Functional neuroimaging of high-risk 6-month-old infants predicts a diagnosis of autism at 24 months of age

Robert W. Emerson,1,5 Chloe Adams,1,5 Tomoyuki Nishino,1,5 Heather Cody Hazlett,1,3 Jason J. Wolff,4 Lonnie Zwaigenbaum,5 John N. Constantino,2,5 Mark D. Shen,1 Meghan R. Swanson,1 Jed T. Elison,6 Sridhar Karandala,2 Annette M. Estes,7 Kelly N. Boster,2,8 Louis Collins,9 Stephen R. Dager,9 Guido Gerig,12 Hongbin Gu,1 Robert C. McKinstry,8 Sarah Paterson,13 Robert T. Schultz,13 Martin Styner,1 IBIS Network, Bradley L. Schlaggar,2,8,14 John R. Pruett Jr.,2,8 Joseph Piven1,3,†

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by social deficits and repetitive behaviors that typically emerge by 24 months of age. To develop effective early interventions that can potentially ameliorate the defining deficits of ASD and improve long-term outcomes, early detection is essential. Using prospective neuroimaging of 59 6-month-old infants with a high familial risk for ASD, we show that functional connectivity magnetic resonance imaging correctly identified which individual children would receive a research clinical best-estimate diagnosis of ASD at 24 months of age. Functional brain connections were defined in 6-month-old infants that correlated with 24-month scores on measures of social behavior, language, motor development, and repetitive behavior, which are all features common to the diagnosis of ASD. A fully cross-validated machine learning algorithm applied at age 6 months had a positive predictive value of 100% [95% confidence interval (CI), 62.9 to 100], correctly predicting 9 of 11 infants who received a diagnosis of ASD at 24 months (sensitivity, 81.8%; 95% CI, 47.8 to 96.8). All 48 6-month-old infants who were not diagnosed with ASD were correctly classified (specificity, 100% [95% CI, 90.8 to 100]; negative predictive value, 96.0% [95% CI, 85.1 to 99.3]). These findings have clinical implications for early risk assessment and the feasibility of developing early preventative interventions for ASD.

INTRODUCTION

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by deficits in social behavior and the presence of restrictive and repetitive behaviors (1). It is estimated that 1 in 68 children are affected by the disorder (2), and despite tremendous research efforts, ASD still incurs substantial burden to affected individuals, their families, and the community (3, 4). Intervention is critically important, and there is a general consensus that early detection paired with early intervention would have a significant impact on improving outcomes (5–7).

One barrier to early (that is, before 24 months) detection is that the defining behavioral characteristics of ASD generally unfold during the second year of life, typically showing consolidation of the full behavioral syndrome by about 24 months of age or later (8, 9). Behavioral differences in ASD have been observed as early as 6 months of age in characteristics such as gross motor ability, visual reception, and patterns of eye tracking (10–14); however, these associated characteristics have not been able to predict which children will later receive a diagnosis. Given the known plasticity of the brain and behavior during the first year of life, together with the absence of the defining features of the disorder, intervention during this presymptomatic phase, before consolidation of the full syndrome of ASD, is likely to show considerably stronger benefits compared with later treatments (5).

Research on neurodegenerative disorders has shown that changes in the brain are often seen preceding clinical manifestations. For example, in Parkinson’s disease, about 50% of the neurons in the substantia nigra are lost before clinical features become apparent (15). This suggests that brain-related changes appear earlier than behavioral changes and may be useful in predicting future behavioral diagnosis. In ASD, a number of selected morphological (16–18) and electrophysiological (19) brain differences have been reported as early as 6 months of age in infants later diagnosed with ASD; however, the reported group differences in these specific brain structures have not yet shown the sensitivity and specificity required to be effective for the early detection of ASD.

