Utilizing brain measures for large-scale classification of Autism applying EPIC

Marc B. Harrison*[a], Brandalyn C. Riedel *[a], Gautam Prasad[a], Neda Jahanshad[a] Joshua Faskowitz[a], Paul M. Thompson[a]

[a] Imaging Genetics Center, Mark and Mary Stevens Neuroimaging and Informatics Institute, Keck School of Medicine, University of Southern California, Marina del Rey, CA 90292, USA

ABSTRACT

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder with atypical cortical maturation leading to a deficiency in social cognition and language. Numerous studies have attempted to classify ASD using brain measurements such as cortical thickness, surface area, or volume with promising results. However, the underpowered sample sizes of these studies limit external validity and generalizability at the population level. Large scale collaborations such as Enhancing NeuroImaging Genetics through Meta Analysis (ENIGMA) or the Autism Brain Imaging Data Exchange (ABIDE) aim to bring together like-minded scientists to further improve investigations into brain disorders. To the best of our knowledge, this study represents the largest classification analysis for detection of ASD vs. healthy age and sex matched controls using cortical thickness brain parcellations and intracranial volume normalized surface area and subcortical volumes. We were able to increase classification accuracy overall from 56% to 60% and for females only by 6%. These novel findings using Evolving Partitions to Improve Connectomics (EPIC) underscore the importance of large-scale data-driven approaches and collaborations in the discovery of brain disorders.

Keywords: Autism Spectrum Disorder, EPIC, classification, machine learning

1. INTRODUCTION

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental disorder currently reported to impact 1 in 68 children in the United States[1]. The growing prevalence of this disorder has led to an increase in data-driven neuroimaging studies that have looked at potential ways to classify ASD on metrics such as volume based morphometry[2], cortical thickness[3], DTI[4], and functional connectivity[5]. Importantly, as traditional qualitative radiological methods are unable to distinguish neuroanatomical scans of patients with autism from healthy controls, diagnoses are currently made based on clinical evaluations. However, identification of brain imaging markers of autism could help to identify mechanisms of risk, as well as lead to improvements in evaluating therapeutic interventions and patient outcomes. While multiple studies have produced insightful associations of the brain with ASD, classification studies often remain underpowered. Insufficiently powered studies in neuroscience and imaging have led to the creation of large-scale consortium projects such as ENIMGA (http://enigma.ini.usc.edu/) and open access datasets (http://fcon_1000.projects.nitrc.org/) that provide an opportunity to improve disease detection and increase external validity. Data-driven techniques such as machine learning that look at features of a neurological disorder and learn the best way to classify disease based on these features have been used to improve detection of ASD. While some studies have produced intriguing results with high classification accuracy, these studies are typically carried out with small disease populations[6,7]. Consequently, a recent study looking at the ability of significant variables to translate into predictors for a given disease found that often this criterion doesn't hold, suggesting a novel approach to classification was warranted[8]. Indeed, a large selection of the ASD literature relies extensively on linear regression to localize brain areas contributing to ASD[3,9]. Improvements in detection of variables associated with ASD, do not necessarily lead to improvements in prediction of the disease. In our study we utilized an open-access worldwide dataset from 23 sites involved in the Autism Brain Imaging Data Exchange (ABIDE). The total sample size used here is N=1114, with 564 individuals with ASD, and 550 age- and sex-matched controls. To the best of our knowledge, this is the largest study to date that has used brain MRI features for the classification of ASD.

2. METHODOLOGY

2.1 Sample data

1592 subjects from ABIDE I & II with imaging data were downloaded from the International Data Sharing Initiative. All data were fully anonymized and collected in accordance to each participating site's institutional review board for human subjects with informed consent procedures. Imaging protocols for each site are available at (http://fcon_1000.projects.nitrc.org/indi/abide/). Structural T1-weighted (T1w) images for all available subjects were aligned using FSL's reorient2std and robustfov to center images[10]. T1w images were then run through FreeSurfer's (version 5.3) recon-all pipeline to partition the cortex into 34 regions of interest per hemisphere using a standard atlas. Thereafter, statistical mapping of distinct subcortical volumes (SV), surface areas (SA), and cortical thickness (CT) were extracted[11]. Each subject's FreeSurfer segmentations were visually inspected for accuracy using ENIGMA quality control protocols. Surface area and subcortical volume were then normalized relative to their intra-cranial volume (ICV) to control for head size and age effects. In order to achieve stratified k-fold cross-validation during classification, and an equal proportion of patients and controls within each training and test fold, we matched group sample size within sex and between Autism and the control group. Group differences were found for full-scale intelligence quotient (FIQ). As ASD is a developmental disorder, we chose to look at classification of individual's ages 5-21 (N = 1,114). The subject demographics are summarized in **Table 1.**

