mean accuracies between the multiparameter multiscale networks (multi-multi) and the single-parameter multiscale networks (multi-FA, multi-MD, and multilength)

Method	Multi-FA	Multi-MD	Multilength
Multi-multi	0.004	0.012	0.006

The t-tests were the pair-wise t-tests on the classification results from the total 50 testing folds, that is, over 10 iterations of 5-fold cross validation.

task. Figure 7 shows the comparison of the classification performance of the multiparameter multiscale network framework (Multi-Multi) and the results of single-parameter multiscale networks (i.e., Multi-FA in Figure 4, Multi-MD in Figure 5, and Multilength in Figure 6). The accuracy of Multi-Multi (75.6%) had an improvement of 2.5, 2.5, and 2% against that of Multi-FA, Multi-MD, and Multilength, respectively. Other statistical measures, sensitivity (72%), specificity (79.3%), PPV (78.7%), NPV (75.5%), and AUC (80.3%), were also the best. We further conducted pair-wise *t*-tests on the accuracy results of Multi-Multi against Multi-FA, Multi-MD, and Multilength over the 50 testing folds conducted in our classification protocol. The results are shown in Table II. The fact that all the p-values < 0.05 further indicated the statistically significant improvements using the proposed multiparameter multiscale network framework. The processing time of the classification (5-fold cross validations with 10 iterations) with our proposed method was 7 min in the Matlab environment on an iMac equipped with a 3.4 GHz Intel Core i7 CPU and 16 GB RAM. The same task with a single-scale network took 2 min. The computing time was tripled, but we think that it was still manageable and worthy spending, considering the much improved performance by our method.

Feature Selection

To evaluate the effect of feature selection on the final classification results, we further conducted the classifications with and without feature selection. Table III proves that the feature selection greatly improved the performance of classification. With feature selection, the accuracy was improved by 16% (i.e., a relative 27% increase), the AUC by 21% (i.e., a relative 33% increase), and the sensitivity by 26% (i.e., a relative 57% increase). The standard deviations of all the measures were also smaller, which indicated a more stable performance in each fold with feature selection.

TABLE III. Comparison of the classification performance with and without feature selection

Mean (standard deviation)	No feature selection	Feature selection
ACC	0.591 (0.036)	0.756 (0.020)
SEN	0.460 (0.064)	0.720 (0.028)
SPE	0.723 (0.065)	0.793 (0.033)
PPV	0.656 (0.076)	0.787 (0.036)
NPV	0.577 (0.057)	0.755 (0.028)
AUC	0.591 (0.036)	0.803 (0.016)

TABLE IV. Comparison of the classification performance between linear and logistic regression

Mean (standard deviation)	Linear regression	Logistic regression
ACC	0.711 (0.027)	0.756 (0.020)
SEN	0.680 (0.045)	0.720 (0.028)
SPE	0.743 (0.033)	0.793 (0.033)
PPV	0.734 (0.038)	0.787 (0.036)
NPV	0.715 (0.033)	0.755 (0.028)
AUC	0.783 (0.026)	0.803 (0.016)

We also investigated how different regression models could affect the classification results. Table IV compares the performance of logistic regression with the commonly used linear regression. Logistic LASSO regression yielded better classification with accuracy, sensitivity, and specificity gains of 4.5, 4, and 5%, respectively.

Choice of Classifier

We chose the RBF kernel for SVM since it usually performs better than a linear kernel. Table V compares the performance between these two kernels with our dataset. Better performance was seen in all the measures when the RBF kernel was used. In particular, the accuracy with RBF kernel was 0.76, while the accuracy with linear kernel was only 0.70. The relative improvement was 8%. Specificity and PPV had the most relative increase, that is, by 14 and 11%, respectively.

Table VI compares the performance of the single kernel and the multikernel SVM. The single kernel SVM classification was implemented by concatenating all the features from the nine networks into a long feature vector and feeding it into the RBF-kernel SVM classifier. Again, all statistics were improved by using a multikernel SVM (SimpleMKL in our case). The accuracy jumped from 0.72 to 0.76, a relative 6% increment. The increments of AUC, specificity, and PPV were among the largest, which were 12, 11, and 10%, respectively.

TABLE V. Comparison of the classification performance between linear and RBF kernels

Mean (standard deviation)	Linear	RBF
ACC	0.703 (0.032)	0.756 (0.020)
SEN	0.708 (0.051)	0.720 (0.028)
SPE	0.698 (0.045)	0.793 (0.033)
PPV	0.709 (0.029)	0.787 (0.036)
NPV	0.723 (0.051)	0.755 (0.028)
AUC	0.781 (0.031)	0.803 (0.016)

TABLE VI. Comparison of the classification performance between the single kernel SVM and the multikernel SVM

Mean (standard deviation)	Single kernel SVM	Multikernel SVM
ACC	0.715 (0.019)	0.756 (0.020)
SEN	0.715 (0.038)	0.720 (0.028)
SPE	0.715 (0.021)	0.793 (0.033)
PPV	0.715 (0.020)	0.787 (0.036)
NPV	0.736 (0.031)	0.755 (0.028)
AUC	0.715 (0.019)	0.803 (0.016)

Multiparameter Multiscale vs. Multiparameter Selected-Scale

For each parameter (FA, MD, and length), we also picked the network with the best performance (highest accuracy) in multiscale networks. The 403-ROI networks performed the best for each parameter. We combined these three 403-ROI networks to form the new multiparameter networks. They were fed into the same multikernel SVM classifier and the performance was compared with our proposed algorithm. The results are shown in Table VII. Our proposed method achieved a better accuracy by almost 4%.

Number of Networks

We have shown that using multiscale networks constructed from different scales of the infant AAL ROI improved the accuracy of classification. The 403-ROI networks were better networks than both the 203-ROI networks and the original 90-ROI networks, in terms of classification performance. Here, we further partitioned the infant AAL atlas into 604 ROIs in the same way as described in the "multiscale ROI Parcellation" section to examine whether adding the networks constructed from those ROIs would further improve the classification result. Table VIII shows the classification results using the single-parameter (FA, MD, and length) networks constructed from these 604 ROIs. Despite the finer ROIs, the performance actually deteriorated. For instance, the accuracies of the 604-ROI networks were 67.6, 65.6, and 66.4%, for FA, MD, and length, respectively, while those of the 403-ROI networks were 69.3, 70.0, and 70.3%, respectively. We further integrated the three 604-ROI networks into our multiparameter multiscale network framework. With a total of 12 networks, the mean accuracy went down from 75.6 to 70%.