Given the complexity and heterogeneity of ASD, methods for the early detection of ASD using brain metrics will likely require information that is multivariate, complex, and developmentally sensitive. Recent research using functional connectivity magnetic resonance imaging (fcMRI) has linked the functional organization of the human brain to individual cognitive profiles (20–22). These measures of brain functional connectivity are reliable (23) and can accommodate participants as young as neonates (24). Furthermore, in conjunction with machine learning

1 Carolina Institute for Developmental Disabilities, University of North Carolina at Chapel Hill, Chapel Hill, NC 27510, USA. 2 Department of Psychiatry, Washington University School of Medicine, St. Louis, MO 63110, USA. 3 Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC 27514, USA. 4 Department of Educational Psychology, University of Minnesota, Minneapolis, MN 55455, USA. 5 Department of Pediatrics, University of Alberta, Edmonton, Alberta, Canada. 6 Institute of Child Development, University of Minnesota, Minneapolis, MN 55455, USA. 7 Department of Speech and Hearing Sciences, University of Washington, Seattle, WA 98105, USA. 8 Mallinckrodt Institute of Radiology, Washington University, Washington University School of Medicine, St. Louis, MO 63110, USA. 9 Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada. 10 Center on Human Development and Disability, University of Washington, Seattle, WA 98105, USA. 11 Department of Radiology, University of Washington, Seattle, WA 98105, USA. 12 Tandon School of Engineering, New York University, Brooklyn, NY 11201, USA. 13 Center for Autism Research, Department of Pediatrics, Children’s Hospital of Philadelphia, and University of Pennsylvania, Philadelphia, PA 19104, USA. 14 Department of Neurology, Washington University School of Medicine, St. Louis, MO 63110, USA.

*Corresponding author. Email: remerson@med.unc.edu
†These authors contributed equally to this work.
‡Co-senior authors.
approaches, fMRI data have provided predictions of brain maturation (25,26) and diagnostic category (27–32) at the single-subject level. By training machine learning algorithms to identify underlying patterns that separate individuals into these different groups, researchers can predict which group a new individual will likely be in (33).

Here, we postulated that brain functional connectivity at 6 months of age would capture the complexity of ASD and provide a robust method for predicting later diagnosis. Our results revealed that machine learning, applied to fMRI data at 6 months of age in infants at high familial risk for ASD, can accurately predict an ASD diagnosis at 24 months of age.

RESULTS

A cohort of 59 infants with a high familial risk for ASD was included in this study. There were 11 infants diagnosed with ASD at 24 months of age and 48 infants who did not have ASD at 24 months of age. Prospective neuroimaging data were collected from each infant at 6 months of age while they were sleeping naturally. Cognitive, behavioral, and diagnostic assessments were completed at their follow-up visit at 24 months of age.

Brain functional connectivity and infant behavior

A set of 230 regions that were previously defined across the whole brain (see Materials and Methods) were used to create functional connectivity matrices from each participant’s functional MRI data at their 6-month visit. This resulted in 26,335 pairs of regions used for further analysis that represented the whole-brain functional organization of an individual infant. From this complete set, two subsets of brain-behavior features were defined separately for visualizing group discrimination and for the classification analysis.

To visualize the ability of early brain features to discriminate between ASD and non-ASD groups of infants, we computed brain-behavior correlations with each participant’s 24-month assessment scores of social interactions, communication, motor development, and repetitive behavior. Table 1 shows the average raw scores and SE for each of these behavioral measures by group. In addition, between-groups t tests were calculated across all the functional connections, and the intersection of functional connections that showed both a nominal brain-behavior correlation and a between-groups difference \( (P < 0.05) \) was used to define a feature space across all participants. This resulted in a total of 974 functional connections in the 6-month-old brain that showed a relationship with behavior at 24 months and were different between groups. Together, these functional connections constituted <4% of the potential 26,335 total functional connections studied. The participants’ scores on the first and second principal components of this feature space were plotted against each other in Fig. 1, revealing an evident linear separation between the ASD and non-ASD groups.

Predicting individual 24-month clinical diagnoses

To determine whether 6-month-old functional connectivity features were capable of predicting the clinical diagnostic outcome of an individual infant, we used a fully cross-validated approach with a “nested” leave-one-out procedure. In this procedure, the diagnostic outcome of each infant was predicted from an independent training sample, without being used to define features or build the classifier. The features were chosen within each training sample as showing a brain-behavior correlation, creating a feature space that reflected functional connections in 6-month-old infants that showed a relationship with 24-month-old behaviors. This process created 59 sets of features that were used to train individual classifiers. This procedure (detailed in Materials and Methods) allowed