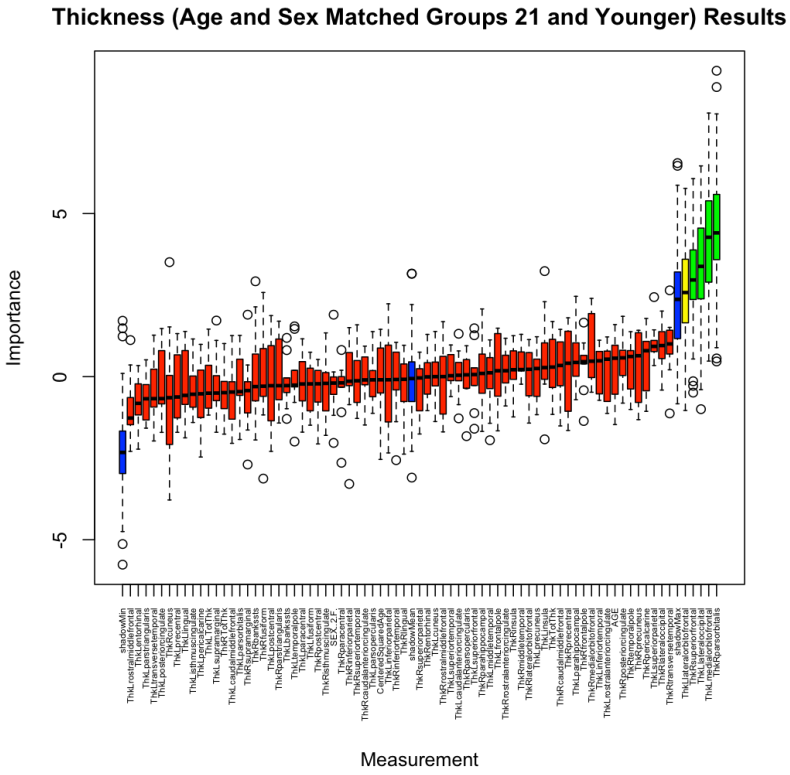
Table 1. Demographics of study, *p <0.05

		Total	Age (Mean ±SD)	Age range	FIQ (Mean ±SD)
Controls	(N=550)		13.2 (3.8)	5-22	113.2 (12.7)
	Male	478	13.2 (3.8)		113.2 (12.6)
	Female	72	12.0 (3.6)		113.2 (13.2)
Autism	(N=564)		13.2 (3.9)	5-22	106.0 (17.0)*
	Male	492	13.4 (3.9)		106.2 (17.1)*
	Female	72	11.8 (3.3)		104.7 (16.9)*

2.2 Variable selection with Boruta

Feature selection is the first general step for classification tasks, often boosting detection of disease, and improving clinical relevancy. Boruta is an all-relevant feature selection wrapper algorithm, capable of working with any classification method that outputs a measure of variable importance. Here, we used Random Forest with Boruta on a random subset of our data to accomplish variable selection. As a developmental disorder with clear sex differences, we chose to look at classification of individual's ages 5-21 both as a whole, and within sexes. Therefore, we ran Boruta for variable selection with a subset of 3 groups: males and females combined, males only, and females only. Boruta performs a top-down search for relevant features by comparing the original attributes with randomly permuted copies of the attribute, and ranking features against the permuted shadow attributes by their average importance over 500 runs. Attributes that perform significantly worse than their shadow attributes are given an importance of 0. Attributes that are significantly better than their shadow attributes are ranked with greater importance as seen in **Figure 1**.

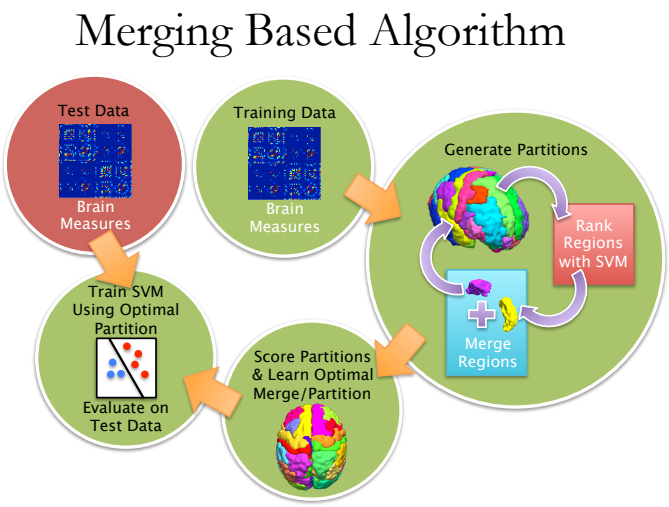
Figure 1. Bourta features selected for thickness