TABLE VII. Comparison of the classification performance between the multiparameter selected-scale networks (i.e., the three best single networks using FA, MD, and length) and our proposed multiparameter multiscale networks

Mean (standard deviation)	Multiparameter selected-scale	Multiparameter multiscale
ACC	0.719 (0.033)	0.756 (0.020)
SEN	0.700 (0.048)	0.720 (0.028)
SPE	0.740 (0.066)	0.793 (0.033)
PPV	0.745 (0.041)	0.787 (0.036)
NPV	0.722 (0.030)	0.755 (0.028)
AUC	0.792 (0.027)	0.803 (0.016)

TABLE VIII. Classification performance of the 604-ROI network for each parameter (FA, MD, and length)

Mean (standard deviation)	604 ROIs - FA	604 ROIs - MD	604 ROIs - Length
ACC	0.676 (0.039)	0.656 (0.030)	0.664 (0.038)
SEN	0.638 (0.041)	0.610 (0.070)	0.655 (0.048)
SPE	0.715 (0.064)	0.703 (0.076)	0.673 (0.077)
PPV	0.704 (0.063)	0.684 (0.055)	0.680 (0.052)
NPV	0.676 (0.036)	0.656 (0.036)	0.676 (0.040)
AUC	0.676 (0.039)	0.656 (0.030)	0.664 (0.038)

Choice of Parameters

Fiber count is another popular measure in WM connectivity analysis [Jahanshad et al., 2011; Zhan et al., 2013a,b,c]. We also investigated how good fiber count acted as a feature in our infant ASD classification. The fiber count connectivity network refers to the matrix consisting of the relative fiber count connecting pairwise ROIs (normalized by the total number of fibers in the whole-brain tractography). We performed classification with the 90-ROI, 203-ROI, and 403-ROI fiber count networks, respectively. The results are shown in Table XI. The mean accuracies from those fiber count networks were much lower than those built on other parameters we used in the proposed algorithm (FA, MD, and length). Therefore, we did not include fiber count in our classification model.

Most Discriminative Connections

We summed the counts of each connection selected by our proposed method for each network scale (90, 203, and 403 ROIs) over the 50 folds. The counts for FA, MD, and fiber length were combined at each level. We ranked the

TABLE XI. Classification performance of the normalized fiber count networks in three different scales

Mean (standard deviation)	403 ROIs	203 ROIs	90 ROIs
ACC	0.586 (0.027)	0.531 (0.046)	0.586 (0.055)
SEN	0.583 (0.065)	0.560 (0.062)	0.628 (0.103)
SPE	0.590 (0.068)	0.503 (0.081)	0.545 (0.078)
PPV	0.595 (0.046)	0.536 (0.056)	0.588 (0.055)
NPV	0.603 (0.028)	0.538 (0.048)	0.627 (0.081)
AUC	0.586 (0.030)	0.531 (0.046)	0.586 (0.055)

TABLE X. The ROIs used in Figure 8

Index (odd -> left, even -> right)	Region	Abbreviation	
		Left	Right
1, 2	Precentral gyrus	PreCG.L	PreCG.R
3, 4	Superior frontal gyrus (dorsal)	SFGdor.L	SFGdor.R
5, 6	Orbitofrontal cortex (superior)	ORBsupb.L	ORBsupb.R
7, 8	Middle frontal gyrus	MFG.L	MFG.R
9, 10	Orbitofrontal cortex (middle)	ORBmid.L	ORBmid.R
11, 12	Inferior frontal gyrus (opercular)	IFGoperc.L	IFGoperc.R
13, 14	Inferior frontal gyrus (triangular)	IFGtriang.L	IFGtriang.R
15, 16	Orbitofrontal cortex (inferior)	ORBinf.L	ORBinf.R
17, 18	Rolandic operculum	ROL.L	ROL.R
19, 20	Supplementary motor area	SMA.L	SMA.R
21, 22	Olfactory	OLF.L	OLF.R
23, 24	Superior frontal gyrus (media)	SFGmed.L	SFGmed.R
25, 26	Orbitofrontal cortex (medial)	ORBmed.L	ORBmed.R
27, 28	Rectus gyrus	REC.L	REC.R
29, 30	Insula	INS.L	INS.R
31, 32	Anterior cingulate gyrus	ACG.L	ACG.R
33, 34	Middle cingulate gyrus	MCG.L	MCG.R
35, 36	Posterior cingulate gyrus	PCG.L	PCG.R
37, 38	Hippocampus	HIP.L	HIP.R
39, 40	ParaHippocampal gyrus	PHG.L	PHG.R
41, 42	Amygdala	AMYG.L	AMYG.R
43, 44	Calcarine	CAL.L	CAL.R
45, 46	Cuneus	CUN.L	CUN.R
47, 48	Lingual gyrus	LING.L	LING.R
49, 50	Superior occipital gyrus	SOG.L	SOG.R
51, 52	Middle occipital gyrus	MOG.L	MOG.R
53, 54	Inferior occipital gyrus	IOG.L	IOG.R
55, 56	Fusiform gyrus	FFG.L	FFG.R
57,58	Postcentral gyrus	PoCG.L	PoCG.R
59, 60	Superior parietal gyrus	SPG.L	SPG.R
61, 62	Inferior parietal lobule	IPL.L	IPL.R
63, 64	Supramarginal gyrus	SMG.L	SMG.R
65, 66	Angular gyrus	ANG.L	ANG.R
67, 68	Precuneus	PCUN.L	PCUN.R
69, 70	Paracentral lobule	PCL.L	PCL.R
71, 72	Caudate	CAU.L	CAU.R
73, 74	Putamen	PUT.L	PUT.R
75, 76	Pallidum	PAL.L	PAL.R
77, 78	Thalamus	THA.L	THA.R
79, 80	Heshl gyrus	HES.L	HES.R
81, 82	Superior temporal gyrus	STG.L	STG.R
83, 84	Temporal pole (superior)	TPOsup.L	TPOsup.R
85, 86	Middle temporal gyrus	MTG.L	MTG.R
87, 88	Temporal pole (middle)	TPOmid.L	TPOmid.R
89, 90	Inferior temporal	ITG.L	ITG.R