<table>
<thead>
<tr>
<th>Assessment</th>
<th>n (ASD+/ASD−)</th>
<th>Measure</th>
<th>ASD</th>
<th>Non-ASD</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSBS: social communication</td>
<td>50 (8/42)</td>
<td>Social interaction</td>
<td>0.75 ± 0.3</td>
<td>2.51 ± 0.2</td>
</tr>
<tr>
<td>CSBS: social communication</td>
<td>50 (8/42)</td>
<td>Joint attention</td>
<td>1.63 ± 0.6</td>
<td>3.66 ± 0.3</td>
</tr>
<tr>
<td>MSEL: cognitive ability</td>
<td>57 (10/47)</td>
<td>Expressive language</td>
<td>18.91 ± 1.4</td>
<td>20.82 ± 0.6</td>
</tr>
<tr>
<td>MSEL: cognitive ability</td>
<td>58 (11/47)</td>
<td>Fine motor</td>
<td>24.08 ± 1.0</td>
<td>23.84 ± 0.3</td>
</tr>
<tr>
<td>MSEL: cognitive ability</td>
<td>53 (10/43)</td>
<td>Gross motor</td>
<td>24.64 ± 0.5</td>
<td>25.79 ± 0.4</td>
</tr>
<tr>
<td>MSEL: cognitive ability</td>
<td>57 (10/47)</td>
<td>Visual reception</td>
<td>20.18 ± 2.0</td>
<td>23.58 ± 0.6</td>
</tr>
<tr>
<td>MSEL: cognitive ability</td>
<td>58 (11/47)</td>
<td>Receptive language</td>
<td>24.42 ± 0.8</td>
<td>26.53 ± 0.5</td>
</tr>
<tr>
<td>RBS-R: repetitive behavior</td>
<td>47 (9/38)</td>
<td>Self-injurious</td>
<td>2.3 ± 1.0</td>
<td>0.16 ± 0.1</td>
</tr>
<tr>
<td>RBS-R: repetitive behavior</td>
<td>47 (9/38)</td>
<td>Stereotyped</td>
<td>3.4 ± 0.8</td>
<td>0.31 ± 0.1</td>
</tr>
<tr>
<td>RBS-R: repetitive behavior</td>
<td>47 (9/38)</td>
<td>Sameness</td>
<td>4.9 ± 0.9</td>
<td>0.9 ± 0.2</td>
</tr>
<tr>
<td>RBS-R: repetitive behavior</td>
<td>47 (9/38)</td>
<td>Ritualistic</td>
<td>2.2 ± 0.5</td>
<td>0.5 ± 0.1</td>
</tr>
<tr>
<td>RBS-R: repetitive behavior</td>
<td>47 (9/38)</td>
<td>Compulsive</td>
<td>3.4 ± 0.4</td>
<td>0.4 ± 0.1</td>
</tr>
<tr>
<td>RBS-R: repetitive behavior</td>
<td>47 (9/38)</td>
<td>Repetitive</td>
<td>1.8 ± 0.9</td>
<td>0.41 ± 0.1</td>
</tr>
</tbody>
</table>
The public health importance of ASD has been increasingly recognized over the last 15 to 20 years (2, 3). Treatment studies have shown modest effects in improving the core characteristics of ASD (34, 35). Research on infants at high familial risk for ASD has revealed a seemingly narrow window of opportunity, before the age of 24 months, when intervention may have the potential to ameliorate the unfolding of the core features of this disorder (5, 6). Intervention studies with infants at high familial risk for ASD (6, 7) suggest that behavioral intervention in the latter part of the first, and early second, year may be more effective than later (post-diagnosis) intervention. Unfortunately, early behavioral markers have not had sufficient power as predictors of later diagnosis to be clinically useful, and so, to date, methods for presymptomatic detection have not been available.

Our results suggest that early brain metrics, identified on the basis of their association with later ASD-related behaviors, are able to accurately predict an individual infant’s 24-month diagnosis of ASD, by 6 months of age. We focused on predicting diagnostic outcome at 24 months of age, a time when the full syndrome of ASD begins to consolidate and can be reliably diagnosed (8, 9). These findings converge with another MRI study showing that structural information at 6 and 12 months of age can accurately predict an ASD diagnosis (36). Using functional neuroimaging, the current study extends these previous findings by using brain data from a single time point (6 months of age) to accurately predict an ASD diagnosis in 9 of 11 infants at high familial risk for ASD. These findings demonstrate the potential for early detection of autism in infants at high familial risk and serve as a proof of concept that patterns of infant brain measures precede the defining behavioral characteristics of ASD.

