

# 18-Month Predictors of Later Outcomes in Younger Siblings of Children With Autism Spectrum Disorder: A Baby Siblings Research Consortium Study

Katarzyna Chawarska, PhD, Frederick Shic, PhD, Suzanne Macari, PhD, Daniel J. Campbell, PhD, Jessica Brian, PhD, Rebecca Landa, PhD, Ted Hutman, PhD, Charles A. Nelson, PhD, Sally Ozonoff, PhD, Helen Tager-Flusberg, PhD, Gregory S. Young, PhD, Lonnie Zwaigenbaum, MD, Ira L. Cohen, PhD, Tony Charman, PhD, Daniel S. Messinger, PhD, Ami Klin, PhD, Scott Johnson, PhD, Susan Bryson, PhD

**Objective:** Younger siblings of children with autism spectrum disorder (ASD) are at high risk (HR) for developing ASD as well as features of the broader autism phenotype. Although this complicates early diagnostic considerations in this cohort, it also provides an opportunity to examine patterns of behavior associated specifically with ASD compared to other developmental outcomes. **Method:** We applied Classification and Regression Trees (CART) analysis to individual items of the Autism Diagnostic Observation Schedule (ADOS) in 719 HR siblings to identify behavioral features at 18 months that were predictive of diagnostic outcomes (ASD, atypical development, and typical development) at 36 months. **Results:** Three distinct combinations of features at 18 months were predictive of ASD outcome: poor eye contact combined with lack of communicative gestures and giving; poor eye contact combined with a lack of imaginative play; and lack of giving and presence of repetitive behaviors, but with intact eye contact. These 18-month behavioral profiles predicted ASD versus non-ASD status at 36 months with 82.7% accuracy in an initial test sample and 77.3% accuracy in a validation sample. Clinical features at age 3 years among children with ASD varied as a function of their 18-month symptom profiles. Children with ASD who were misclassified at 18 months were higher functioning, and their autism symptoms increased between 18 and 36 months. **Conclusion:** These findings suggest the presence of different developmental pathways to ASD in HR siblings. Understanding such pathways will provide clearer targets for neural and genetic research and identification of developmentally specific treatments for ASD. *J. Am. Acad. Child Adolesc. Psychiatry*, 2014;53(12):1317–1327. **Key Words:** ASD, high-risk siblings, infants, broader autism phenotype, predictors of outcomes

Autism spectrum disorder (ASD), characterized by impairments in communication and social reciprocity and the presence of repetitive behaviors and sensory interests,<sup>1</sup> is one of the most prevalent<sup>2</sup> and heritable<sup>3</sup> neurodevelopmental disorders. Compared to the general population, younger siblings of children with ASD are at increased risk for developing ASD

(18.7%)<sup>4</sup> or social-communicative and other developmental vulnerabilities.<sup>5-7</sup> Prospective studies of such high-risk (HR) cohorts provide unique opportunities to investigate the developmental dynamics underlying the emergence of autistic psychopathology as well as resilience.<sup>8,9</sup> A better understanding of these dynamics could lead to identification of early predictors of ASD in HR siblings, which in turn would enhance screening and diagnostic practices in this cohort and might inform identification of novel treatment targets. Considering the importance of early intervention in shaping brain development and



Supplemental material cited in this article is available online.

behavioral outcomes in young children with ASD,<sup>10</sup> identification of the affected siblings as soon as their symptoms begin to emerge offers the potential for altering their long-term outcomes in a clinically meaningful manner.