connections by the counts of selection. For the purpose of visualization, Figure 8a–c illustrates the top 20 most discriminative representations of connectogram, a circular representation tool called Circos (www.cpan.org/ports) [Krzywinski et al., 2009], for the 90-ROI, 203-ROI, and 403-ROI networks, respectively. Note that for the 90-ROI networks, the connections between the original ROIs were reported. While, in the 203-ROI and 403-ROI networks, the connections in the original 90 ROIs that contained those selected connections between sub-ROIs were reported for comparison. The connections in the same hemisphere were represented in red while inter-hemisphere in black. Thickness of each line reflected its selection frequency, that is, the thicker the line, the higher the frequency. A total of 65% connections (13 pairs) overlapped between the 90-ROI scale and the 203-ROI scale, 20% (4 pairs) between the 90-ROI scale and the 403-ROI scale, and 15% (3 pairs) between the 203-ROI scale and the 403-ROI scale. 10% connections (2 pairs)

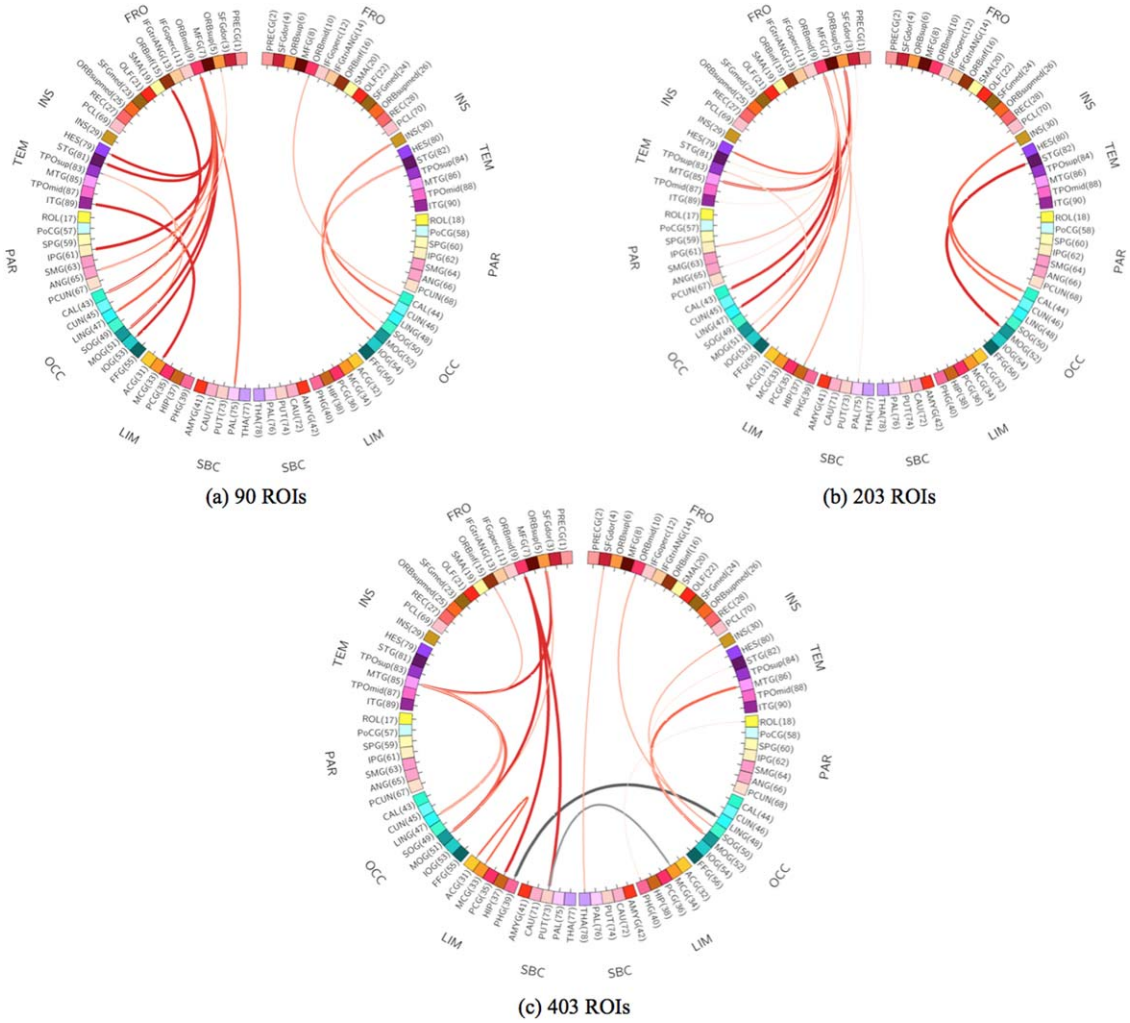

Figure 8.

Connectogram of the top 20 discriminative connections selected by our framework in the (**a**) 90-ROI, (**b**) 203-ROI, and (**c**) 403-ROI networks, respectively. The intrahemisphere and interhemisphere connections are shown in red and black colors, respectively. For the abbreviations of the ROIs, please refer to Table X. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

appeared in all the three scales. They were the connections between the left middle occipital gyrus and the left middle orbitofrontal cortex, as well as between the right superior temporal gyrus and the right middle occipital gyrus. The ROIs used in Figure 8 are listed in Table X.

DISCUSSION

The population size of our study (40 high-risk vs. 40 low-risk) has been commonly seen in previous classification studies. For example, for schizophrenia classification, Ardekani et al. [2011] used 50 healthy controls and 50 patients, and Calhoun et al. [2008] used 26 healthy controls, 21 schizophrenia patients, and 14 bipolar patients. For ASD, Bosl et al. [2011] used 33 normal controls and 46 ASD patients, and Ingalhalikar et al. [2012] used 30 normal controls and 45 ASD patients.

The way to divide sub-ROIs is a critical step for building the multiscale connectivity networks. If we use a fixed number to divide every ROI, the sub-ROIs generated from the smaller ROIs may be too little to have any fibers traversing them, thus resulting in the heavy weights on the connections between the sub-ROIs generated from those larger ROIs. However, those larger regions may not be discriminative in classification. It is more reasonable to have each sub-ROI be as the equal size as possible so that it can be weighted fairly in a connectivity network. The 203 and

403 sub-ROIs were the results of an approximately size of 20 and 16 mm for each sub-ROI, respectively. They were anatomically meaningful considering that the structure of the original AAL ROIs was retained. For example, the superior temporal gyrus was a single ROI in the original AAL atlas while we subdivided it into multiple sub-ROIs with similar sizes. In this way, we retained the anatomical information of each sub-ROI.