Infants with high familial risk for ASD begin life with about a 20% chance of developing ASD (37) compared to ~1.5% in infants with low or unknown risk (2). Because of the 1 in 68 prevalence of ASD in the population, the clinical application of functional neuroimaging is likely to be the most valuable in evaluating infants at high familial risk. Current intervention research that has focused on the first year of life has been limited to studying entire cohorts of infants at high familial risk, with little to no ability to assess a specific individual’s likelihood of receiving a diagnosis beyond the expected recurrence risk. In our sample, even in the lower bound of the CI, the positive predictive value of this classifier shows a higher ability to correctly detect ASD in 6-month-old high-risk infants than has been possible with behavioral screening alone in this age range (38). If these results are replicated in a new high-risk infant cohort, functional neuroimaging at 6 months of age could provide a clinically valuable tool for the detection of ASD in high-risk infants before the development of the full syndrome. This would open the door to randomized controlled trials aimed at identifying effective interventions by recruiting high-risk infants who have been identified as having an even greater risk based on their 6-month neuroimaging assessment.

Although we have taken many precautions to test the internal validity of our classification analysis, there are several limitations that will need to be addressed by future research before the clinical utility of this method can be fully realized. Although our results are strong within this sample of high-risk infants, these findings need to be replicated and extended to an independent high-risk sample of infants. In addition, there is uncertainty associated with a 24-month diagnosis of ASD. Future research will have to address the meaning of this uncertainty with regard to the negative predictive value of the classifier. An effective classifier in the general population would likely require a much
**Fig. 2. Differences in functional connectivity in ASD infants versus non-ASD infants.** Each panel represents the functional connections that show a relationship to scores on each of the behavioral assessments (see Table 1): CSBS (top), MSEL (middle), and RBS-R (bottom). For each assessment, the functional connections associated with individual measures were combined and projected onto a Talairach brain, with the right hemisphere marked (R). The color and thickness of each connection signify the sign and strength of the \( t \) value it represents. Unpaired two-sample \( t \) tests were used to test the difference between group means (ASD versus non-ASD) for each functional connection. Red signifies a connection that shows more negative connectivity in the ASD infant group on average, whereas blue signifies more positive connectivity. \( t \) values were set to a threshold of \( P < 0.005 \) (uncorrected), and the thickness of each bar represents its strength. Coordinates for each sphere are listed in table S1. These calculations are only for visualization and should not be interpreted as differences directly contributing to any individual’s classification.
larger sample to demonstrate its ability to capture the full breadth of the heterogeneity in ASD. Finally, MRI is likely too expensive to be feasible as a general screening tool; however, as genetic information or more advanced screening techniques become available, neuroimaging may be useful as a secondary confirmation of enhanced risk. If these findings could be generalized to more cost-effective and mobile neuroimaging technologies, it would greatly increase the accessibility of early screening. Even with the limited sample size of the present study, the ability of the classifier to predict an individual infant’s later diagnosis is substantial. This high accuracy was maintained when the classifier was trained on a smaller subsample and was used to predict 10 independent infants, suggesting that these results are fairly robust. Therefore, despite the noted limitations, our results suggest that early differences in the brain’s functional connections are useful in predicting a later diagnosis of ASD as early as 6 months of age, well before the onset of the defining behavioral characteristics of ASD. As the field begins to incorporate other risk markers (for example, parental or genetic factors, multimodal imaging, and other measures of infant behavior), we will improve our ability to determine which infants may benefit from early functional neuroimaging.

Our results show that functional neuroimaging with 6-month-old infants at high familial risk for ASD can accurately predict which individuals receive a clinical diagnosis of ASD at 24 months of age. Ultimately, this study represents an initial first step toward developing the earliest diagnostic methods available and may yield the clues necessary to build efficacious early interventions based on individual risk profiles.

Cognitive and behavioral assessments
The RBS-R (42, 43) is a parent/caregiver-rated measure covering a broad range of repetitive behaviors. The RBS-R is a questionnaire that focuses exclusively on restricted/repetitive behaviors. It includes 43 items rated on a four-point scale: 0, behavior does not occur; 1, behavior occurs and is a mild problem; 2, behavior occurs and is a moderate problem; 3, behavior occurs and is a severe problem. Items are grouped into six conceptually derived subscales: stereotyped behavior, self-injurious behavior, compulsive behavior, ritualistic behavior, sameness behavior, and restricted behavior. Scores on these subscales were used to determine brain-behavior features used in the analysis.