2.3 Determining the optimal combinations of features using EPIC

To optimize the interpretability of results and allow for analyses of SV, CT, and SA together, we modified a new approach, Evolving Partitions to Improve Connectomics (EPIC)[12], to perform merges within each measurement type. The framework behind EPIC involves learning a low-dimensional way to partition the cortex to optimize the classification of disease. EPIC takes as input the MRI parcellations, their diagnostic class labels, and the patients' age, and sex. The EPIC algorithm consists of a top level cross-validation process to separate data into 50 distinct training and test sets that follow from a 5 times repeated stratified 10-fold cross-validation. Training and test data were split such that all patients and controls were in only one random test fold. In each fold of our algorithm, the training data were composed of a random subset of subjects and were used to obtain coefficient weights from the linear support vector machine (SVM) classifier for the corresponding feature. SVM works by finding the optimal hyperplane separating the two classes of patients and controls. Our adaptive algorithm then iteratively merges the two lowest weighted regions to form a new composite feature and learns the optimal number of merges by selecting the configuration that maximizes accuracy between folds. This process is repeated N-1 times, where N corresponds to the total number of features within each imaging class. For each partition, or way of grouping variables, we evaluate its ability to detect group differences between diagnostic classes based on its classification accuracy. Here, this process is carried out for 5 sub-folds and repeated twice. The classification performance for a given partition is then averaged between the sub-folds, and then used to learn a partition on the full training set and evaluated on new test data as seen in **Figure 2** below. Importantly, classification using high-dimensional feature sets is typically aided by prior feature subset selection. This feature selection helps to mitigate both the "curse of dimensionality," in which data sparsity scales exponentially with its dimensionality, and the potential for overfitting to training data.

Figure 2. EPIC workflow



3. RESULTS

3.1 Autism classification

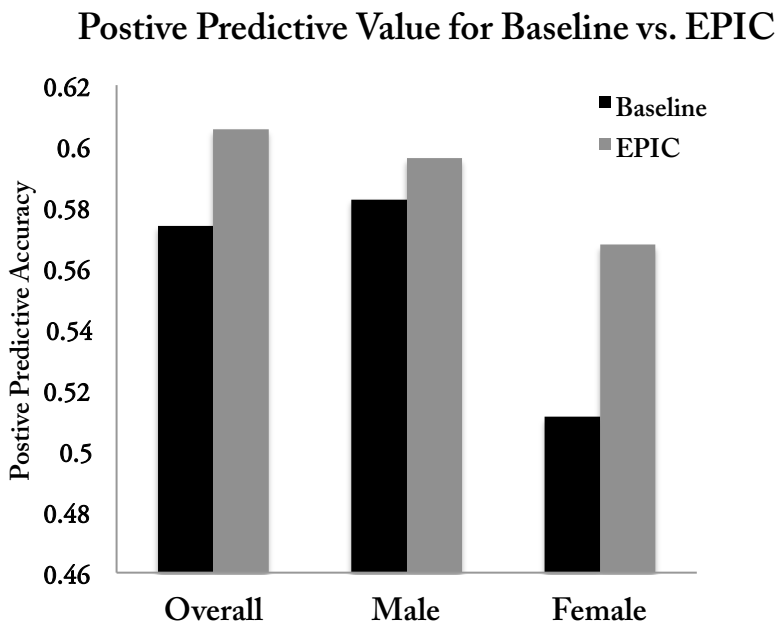
We performed stratified 10-fold cross-validation using the features identified through the 3 independent Boruta runs. **Table 2** summarizes the classification results using the selected measures. To evaluate the robustness of our methods, we repeated the 10-fold cross-validation process 5 times, and compared the average classification results to baseline linear SVM performance using the original features. From baseline performance, we were able to modestly improve classification in all 3 analyses. In the training phase of the algorithm, EPIC chose a reduced dimensionality representation that was roughly 45-55% the size of the original feature set in the final iterations for males, females, and overall. All results also included features that were not merged for each type of measurement (SV, CT, and SA). EPIC was able to increase accuracy by 3.4% on average and performed better than chance using a two-sample t-test (t=-6.09, p<. 001).