The original 90-ROI infant AAL cortical parcellation might not be the most suitable for identifying differences between high-risk and low-risk ASD infants. Smaller networks (with larger ROIs) might be more beneficial for group studies since more stable features within groups can be extracted, while larger networks (with smaller ROIs) provide more detailed connectivity information at the individual level and more precise distinction can be captured [Cammoun et al., 2012]. The right scale depends on the studies [Fornito et al., 2010; Zalesky et al., 2010]. In this study, we further divide the original ROIs into smaller ROIs to form atlases of three different scales, that is, 90 ROIs, 203 ROIs, and 403 ROIs. The 403-ROI networks give the highest classification accuracy for all three measures (FA, MD, and length) (see Figures 4–6). The multiscale networks improve the classification accuracy for each metric (FA, MD, and length) by 5.4, 4.5, and 4.6%, compared with the best single networks, respectively. The AUCs are improved by 13.8, 12, and 11.7%, respectively (see Figures 4–6). Integrating multiscale networks further improves classification performance. We note that, however, adding the 604-ROI networks does not further improve the classification accuracy (see the "Number of Atlases" section).

Wee et al. [2011] demonstrated that utilizing networks constructed from multiple DWI-derived parameters enhances classification performance. In this work, we confirm this and show that, using the multiparameter networks based on FA, MD, and length, we can improve the accuracy by 2.5, 2.5, and 2% compared with any single parameter network (see Figure 7).

The top 20 most discriminative connections (see Figure 8) from the networks of three different scales overlap, that is, the top selected connections in the original 90-ROI network contain many selected sub-ROI connections in the 203-ROI and the 403-ROI networks. The relevant regions include (1) multiple cortical regions, such as the orbital part of the left frontal lobe, the bilateral temporal lobes, the bilateral occipital lobes, the bilateral calcarine fissures, the bilateral lingual gyri, the right cuneus, the right insula and (2) several subcortical regions, such as the left globus pallidum, the left putamen, and the bilateral parahippocampal gyri. Those findings are consistent with the regions identified in a number of ASD infant or child studies. The temporal lobes, the occipital lobes, the calcarine fissures, the lingual gyri, and the cuneus are involved in processing auditory and visual stimuli, language and nonlinguistic social stimuli. Abnormalities exist in all these regions in ASD patients [Johnson et al., 2005; Redcay, 2008; Schultz et al., 2000; Wetherby et al., 2004]. The frontal cortical connectivity abnormalities were also well documented in ASD [Just et al., 2012]. Among the subcortical regions, insula is positioned as a hub to mediate external and internal cognitive processes. Its dysfunction plays an important role in ASD [Uddin and Menon, 2009]. The basal ganglia region that contains globus pallidum and putamen also shows shape deformation in ASD patients [Qiu et al., 2010]. Attenuated neural activities in parahippocampal regions may explain some aspects of ASD [Sumiyoshi et al., 2011].

It is also interesting to notice the asymmetry between the left hemisphere and the right hemisphere in the three connectograms in Figure 8. The majority of the top 20 discriminative connections lie in the left hemisphere in all three-scale networks. This is consistent to the left-hemisphere hypothesis for ASD [Chandana et al., 2005; Chugani et al., 1997]. The hypothesis can explain many of the symptoms of ASD, most notably the language deficits, as language is a left-brain skill. On the other hand, while more connections in the right hemisphere become discriminative in the 403-ROI networks with additional interhemisphere connections. This demonstrates that subtler changes in the connectome can be detected with smaller ROIs.

Our framework uses features extracted from multiple networks constructed from multiscale ROIs. The number of features increases to tens of thousands when more ROIs are used. Our feature selection step removes irrelevant features and eventually retains only hundreds of features. Moreover, the classification accuracy is improved 16%, that is, from 60 to 76%, demonstrating the effectiveness of our feature selection strategy. The *t*-test acts as an effective feature ranking method that is preferable in high dimensional problems because of its computational scalability. Feature subset selection methods such as LASSO select a subset of features that may have better discriminative power jointly. The latter methods usually have better classification performance, but its high computational cost limits their applications to those problems. In our case, approximately a quarter of a million of features were retrieved from the nine networks. Simply applying the subset feature selection may not be realistic. It is possible that the *t*-test may miss some of the features that were irrelevant individually yet relevant when considered jointly. However, the *t*-test filtered much more irrelevant features and reduced the computational cost for the LASSO logistic regression. It has been proved to be an effective feature selection step in other studies [Fan et al., 2007; Wee et al., 2014]. The fact that the logistic regression performs better than the linear regression (76 vs. 71%) implies that the relationship between the features and the labels may not be linear. This observation is also consistent with the observation that the SVM with RBF kernel performs better than the SVM with linear kernel (76 vs. 70%).

The fiber count is a common measure used in the connectivity analysis. However, the mean values of FA, MD, and fiber length take the whole tract path between two ROIs into account. Adding or missing a few more fibers will probably not change their overall characteristics as

much as that of fiber count. Therefore, they provide more reliable connectivity features than fiber count, and it has been demonstrated by our experiment results (i.e., the mean ACC of FA at the 90-ROI scale was 0.675 vs. that of fiber count at 0.586).

Many factors, such as field strength [Dennis et al., 2013; Zhan et al., 2013a], scanners [Zhan et al., 2014a], feature space [Zhan et al., 2014b], and imaging acquisition parameters [Zhan et al., 2012, 2013b], affect the construction of connectivity network. The choice of diffusion models (e.g., tensor, multitensor, and spherical deconvolution), tracking parameters, seed locations, maximum turning angle, and stopping criteria will also introduce variability in network construction and classification results [Dennis et al., 2015b; Zhan et al., 2013c, 2015]. Tractography is even more challenging given that the image quality at the 6-month-old is generally insufficient, both in terms of contrast and resolution. In the future, more efforts need to be dedicated to refining these tractography and imaging parameters.