The MSEL (44) is a standardized, normed, developmental assessment that provides an overall index of cognitive ability and delay. Children were assessed at 24 months of age, and their scores on the receptive language, expressive language, visual reception, fine motor, and gross motor subscales were used to determine brain-behavior features used in the analysis.

The Communication and Symbolic Behavior Scales Developmental Profile (45) is designed to elicit social and communicative behaviors in infants and was administered at each participant’s 24-month visit. Specifically, their scores on items measuring initiation of joint attention and social interaction were used to determine brain-behavior features in the primary analysis. These items were chosen to reflect specific aspects of behavior that we reasoned to be particularly relevant to social development at about 2 years of age.

Image acquisition
All scans were acquired at IBIS Network clinical sites using cross-site calibrated 3T Siemens TIM Trio scanners (Siemens Medical Solutions) equipped with standard 12-channel head coils. Images were acquired during natural infant sleep without sedation. The IBIS imaging protocol included anatomical images (T1- and T2-weighted), diffusion tensor images [25-direction and 65-direction HARDI DWI (high-angular resolution diffusion imaging/diffusion-weighted imaging)], and resting-state fMRI. This study used a three-dimensional T2-weighted sequence [echo time (TE), 497 ms; repetition time (TR), 3200 ms; matrix, 256 × 256 × 160; voxel size, 1 mm3; sagittal acquisition] and a gradient-echo echo planar image functional sequence (TE, 27 ms; TR, 2500 ms; field of view, 256 mm; matrix, 64 × 64; voxel size, 4 × 4 × 4 mm3; flip angle, 90°; bandwidth, 1906 Hz). All included infants provided data collected during at least two fMRI runs, each run comprising 130 temporally contiguous frames (5.4 min).

fMRI preprocessing
Initial fMRI data preprocessing followed previously described procedures (25, 46, 47) including (i) compensation for slice-dependent time shifts using sinc interpolation, (ii) correction of systematic odd-even slice intensity differences caused by interleaved acquisition, and (iii) spatial realignment to compensate for head motion within and across fMRI runs. Atlas registration of the functional data was achieved by a sequence of affine transforms (individual fMRI average volume → individual T2-weighted → atlas-representative target). All data were registered to an age-specific (6-month) target atlas to handle shape differences across developmental age categories (48). The volumetric time series were resampled in atlas space (3-mm3 voxels) using a resampling procedure that applied all affine registration transform and correction for head movement in a single step. Each atlas-transformed functional data set was visually inspected in sagittal, transverse, and coronal views to exclude potential errors not otherwise identified.

MATERIALS AND METHODS
Study design
Participants were part of the Infant Brain Imaging Study (IBIS), an ongoing longitudinal study of infants at low and high familial risk for ASD. Infants were recruited, screened, and assessed at one of four clinical sites: University of North Carolina, University of Washington, The Children’s Hospital of Philadelphia, and Washington University in St. Louis. The research protocol was approved by the institutional review board at all clinical sites, and parents provided written informed consent after receiving a detailed description of the study. Data were used for research purposes only.

A cohort of 59 (18/41, female/male) infants at high familial risk for ASD was included in this study: 11 infants diagnosed with ASD at 24 months (11 male) and 48 non-ASD infants at 24 months (30 male). High-risk infants were defined as having at least one sibling with an ASD diagnosis. Participants were excluded for comorbid medical or neurological diagnoses influencing growth, development, or cognition; previous genetic conditions; premature birth or low birth weight; maternal substance abuse during pregnancy; contraindication for MRI; or family history of psychosis, schizophrenia, or bipolar disorder.

Diagnostic testing
All infants included in these analyses participated in a comprehensive battery of behavioral assessments including the Autism Diagnostic Observation Schedule (ADOS) (39) and Autism Diagnostic Interview–Revised (40) at 24 months. The ADOS and all other testing and interview data were independently reviewed by experienced clinicians for Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR) (41) criteria for autistic disorder or pervasive developmental disorder not otherwise specified. All ASD-positive infants were assigned a diagnosis according to clinical best estimate using DSM-IV-TR at 24 months of age.
The final visualization feature space was then defined as the intersection frame-to-frame displacement (FD) measure \( t(57) = 0.43; P = 0.79 \) or the number of total frames censored \( t(57) = 0.45; P = 0.75 \).