Table 2. Baseline vs. EPIC results

		Accuracy	Sensitivity	Specificity	Positive Predictive Value
Baseline					
Anatomic		0.5676	0.5607	0.5745	0.5738
	Males only	0.5653	0.5265	0.6042	0.5825
	Females only	0.5262	0.5142	0.5142	0.5111
EPIC					
Adaptive Average		0.5948	0.5678	0.6218	0.6055
	Males only	0.5741	0.5142	0.6341	0.5960
	Females only	0.5794	0.5714	0.5714	0.5676

In addition, the positive predictive value for classifying MDD increased for all groups with females only improving by 5.65 % when using EPIC as shown in Figure 2.

Figure 2. Positive predictive results



4. CONCLUSION

Our results represent the most comprehensive and largest classification analysis using CT, SA, and SV ever performed with Autism. In females with ASD, EPIC was able to improve classification accuracy using features selected measures from boruta from 51% to 57% and overall from 56% to 60%. As we only considered T1-weighted MRI-based brain measures including cortical thickness, surface areas and subcortical volume in this study future directions will aim to include functional connectivity and larger sample sizes in order to further improve classification of Autism. Additionally, we plan to apply this methodology on additional neurological disorders with ENIGMA including Major Depressive Disorder.

REFERENCES

[1] Centers for Disease Control and Prevention. Prevalence of autism spectrum disorders — autism and developmental disabilities monitoring network, 14 sites, United States. *MMWR Surveill Summ* 2012; 61: 1–19. (2008)

[2] Nickl-Jockschat, T., Hablel, T., Michel, T.M., Manning, J., Laird, A.R., Fox, P.T., Schneider, F., and Eickhoff, S.B., "Brain Structure Anomalies in Autism Spectrum Disorder—A Meta-Analysis of VBM Studies Using Anatomic Likelihood Estimation," Hum Brain Mapp. 33(6), 1470-1489 (2012).

[3] Ecker, E., Shahidiani, A., Feng, Y., Murphy, C., D'Almeida, V., Deoni, S., … and Murphy, D.G.M., "The effect of age, diagnosis, and their interaction on vertex-based measures of cortical thickness and surface area in autism spectrum disorder," J Neural Transm 121, 1157-1170 (2014).

[4] Travers, B.G., Adluru, N., Ennis, E., Tromp, D.P.M., Destiche, D., Doran, S., Bigler, E.D., …and Alexander, A.L., "Diffusion Tensor Imaging in Autism Spectrum Disorder: A Review," Autism Res. 5, 289-313 (2012).
[5] Rudie, J.D., Brown, J.A., Beck-Pancer, D., Hernandez, L.M., Thompson, P.M., Bookheimer., S.Y., and Dapretto, M., "Altered functional and structural brain network organization in autism," NeuroImage: Clinical 2 79-94 (2013).
[6] Ecker, C., Rocha-Rego, V., Johnston, P., Mourao-Miranda, J., Marquand, A., Daly, E.M., Brammer, M.J., and the MRC AIMS Consortium "Investigating the predictive value of whole-brain structural MR scans in autism: A pattern classification approach," NeuroImage 49, 44-56 (2010).
[7] Uddin, L.Q., Menon, V., Young, C.B., Ryali, S., Chen, T., Khouzam, A., …and Hardan, A.Y., "Multivariate Searchlight Classification of Structural Magnetic Resonance Imaging in Children and Adolescents with Autism," Biol. Psychiatry 70, 833-841 (2011).
[8] Lo, A., Chernoff, H., Zheng, T., and Lo, S., "Why significant variables aren't automatically good predictors," PNAS 112(45), 13892-13897 (2105).
[9] Ecker, C., Johnston, P., Palaniyappan, L., Williams, S.C., Baran-Cohen S., and D.G.M., "Brain Surface Anatomy in Adults with Autism," Jama. Psychiatry 70:1, 59-70 (2013).
[10] Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., ... and P.M. Matthews. "Advances in functional and structural MR image analysis and implementation as FSL," NeuroImage, 23(S1), 208-19 (2004)
[11] Fischl, B. (2012). FreeSurfer. Neuroimage, *62*(2), 774-781.
[12] Prasad, G., Joshi, S.H., and Thompson, P.M., "Optimizing brain connectivity networks for disease classification using EPIC," IEEE International Symposium on Biomedical Imaging, 834-837 (2014).