CONCLUSION

We propose a classification framework that uses the complementary information from multiparameter multiscale structural networks for identification of 6-month-old infants who are at high-risk for ASD. The experiment results confirm the effectiveness of the proposed framework. The discriminative connections identified in our study may serve as potential imaging connectomic biomarkers for ASD diagnosis or prediction. Our method can also be generalized for detecting other WM related diseases.

ACKNOWLEDGMENTS

Data and/or research tools used in the preparation of this manuscript were obtained from the NIH supported National Database for Autism Research (NDAR). NDAR is a collaborative informatics system created by the National Institutes of Health to provide a national resource to support and accelerate research in autism. Dataset have been deposited in NDAR Study 385. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or of the Submitters submitting original data to NDAR. The authors declare that there is no conflict of interest.

REFERENCES

Abrahams BS, Geschwind DH (2008): Advances in autism genetics: On the threshold of a new neurobiology. Nat Rev Genet 9: 341–355.

Alexander AL, Lee JE, Lazar M, Boudos R, DuBray MB, Oakes TR, Miller JN, Lu J, Jeong EK, McMahon WM, Bigler ED, Lainhart JE (2007): Diffusion tensor imaging of the corpus callosum in autism. Neuroimage 34:61–73.

Ardekani BA, Tabesh A, Sevy S, Robinson DG, Bilder RM, Szeszko PR (2011): Diffusion tensor imaging reliably differentiates patients with schizophrenia from healthy volunteers. Hum Brain Mapp 32:1–9.

Aurenhammer F (1991): Voronoi diagrams – a survey of a fundamental geometric data structure. ACM Comput Surv 23:345–405.

Avants BB, Epstein CL, Grossman M, Gee JC (2008): Symmetric diffeomorphic image, registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. Med Image Anal 12:26–41.

Bailey A, Le Couteur A, Gottesman I, Bolton P, Simonoff E, Yuzda E, Rutter M (1995): Autism as a strongly genetic disorder: Evidence from a British twin study. Psychol Med 25:63–77.

Barnea-Goraly N, Lotspeich LJ, Reiss AL (2010): Similar white matter aberrations in children with autism and their unaffected siblings: A diffusion tensor imaging study using tract-based spatial statistics. Arch Gen Psychiatry 67:1052–1060.

Basser PJ, Mattiello J, LeBihan D (1994): MR diffusion tensor spectroscopy and imaging. Biophys J 66:259–267.

Belmonte MK, Allen G, Beckel-Mitchener A, Boulanger LM, Carper RA, Webb SJ (2004): Autism and abnormal development of brain connectivity. J Neurosci 24:9228–9231.

Bosl W, Tierney A, Tager-Flusberg H, Nelson C (2011): EEG complexity as a biomarker for autism spectrum disorder risk. BMC Med, 9–18. doi: 10.1186/1741-7015-9-18.

Calhoun VD, Maciejewski PK, Pearlson GD, Kiehl KA (2008): Temporal lobe and "default" hemodynamic brain modes discriminate between schizophrenia and bipolar disorder. Hum Brain Mapp 29:1265–1275.

Cammoun L, Gigandet X, Meskaldji D, Thiran JP, Sporns O, Do KQ, Maeder P, Meuli R, Hagmann P (2012): Mapping the human connectome at multiple scales with diffusion spectrum MRI. J Neurosci Methods 203:386–397.

Centers for Disease Control and Prevention (2014): Prevalence of autism spectrum disorder among children aged 8 years – autism and developmental disabilities monitoring network, 11 sites, United States, 2010. Morbidity and Mortality Weekly Report, Surveillance Summaries 63:1–21.

Chandana SR, Behen ME, Juhász C, Muzik O, Rothermel RD, Mangner TJ, Chakraborty PK, Chugani HT, Chugani DC (2005): Significance of abnormalities in developmental trajectory and asymmetry of cortical serotonin synthesis in autism. Int J Dev Neurosci 23:171–182.

Cheon KA, Kim YS, Oh SH, Park SY, Yoon HW, Herrington J, Nair A, Koh YJ, Jang DP, Kim YB, Leventhal BL, Cho ZH, Castellanos FX, Schultz RT (2011): Involvement of the anterior thalamic radiation in boys with high functioning autism spectrum disorders: A diffusion tensor imaging study. Brain Res 1417:77–86.

Chugani DC, Muzik O, Rothermel R, Behen M, Chakraborty P, Mangner T, da Silva EA, Chagani HT (1997): Altered serotonin synthesis in the dentatothalamocortical pathway in autistic boys. Ann Neurol 42:666–669.

Constantino JN, Zhang Y, Frazier T, Abbacchi AM, Law P (2010): Sibling recurrence and the genetic epidemiology of autism. Am J Psychiatry 167:1349–1356.

Courchesne E, Pierce K (2005): Why the frontal cortex in autism might be talking only to itself: Local over-connectivity but long-distance disconnection. Curr Opin Neurobiol 15:225–230.

Daianu M, Jahanshad N, Nir TM, Toga AW, Jack JCR, Weiner MW, Thompson PM (2013): Alzheimer's disease neuroimaging initiative. Breakdown of brain connectivity between normal aging and Alzheimer's disease: A structural k-core network analysis. Brain Connect 3:407–422.

Daianu M, Jahanshad N, Nir TM, Jack CRJ, Weiner MW, Bernstein MA, Thompson PM, Alzheimer's Disease Neuroimaging Initiative (2015): Rich club analysis in the Alzheimer's disease connectome reveals a relatively undisturbed structural core network. Hum Brain Mapp 36:3087–3103.

Davatzikos C, Fan Y, Wu X, Shen D, Resnick SM (2008): Detection of prodromal Alzheimer's disease via pattern classification of magnetic resonance imaging. Neurobiol Aging 29:514–523.

Dennis EL, Zhan L, Jahanshad N, Mueller BA, Jin Y, Lenglet C, Yacoub E, Sapiro G, Ugurbil K, Harel N, Toga AW, Lim KO, Thompson PM (2013): Rich club analysis of structural brain connectivity at 7 Tesla versus 3 Tesla. Comput Diffusion MRI Brain Connectivity Math Visualization 209–218.

Dennis EL, Ellis MU, Marion SD, Jin Y, Moran L, Olsen A, Kernan C, Babikian T, Mink R, Babbitt C, Johnson J, Giza CC, Thompson PM, Asarnow RF (2015a): Callosal function in pediatric traumatic brain injury linked to disrupted white matter integrity. J Neurosci 35:10202–10211.