Early identification of HR siblings who are most likely to develop ASD is impeded by the complexity of clinical presentation and developmental trajectories unique to this cohort. A large minority of HR siblings without ASD exhibit broader autism phenotype features, e.g., elevated scores on diagnostic instruments of autism severity and lower verbal skills,<sup>5-7,11,12</sup> limited functional play,<sup>13</sup> imitation,<sup>14</sup> and repetitive behaviors,<sup>15</sup> and therefore it may be difficult to differentiate them in the second year of life from children who eventually develop ASD. Moreover, it is not clear when and how the behavioral symptoms begin to emerge. Symptoms may emerge after a period of relatively typical development either as a result of a loss of skills<sup>16</sup> or a failure to acquire new skills.<sup>17,18</sup> The conceptualization of onset patterns is further complicated by the finding that some symptoms may show a gradual decrease in frequency (e.g., eye contact), whereas others fail to increase at the rate observed in typical controls<sup>19</sup> (e.g., social vocalizations), suggesting that the departure from typical trajectories in specific domains may follow different patterns and take place during different developmental periods. The complexity of broader autism phenotype expression among HR siblings and variable symptom onset patterns may contribute to difficulties in identifying features early in development that are consistently associated with ASD among HR siblings.

In a recent study aimed at identifying predictive signs of ASD in HR siblings based on clinical presentation as captured by individual items of the Autism Diagnostic Observation Schedule–Toddler Module (ADOS-T)<sup>20</sup>, Macari *et al.*<sup>11</sup> used a nonparametric decision-tree learning algorithm (Classification and Regression Trees [CART]).<sup>21,22</sup> A combination of several features—including the limited ability to engage in play, paucity of communicative gestures, limited imitation skills, and atypical intonation—differentiated siblings with ASD from the remaining sample around the first birthday with a high degree of sensitivity and specificity. Although promising with regard to a methodological approach to identifying early prognostic features of ASD, the study was considered preliminary because of its small sample size ( $N = 84$ ) and focus on predicting

outcome at 24 months. Taking advantage of the large and prospectively characterized samples collected through the Baby Siblings Research Consortium (BSRC),<sup>7,19</sup> the present study is focused on identifying features in 18-month-old HR siblings predictive of ASD diagnosis at age 3 years. We focused on detection of behavioral features at 18 months, as this represents an age by which many parents begin to note concerns regarding their children's development but before most affected children receive diagnosis and enter treatment.<sup>23,24</sup> Risk assessment at 18 months was based on coding of 29 individual items from the ADOS Module 1<sup>25</sup>, with scores ranging from 0 to 3. The items provided a broad range of autism-relevant behaviors measured in a standardized fashion. To identify predictors of ASD, atypical (ATYP) and typical developmental (TD) diagnostic outcome at 3 years based on the ADOS items, we used a CART approach. The advantage of this approach is that it allows selection of the most predictive features from a multiplicity of behavioral symptoms and their interactions, resulting in a parsimonious mapping of different sets of predictors for later outcomes. It also allows several distinct combinations of features to be related to a single diagnostic outcome, a characteristic that is particularly important, given that ASD may arise through multiple etiological pathways.<sup>26</sup> We examined combinations of behavioral features at 18 months that are associated with ASD diagnosis at 3 years, as well as the factors that affect accurate identification at 18 months of HR siblings who receive diagnoses of ASD at 3 years of age.

## METHOD

### Study Participants

Participants were 719 HR siblings (413 males [57.4%] and 306 females [42.6%]) from 8 sites of the BSRC, a network of projects focused on studying the development of younger siblings of children with ASD: University of Toronto/Holland-Bloorview Kids Rehabilitation Hospital, Dalhousie/IWK Health Centre, University of Alberta/ Glenrose Rehabilitation Hospital, Boston University/Boston Children's Hospital, Kennedy Krieger Institute, University of California, Davis, University of California, Los Angeles, and Yale University. Infants were included in the study if they had an older biological sibling with ASD and had complete characterization data including ADOS and MSEL<sup>27</sup> at the ages of 18 months and 3 years. All sites verified the older siblings' diagnostic status based on the ADOS,<sup>25</sup> Autism Diagnostic Interview–Revised,<sup>28</sup>

and/or Social Communication Questionnaire.<sup>29</sup> The exclusion criterion was an identified neurological or genetic condition in the older or younger sibling (e.g., fragile X syndrome or tuberous sclerosis).