**fcMRI preprocessing**

In addition to the previously published procedures (52), further preprocessing was conducted before the computation of region-of-interest (ROI) pair time series correlations. Using only the noncensored frames, the data were voxel-wise demeaned and detrended within runs, and nuisance waveforms were regressed out. Nuisance regressors included (i) the time series of three translation \( X, Y, \) and \( Z \) and three rotation \( \text{Pitch}, \text{Yaw}, \) and \( \text{Roll} \) estimates derived by retroactive head motion correction and Volterra expansion derivatives to comprise 24 total motion regressors (53), and (ii) time series derived from the regions of noninterest (whole brain, white matter, and cerebrospinal fluid) and their first derivatives. After nuisance regression, data in frames marked for censoring were replaced by interpolated values computed by least-squares spectral analysis (54, 55). The fMRI data were then temporally filtered to retain frequencies in the \( 0.009 \text{ Hz} < f < 0.08 \text{ Hz} \) band. As a last step, the data were spatially smoothed using a Gaussian kernel (6 mm full width at half maximum isotropic).

**Definition of ROI and correlation computation**

Following Pruett et al. (25), candidate ROIs \( n = 280 \) were adopted from a combination of meta-analyses of ASD studies (46) and of task data and cortical functional areal parcellations obtained in healthy adults (47). Three viewers inspected ROI placements in age-specific atlas templates. Of the 280 ROIs, 50 were partially outside the whole-brain mask and were removed, leaving 230 usable ROIs (25). ROI representative time series were calculated as the average of the time series of each voxel intersecting the 10-mm-diameter sphere located at a given ROI center. Pairwise Pearson correlation values were generated from each of the 26,335 possible pairs of ROIs and then Fisher \( z \)-transformed to improve normality.

**Visualization features and group discrimination**

To create the visualizations in Fig. 1, a leave-one-out cross-validation analysis was performed within the entire group of 59 infants to identify a set of functional connections that were both related to behavior and showed differences between groups. We generated 59 sets of features by iteratively removing one participant from the analysis. For each iteration, the remaining 58 participants were used to define region pairs whose connectivity showed both a nominal Pearson correlation with behavior \( P < 0.05 \) and difference between groups \( t \text{ test}, P < 0.05 \). The final visualization feature space was then defined as the intersection (100% consensus) of these sets. To demonstrate that these features can discriminate between the infant groups, a principal component analysis was used to define the top two dimensions of variance across all participants. Participant’s scores on the first and second principal components of this feature space were plotted against each other in Fig. 1. Information from the classification analysis (see below) was included to visualize which of the infants were incorrectly classified; however, the set of visualization features is used only in Fig. 1 to demonstrate that there is variability in the functional connectivity at the group level that can discriminate between infants who receive a diagnosis at 24 months from those who do not.

**Classification features and prediction of individual infants**

For the classification analysis, one infant was removed from the group to serve as a test case, whereas the remaining 58 infants were used a training set. For each test case in the classification analysis, starting from the full set of 26,335, features were determined within the training set as the functional connections between ROI pairs that show a nominally significant \( P < 0.05 \) behavioral correlation, and a leave-one-out cross-validation analysis was performed within this training set. To be included in the final set of features for the independent test case, a functional connection had to show a correlation with one of the behavioral measures in all the cross-validation sets of training data. This final set of features provided an independently defined set of features that was used to train a classifier with a linear kernel to discriminate between infants who are and are not diagnosed with ASD at 24 months. Finally, the classifier was used to predict the independent test case. This strategy was repeated using each participant as the test case, creating a fully cross-validated approach with a nested leave-one-out procedure to identify features. As a result, the estimation of accuracy was relatively unbiased in the sense that the training features were selected independently of each test case (55). Similar methods are discussed in detail by Pereira et al. (33), with many of the applications reviewed by Gabrieli et al. (29). In the case where a participant did not have behavioral scores, they did not contribute information to the feature selection step (see Table 1). However, their functional connectivity measures were still used to train the classifier to predict the independent infant.