Dennis EL, Jin Y, Kernan C, Babikian T, Mink R, Babbitt C, Johnson J, Giza CC, Asarnow RF, Thompson PM (2015b): White matter integrity in traumatic brain injury: Effects of permissible fiber turning angle. In: Proceedings of IEEE International Symposium Biomedical Imaging, New York, NY, pp. 930–933.

Dennis EL, Jin Y, Villalon-Reina JE, Zhan L, Kernan CL, Babikian T, Mink RB, Babbitt CJ, Johnson JL, Giza CC, Thompson PM, Asarnow RF (2015c): White matter disruption in moderate/severe pediatric traumatic brain injury: Advanced tract-based analyses. Neuroimage Clin 7:493–505.

Di Martino A, Kelly C, Grzadzinski R, Zuo XN, Mennes M, Mairena MA, Lord C, Castellanos FX, Milham MP (2011): Aberrant striatal functional connectivity in children with autism. Biol Psychiatry 69:847–856.

Duffy FH, Als H (2012): A stable pattern of EEG spectral coherence distinguishes children with autism from neuro-typical controls-A large case control study. BMC Med 10:64.

Elison JT, Paterson SJ, Wolff JJ, Reznick JS, Sasson NJ, Gu H, Botteron KN, Dager SR, Estes AM, Evans AC, Gerig G, Hazlett HC, Schultz RT, Styner M, Zwaigenbaum L, Piven J, IBIS Network (2013) White matter microstructure and atypical visual orienting in 7-month-olds at risk for autism. Am J Psychiatry 170:899–908.

Elsabbagh M, Fernandes J, Jane Webb S, Dawson G, Charman T, Johnson MH, British Autism Study of Infant Siblings Team (2013): Disengagement of visual attention in infancy is associated with emerging autism in toddlerhood. Biol Psychiatry 74: 189–194.

Fan Y, Shen D, Gur RC, Gur RE, Davatzikos C (2007): COMPARE: Classification of morphological patterns using adaptive regional elements. IEEE Trans Med Imaging 26:93–105.

Fornito A, Zalesky A, Bullmore ET (2010): Network scaling effects in graph analytic studies of human resting-state FMRI data. Front Syst Neurosci 17:4–22.

Gillberg C (1993): Autism and related behaviors. J Intellect Disabil Res 37:343–372.

Guo Y, Wu G, Commander LA, Szary S, Jewells V, Lin W, Shen D (2014): Segmenting hippocampus from infant brains by sparse patch matching with deep-learned features. MICCAI LNCS 8674. pp 308–315.

Guyon I, Gunn S Nikravesh M, Zadeh L (2006): Feature Extraction: Foundations and Applications, New York: Springer-Verlag. p 778.

Hagmann P (2005): From diffusion MRI to brain connectomics. Lausanne: Ecole Polytech-nique Fédérale de Lausanne (EPFL).

Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ, Sporns O (2008): Mapping the structural core of human cerebral cortex. PLoS Biol 6:e159.

Haller S, Badoud S, Nguyen D, Garibotto V, Lovblad KO, Burkhard PR (2012): Individual detection of patients with Parkinson disease using support vector machine analysis of diffusion tensor imaging data: Initial results. AJNR Am J Neuroradiol 33:2123–2128.

Ingalhalikar M, Parker D, Bloy L, Roberts TP, Verma R (2012): Diffusion based abnormality markers of pathology: Toward learned diagnostic prediction of ASD. Neuroimage 57:918–927.

Jahanshad N, Aganj I, Lenglet C, Joshi A, Jin Y, Barysheva M, McMahon KL, de Zubicaray GI, Martin NG, Wright MJ, Toga AW, Sapiro G, Thompson PM (2011): Sex differences in the human connectome: 4-Tesla high angular resolution diffusion imaging (HARDI) tractography in 234 young adult twins. In: Proceedings of IEEE International Symposium on Biomedical Imaging, Chicago, IL. pp 939–943.

Jin Y, Shi Y, Zhan L, Li J, de Zubicaray GI, McMahon KL, Martin NG, Wright MJ, Thompson PM (2012): Automatic population HARDI white matter tract clustering by label fusion of multiple tract atlases. MICCAI Multimodal Brain Image Analysis Workshop, LNCS 7509, pp 147–156.

Jin Y, Shi Y, Zhan L, de Zubicaray GI, McMahon KL, Martin NG, Wright MJ, Thompson PM (2013): Labeling white matter tracts in HARDI by fusing multiple tract atlases with applications to genetics. In: Proceedings of IEEE International Symposium on Biomedical Imaging, San Francisco, CA. pp 512–515.

Jin Y, Shi Y, Zhan L, Gutman BA, de Zubicaray GI, McMahon KL, Wright MJ, Toga AW, Thompson PM (2014): Automatic clustering of white matter fibers in brain diffusion MRI with an application to genetics. Neuroimage 100:75–90.

Jin Y, Shi Y, Zhan L, Thompson PM (2015): Automated Multi-atlas labeling of the fornix and its integrity in Alzheimer's disease. In: Proceedings of IEEE International Symposium Biomedical Imaging, New York, NY, pp 140–143.

Johnson MH, Griffin R, Csibra G, Halit H, Farroni T, de Haan M, Tucker LA, Baron-Cohen S, Richards J (2005): The emergence of the social brain network: Evidence from typical and atypical development. Dev Psychopathol 17:599–619.

Just MA, Keller TA, Malave VL, Kana RK, Varma S (2012): Autism as a neural systems disorder: A theory of frontal-posterior underconnectivity. Neurosci Biobehav Rev 36:1292–1313.

Klein JC, Behrens TE, Robson MD, Mackay CE, Higham DJ, Johansen-Berg H (2007): Connectivity-based parcellation of human cortex using diffusion MRI: Establishing reproducibility, validity and observer independence in BA 44/45 and SMA/pre-SMA. Neuroimage 34:204–211.

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009): Circos: An information aesthetic for comparative genomics. Genome Res 19:1639–1645.

Landa R, Holman KC, Garrett-Mayer E (2007): Social and communication development in toddlers with early and later diagnosis of autism spectrum disorders. Arch Gen Psychiatry 64: 853–864.