### Measures

**ADOS.** The ADOS is a semi-structured, standardized assessment of social interaction, communication, and play skills as well as repetitive behaviors that are diagnostic of ASD.<sup>25</sup> To quantify symptoms of ASD, at 18 months, all children were administered module 1, which is designed to assess behavioral symptoms in nonverbal and minimally verbal young children. At 3 years, either module 1 (n = 181) or module 2 (for children with early phrase speech; n = 538) was used. To facilitate comparisons of symptom severity across modules and ages, total algorithm scores were converted into calibrated severity scores,<sup>30</sup> which range from 1 to 10.

**MSEL.** The MSEL is a standardized developmental measure for children between birth and 68 months.<sup>27</sup> Scores in 4 MSEL domains were considered in the current analysis: Fine Motor, Visual Reception, Expressive Language, and Receptive Language. Nonverbal Developmental Quotient (DQ; NVDQ) scores were computed as the average of Fine Motor and Visual Reception DQ scores, and verbal DQ (VDQ) scores as the average of Expressive Language and Receptive Language scores.

### Outcome Assessment

Diagnostic groupings were based at 3 years on a combination of clinical best-estimate (CBE) and test scores on ADOS and MSEL (detailed in Table S1, available online). CBE was assigned at each site by an expert clinician based on a combination of measures including medical and developmental history as well as social, cognitive, verbal, and adaptive functioning. Consistent with other BSRC reports,<sup>4</sup> children in the ASD group had to meet CBE criteria for ASD and to have an ADOS severity score in the clinical range. Toddlers with ATYP exhibited abnormal scores on either the ADOS or MSEL or both. The TD group had scores in the typical range on both instruments. Based on these criteria, the sample consisted of 157 siblings (21.8%) with ASD outcomes, 178 siblings (24.8%) with ATYP outcomes, and 384 siblings (53.4%) with TD outcomes.

### Statistical Analysis

We used CART analysis<sup>21,22</sup> to identify the individual items of the ADOS at 18 months of age that best predicted diagnostic outcomes at 3 years of age. CART analysis is a decision-tree technique that uses recursive partitions of the data to predict a categorical or continuous response variable. A decision tree is a flow-chart-like structure in which each internal (non-leaf) node denotes a test on an attribute (e.g., ADOS item X),

each branch represents the outcome of a test (e.g., the score of item X), and each leaf represents a class label (e.g., ASD, ATYP, or TD). At each step, the model selects the best variable and cutoff score among all available predictor variables to make a partition. The nested structure of partitions within CART analysis naturally incorporates interactions among variables in the model, and the option to stop the growth of the tree at any partition (i.e., “pruning” the tree) provides a method of variable selection by predictive importance.

A validation sample comprising 20% of the participants (n = 154), selected randomly from each outcome group (35 ASD, 40 ATYP, and 79 TD) and site, was initially set aside and used only for assessment of the final CART model on out-of-sample data. The remaining participants (n = 565; 122 ASD, 138 ATYP, and 305 TD) were retained in the training set and used to construct the tree. CART models were built using the “tree” package<sup>31</sup> in R<sup>32</sup>. To prevent overfitting the model to the training data, an optimal tree size was identified by examining misclassifications over all 3 groups using 10-fold cross-validation on the training set for tree sizes varying from 2 to 20 leaves. The final tree was pruned to this optimal misclassification-minimizing size, with the fitted label for each leaf in the final tree assigned by a “majority vote” (i.e., defined as the most prevalent diagnostic group in the leaf, with ties broken by giving higher precedence to groups having a higher proportion of their membership within that leaf).

Comparisons of cases that were correctly classified and misclassified by CART were conducted using linear mixed models with Tukey–Kramer correction for multiple post hoc comparisons. Effects of site and gender were included in the models as indicated.