Finally, after completing the classification analysis, we used the intersection of the 59 independent classification feature sets to visualize the functional connections that most likely contributed to the classification accuracy (Fig. 2). Although these features were initially defined by their relationship with behavior only, we performed a between-groups \( t \text{ test} \) and projected the \( t \) values onto a Talairach brain with a threshold of \( P < 0.005 \) (uncorrected). Our logic was that these regions would be the most likely to contribute to the discrimination between groups in the 59 separate support vector machine (SVM) models. This set of classification features represents only a subset of the features used in the individual classifiers. Because each classifier contains a slightly different set of features and is weighted differently, the calculations of group differences should not be interpreted as directly contributing to any individual’s classification. The Talairach coordinates, average connectivity values, \( t \) values, and \( P \) values by group are listed in table S1.

**Predicting individual 24-month clinical diagnoses**

To determine whether brain features were capable of predicting the clinical diagnostic outcome of an individual infant, a classification model, built from an independent group of infants (see above), was applied to their 6-month functional connectivity data. When this process was repeated for all 59 infants, we were able to calculate measures
of classification performance. Sensitivity was calculated as the proportion of infants with ASD that were correctly identified, whereas specificity was calculated as the proportion of infants that did not have ASD that were correctly identified. The positive predictive value was calculated as the proportion of positive predictions that were truly infants with ASD. Conversely, negative predictive value was calculated as the proportion of negative predictions that were truly infants without ASD. Finally, 95% CIs for the reported proportions (specificity, specificity, positive predictive value, and negative predictive value) are calculated according to the efficient score method and corrected for continuity (56).

The significance of the classification accuracy was determined by repeating the entire classification analysis (including feature selection and fitting of the SVM classifier) using randomly shuffled group labels (within the training sets) to predict each test case. This procedure estimated the null distribution of classification accuracy and was repeated 10,000 times to determine what proportion of times a randomly constructed classifier would perform as well as the classifier trained with the correct group labels. This distribution and the observed scores of the correct labels are shown in fig. S2.

**Leave-10-out**

To complete the leave-10-out analysis, the nested cross-validation procedure was repeated with a random set of 10 infants initially removed as the independent test set. To maintain the general population frequency distribution, each set of 10 consisted of two randomly selected ASD-positive and eight randomly selected ASD-negative children. This analysis was run with 1000 random sets, allowing us to assess the distribution of classifier accuracies when more participants were kept independent. This represents a very small sample of the full set of randomized permutations; however, this analysis is meant to serve only as a demonstration of the robustness of the classification analysis.

**Statistical analysis**

All classification analyses were completed using MATLAB’s **Statistics and Machine Learning Toolbox** (Mathworks Inc.). SVMs were trained using a linear kernel using the default setting of the fitcsvm function, and individual participants were predicted using the predict function. The default setting of this algorithm accounts for imbalances in the groups by setting the class prior probabilities to the relative frequencies of each class and then normalizes the weights to sum to the value of the prior probability in the respective class. Scripts were designed in-house, and their workflow is detailed above. Principal components were calculated using the default settings of the pca function to create a linear combination of the features space, which was then used for the visualization in Fig. 1. As described above, CIs for the reported proportions (sensitivity, specificity, positive predictive values, and negative predictive values) are calculated according to the efficient score method and corrected for continuity (56).

**SUPPLEMENTARY MATERIALS**

www.sciencetranslationalmedicine.org/cgi/content/full/9/393/eaag2882/DC1

Materials and Methods

Fig. S1. Individual classification accuracies.

Fig. S2. Null distribution of classification accuracy.

Table S1. The Talairach coordinates for each of the ROIs.

Table S2. Comparison to independent high-risk sample.

Reference (57)

**REFERENCES AND NOTES**


The Autism Diagnostic Observation Schedule


Predicting the future with brain imaging

In a new study, Emerson et al. show that brain function in infancy can be used to accurately predict which high-risk infants will later receive an autism diagnosis. Using machine learning techniques that identify patterns in the brain’s functional connections, Emerson and colleagues were able to predict with greater than 96% accuracy whether a 6-month-old infant would develop autism at 24 months of age. These findings must be replicated, but they represent an important step toward the early identification of individuals with autism before its characteristic symptoms develop.