Lauritsen MB, Pedersen CB, Mortensen PB (2005): Effects of familial risk factors and place of birth on the risk of autism: a nationwide register-based study. J Child Psychol Psychiatry 46:963–971.

Lee SH, Kubicki M, Asami T, Seidman LJ, Goldsteine JM, Mesholam-Gately MJ, McCarley RW, Shenton ME (2013):

Extensive white matter abnormalities in patients with first-episode schizophrenia: A diffusion tensor imaging (DTI) study. Schizophr Res 143:231–238.

Lewis JD, Evans AC, Pruet JR, Botteron K, Zwaigenbaum L, Estes A, Gerig G, Collins L, Kostopoulos P, McKinstry R, Dager S, Paterson S, Schultz RT, Styner M, Hazlett H, Piven J (2014): Network inefficiencies in autism spectrum disorder at 24 months. Transl Psychiatry 4:e388.

Li J, Jin Y, Shi Y, Dinov ID, Wang DJ, Toga AW, Thompson PM (2013): Voxelwise spectral diffusional connectivity and its applications to Alzheimer's disease and intelligence prediction. MICCAI LNCS 8149, pp 655–662.

Liu F, Wee CY, Chen H, Shen D (2014): Inter-modality relationship constrained multi-modality multi-task feature selection for Alzheimer's disease and mild cognitive impairment identification. Neuroimage 84:466–475.

Liu M, Zhang D, Shen D, The Alzheimer's Disease Neuroimaging Initiative (2015): View-centralized multi-atlas classification for Alzheimer's disease diagnosis. Hum Brain Mapp 36:1847–1865.

Lord C, Jones RM (2012): Annual research review: Re-thinking the classification of autism spectrum disorders. J Child Psychol Psychiatry 53:490–509.

Ozonoff S, Iosif AM, Baguio F, Cook IC, Hill MM, Hutman T, Rogers SJ, Rozga A, Sangha S, Sigman M, Steinfeld MB, Yong GS (2010): A prospective study of the emergence of early behavioral signs of autism. J Am Acad Child Adolesc Psychiatry 49:256–266.

Pan S, Iplikci S, Warwick K, Aziz TZ (2012): Parkinson's Disease tremor classification – a comparison between support vector machines and neural networks. Expert Syst Appl 39:10764–10771.

Price T, Wee CY, Gao W, Shen D (2014): Multiple-Network Classification of Childhood Autism Using Functional Connectivity Dynamics. Med Image Comput Comput Assit Interv 17:177–184.

Qiu A, Adler M, Crocetti D, Miller MI, Mostofsky SH (2010): Basal ganglia shapes predict social, communication, and motor dysfunctions in boys with autism spectrum disorder. J Am Acad Child Adolesc Psychiatry 49:539–551.

Rakotomamonjy A, Bach FR, Canu S, Grandvalet Y (2008): SimpleMKL. J Mach Learn Res 9:2491–2521.

Redcay E (2008): The superior temporal sulcus performs a common function for social and speech perception: Implications for the emergence of autism. Neurosci Biobehav Rev 32:123–142.

Ritvo ER, Jorde LB, Mason-Brothers A, Freeman BJ, Pingree C, Jones MB, McMahon WM, Petersen PB, Jenson WR, Mo A (1989): The UCLA-University of Utah epidemiologic survey of autism: Recurrence risk estimates and genetic counseling. Am J Psychiatry 146:1032–1036.

Scholkopf B, Smola A. 2001. Learning with Kernels, Cambridge MA: The MIT Press. p 648.

Schultz RT, Gauthier I, Klin A, Fulbright RK, Anderson AW, Volkmar F, Skudlarski P, Lacadie C, Cohen DJ, Gore JC (2000): Abnormal ventral temporal cortical activity during face discrimination among individuals with autism and Asperger syndrome. Arch Gen Psychiatry 57:331–340.

Shi F, Yap PT, Wu G, Jia H, Gilmore JH, Lin W, Shen D (2011): Infant brain atlases from neonates to 1- and 2-year-olds. PLos One 6:e18746.

Shi F, Yap PT, Gao W, Lin W, Gilmore JH, Shen D (2012a): Altered structural connectivity in neonates at genetic risk for schizophrenia: a combined study using morphological and white matter networks. Neuroimage 62:1622–1633.

Shi F, Wang L, Dai Y, Gilmore JH, Lin W, Shen D (2012b): LABEL: Pediatric brain extraction using learning-based meta-algorithm. Neuroimage 62:1975–1986.

Shi F, Wang L, Peng Z, Wee CY, Shen D (2013): Altered modular organization of structural cortical networks in children with autism. PLos One 8:e63131.

Sled JG, Zijdenbos AP, Evans AC (1998): A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Trans Med Imaging 17:87–97.

Smith SM (2002): Fast robust automated brain extraction. Hum Brain Mapp 17:143–155.

Spencer MD, Holt RJ, Chura LR, Suckling J, Calder AJ, Bullmore ET, Baron-Cohen S (2011): A novel functional brain imaging endophenotype of autism: the neural response to facial expression of emotion. Transl Psychiatry 1, e19.

Sporns O, Tononi G, Kötter R (2005): The human connectome: A structural description of the human brain. PLoS Comput Biol 1:e42.

Sumiyoshi C, Kawakubo Y, Suga M, Sumiyoshi T, Kasai K (2011): Impaired ability to organize information in individuals with autism spectrum disorders and their siblings. Neurosci Res 69: 252–257.

Tibshirani R (1996): Regression shrinkage and selection via the Lasso. J R Stat Soc Ser B 58:267–288.

Tomassini V, Jbabdi S, Klein JC, Behrens TE, Pozzilli C, Matthews PM, Ruthworth MF, Johansen-Berg H (2007): Diffusion-weighted imaging tractography-based parcellation of the human lateral premotor cortex identifies dorsal and ventral subregions with anatomical and functional specializations. J Neurosci 27:10259–10269.

Torgerson CM, Irimia A, Leow AD, Bartzokis G, Moody TD, Jennings RG, Alger JR, Van Horn JD, Altshuler LL (2013): DTI tractography and white matter fiber tract characteristics in euthymic bipolar I patients and healthy control subjects. Brain Imaging Behav 7:129–139.