## RESULTS

**Sample description.** ASD, ATYP, and TD groups were similar with regard to racial (white versus nonwhite) and ethnic (Hispanic versus non-Hispanic) composition (Table 1). Boys were more likely to be diagnosed with ASD than girls (29.3% versus 11.8%) and to have an ATYP outcome (28.1% versus 20.3%). The groups did not differ in the age of recruitment, or in maternal and paternal age. They included similar proportions of college-educated mothers, but fathers of children with TD outcomes were more likely to hold a college degree than fathers of children with ASD or ATYP. There were no differences among groups in age at the 18-month visit, but the groups differed in autism severity scores (ASD>ATYP>TD) and verbal and nonverbal (ASD<ATYP=TYP) scores. There were no significant differences among sites in the proportion of ASD outcomes ( $\chi^2[5] = 5.39$ ,  $p = .369$ ), ADOS-calibrated severity score, verbal

**TABLE 1** Sample Characterization Based on 36-Month Outcomes

Variable	ASD		ATYP		TD		p Value
n (% of total sample)	157	(21.8)	178	(24.8)	384	(53.4)	
Age of recruitment, mo	8.2	(3.8)	7.7	(3.7)	7.5	(4.4)	$p = .25$
Males (%)	121	(77.1) <sup>a</sup>	116	(65.2) <sup>a</sup>	176	(46.0) <sup>b</sup>	$p < .001$
Relative male to female risk ratio	2.5:1		1.5:1		0.63:1		
White, %	81.3		83.7		84.7		$p = .66$
Hispanic, %	13.7		3.8		8.1		$p = .09$
Maternal age, y	34.0	(4.6)	35.2	(4.6)	34.4	(4.4)	$p = .06$
Paternal age, y	36.3	(5.5)	37.3	(5.3)	36.9	(5.2)	$p = .26$
Mothers with college degree or higher, %	71.0		72.9		78.4		$p = .22$
Fathers with college degree or higher, %	68.0 <sup>a</sup>		63.2 <sup>a</sup>		75.7 <sup>b</sup>		$p = .02$
<b>18 Months</b>							
Age, mo	18.4	(0.6)	18.4	(0.6)	18.3	(0.5)	$p = .21$
ADOS Severity Score	4.6	(2.8) <sup>a</sup>	2.8	(1.9) <sup>b</sup>	2.1	(1.5) <sup>c</sup>	$p < .001$
MSEL Verbal DQ	73.5	(23.5) <sup>a</sup>	91.4	(22.5) <sup>b</sup>	97.2	(20.6) <sup>b</sup>	$p < .001$
MSEL Nonverbal DQ	95.5	(12.3) <sup>a</sup>	103.8	(12.5) <sup>b</sup>	105.9	(11.3) <sup>b</sup>	$p < .001$
<b>36 Months</b>							
Age, mo	37.5	(2.4) <sup>a</sup>	38.3	(2.8) <sup>b</sup>	37.7	(2.8) <sup>ab</sup>	$p = .02$
ADOS Severity Score	7.0	(1.7) <sup>a</sup>	4.1	(1.8) <sup>b</sup>	1.3	(0.5) <sup>c</sup>	$p < .001$
MSEL Verbal DQ	81.3	(27.0) <sup>a</sup>	98.1	(16.4) <sup>b</sup>	105.9	(14.6) <sup>c</sup>	$p < .001$
MSEL Nonverbal DQ	86.4	(21.8) <sup>a</sup>	100.7	(16.5) <sup>b</sup>	109.9	(13.7) <sup>c</sup>	$p < .001$

Note: Means are reported with standard deviations in parentheses: mean (SD), except where variable is labeled as a percentage (%). ADOS = Autism Diagnostic Observation Schedule; ASD = autism spectrum disorder; ATYP = atypical development; DQ = Developmental Quotient; MSEL = Mullen Scales of Early Learning; TD = typical development.

<sup>a,b,c</sup> Within each row, means with different superscript letters differ significantly, at least at  $p < .05$ , based on post hoc differences corrected for multiple tests using Tukey honestly significant difference test for continuous variables and Bonferroni correction for categorical variables.

DQ, and nonverbal DQ at either of the ages (all  $p$  values  $> .07$ ).