Tsiaras V, Simos PG, Rezaie R, Sheth BR, Garyfallidis E, Castillo EM, Papanicolaou AC (2011): Extracting biomarkers of autism from MEG resting-state functional connectivity networks. Comput Biol Med 41:1166–1177.

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 15:273–289.

Uddin LQ, Menon V (2009): The anterior insula in autism: Underconnected and under-examined. Neurosci Biobehav Rev 33: 1198–1203.

van den Heuvel MP, Mandl RC, Stam CJ, Kahn RS, Hulshoff Pol EH (2010): Aberrant frontal and temporal complex network structure in schizophrenia: A graph theoretical analysis. J Neurosci 30:15915–15926.

Wee CY, Yap PT, Li W, Denny K, Brwondyke JN, Potter GG, Welsh-Bohmer KA, Wang L, Shen D (2011): Enriched white matter connectivity networks for accurate identification of MCI patients. Neuroimage 54:1812–1822.

Wee CY, Wang L, Shi F, Yap PT, Shen D (2014): Diagnosis of autism spectrum disorders using regional and interregional morphological features. Hum Brain Mapp 35:3414–3430.

Wetherby AM, Woods J, Allen L, Cleary J, Dickinson H, Lord C (2004): Early indicators of autism spectrum disorders in the second year of life. J Autism Dev Disord 34:473–493.

Wing L (1997): The autistic spectrum. Lancet 350:1761–1766.

Wolf JJ, Gu H, Gerig G, Elison JT, Styner M, Gouttard S, Botteron KN, Dager SR, Dawson G, Estes AM, Evans AC, Hazlett HC, Kostopoulos P, McKinstry RG, Paterson SJ, Schultz RT, Zwaigenbaum L, Piven J, IBIS Network (2012): Differences in white matter fiber tract development present from 6 to 24 months in infants with Autism. Am J Psychiatry 169:589–600.

Yap PT, Shen D (2012): Spatial transformation of DWI data using non-negative sparse representation. IEEE Trans Med Imaging 31:2035–2049.

Yap PT, An H, Chen Y, Shen D (2014): Fiber-driven resolution enhancement of diffusion-weighted images. Neuroimage 84:939–950.

Zalesky A, Fornito A, Harding IH, Cocchi L, Yücel M, Pantelis C, Bullmore ET (2010): Whole-brain anatomical networks: Does the choice of nodes matter? Neuroimage 50:970–983.

Zalesky A, Fornito A, Seal ML, Cocchi L, Westin CF, Bullmore ET, Egan GF, Pantelis C (2011): Biol Psychiatry 69:80–89.

Zhan L, Franc D, Patel V, Jahanshad N, Jin Y, Mueller BA, Bernstein MA, Borowski BJ, Jack CR Jr, Toga AW, Lim KO, Thompson PM (2012): How do spatial and angular resolution affect brain connectivity maps from diffusion MRI? In: Proceedings of IEEE International Symposium Biomedical Imaging, Barcelona, Spain. pp 1–4.

Zhan L, Mueller BA, Jahanshad N, Jin Y, Lenglet C, Yacoub E, Sapiro G, Ugurbil K, Harel N, Toga AW, Lim KO, Thompson PM (2013a): Magnetic resonance field strength effects on diffusion measures and brain connectivity networks. Brain Connect 3:72–86.

Zhan L, Jahanshad N, Ennis DB, Jin Y, Bernstein MA, Borowski BJ, Jack CR,J, Toga AW, Leow AD, Thompson PM (2013b): Angular versus spatial resolution trade-offs for diffusion imaging under time constraints. Hum Brain Mapp 34:2688–2706.

Zhan L, Jahanshad N, Jin Y, Toga AW, McMahon MK, de Zubicaray GI, Martin NG, Wright MJ, Thompson PM (2013c): Brain network efficiency and topology depend on the fiber tracking method: 11 tractography algorithms compared in 536 subjects. In: Proceedings of IEEE International Symposium Biomedical Imaging, San Francisco, CA. pp 1134–1137.

Zhan L, Jahanshad N, Jin Y, Nir TM, Leonardo C, Bernstein MA, Borowski BJ, Jack CR Jr, Thompson PM (2014a): Understanding scanner upgrade effects on brain integrity & connectivity measures. In: Proceedings of IEEE International Symposium Biomedical Imaging, Beijing, China. pp 234–237.

Zhan L, Nie Z, Ye J, Jin Y, Jahanshad N, Prasad G, de Zubicaray GI, McMahon KL, Martin NG, Wright MJ, Thompson PM (2014b): Multiple stages classification of Alzheimer's disease based on structural brain networks using generalized low rank approximation (GLRAM). Comput Diffus MRI Math Visualization 35–44.

Zhang D, Wang Y, Zhou L, Yuan H, Shen D, The Alzheimers Disease Neuroimaging Initiative (2011): Multimodal classification of Alzheimer's disease and mild cognitive impairment. Neuroimage 55:856–867.

Zhu D, Li K, Guo L, Jiang X, Zhang T, Zhang D, Chen H, Deng F, Faraco C, Jin C, Wee CY, Yuan Y, Lv P, Yin Y, Hu X, Duan L, Hu X, Han J, Wang L, Shen D, Miller LS, Li L, Liu T (2013): DICCCOL: Dense individualized and common connectivity-based cortical landmarks. Cereb Cortex 23:786–800.

Zhu D, Li K, Terry DP, Puente AN, Wang L, Shen D, Miller LS, Liu T (2014): Connectome-scale assessments of structural and functional connectivity in MCI. Hum Brain Mapp 35:2911–2923.

Zhu X, Suk HI, Shen D (2014): A novel matrix-similarity based loss function for joint regression and classification in AD diagnosis. Neuroimage 100:91–105.

Zhu X, Suk HI, Lee SW, Shen D (2015a): Canonical feature selection for joint regression and multi-class identification in Alzheimer's disease diagnosis. Brain Imaging Behav (in press). doi: 10.1007/s11682-015-9430-4.

Zhu X, Suk HI, Lee SW, Shen D (2015b): Subspace regularized sparse multi-task learning for multi-class neurodegenerative disease identification. IEEE Trans Bio Eng (in press). doi: 10.1109/TBME.2015.2466616.

Zwaigenbaum L, Bryson S, Rogers T, Roberts W, Brian, J Szatmari P (2005): Behavioral manifestations of autism in the first year of life. Int J Dev Neurosci 23:143–152.