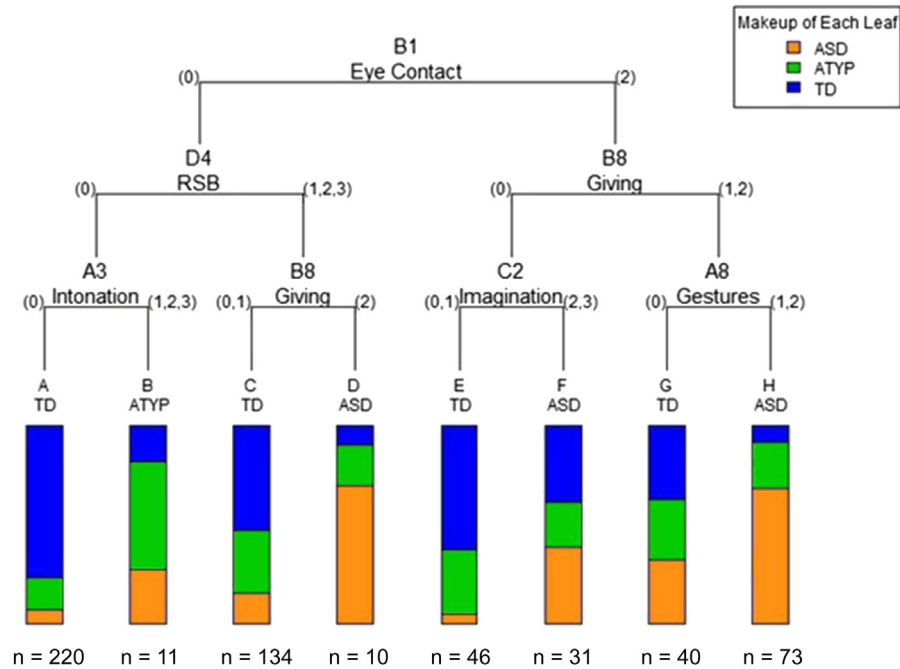
*Predicting ASD diagnosis based on behavioral features at 18 months.* CART cross-validation analysis indicated that the optimal misclassification error rate was obtained for a tree size of 8 leaves (Figure 1, Table 2). Three leaves received a label of "ASD," capturing 57% of ASD cases (H, F, and D). A majority (93%) of TD siblings was classified correctly into leaves A, C, E, and G; thus these leaves received a label of "TD." Very few (4%) children with ATYP outcomes were assigned their own leaf (B); instead, the ATYP siblings were distributed largely between TD (77%) and ASD (19%) leaves. Inclusion of site in the model did not alter the tree structure, suggesting that site was not an important predictor of classification outcome under CART analysis. Taken together, at 18 months, CART analysis predicted ASD versus non-ASD (ATYP and TD combined) diagnostic status at 36 months with an overall accuracy of 82.7%, sensitivity of 56.6%, specificity of 89.8%, positive predictive value (PPV) of 60.5%, and negative predictive value (NPV) of 88.2%. Analogous prediction rates on the validation subsample were overall accuracy

of 77.3%, sensitivity of 45.7%, specificity of 86.6%, PPV of 50.0%, and NPV of 84.4%.

The analysis identified 3 combinations of behavioral features predictive of diagnostic outcome at 36 months (Figure 1, Table 2). The largest proportion of affected siblings (41%, leaf H) was classified based on a combination of poorly modulated eye contact to initiate, terminate, or regulate social interactions (item B1 of Module 1, score 2), the absence of giving objects to others to share (B8, score 1 or 2), and limited use of emotional or descriptive communicative gestures other than pointing (A8, score 1 or 2). Relative risk for ASD in this leaf was 3.17; that is, compared to all siblings in our sample, siblings in leaf H were more than 3 times more likely to be diagnosed with ASD. Toddlers in this leaf had ADOS-calibrated severity scores that were high and stable from 18 to 36 months, verbal scores that were very low, and nonverbal scores that were low to average (Table 3). An additional 10% of the children with ASD were classified by their impaired eye contact (B1, score 2), and the lack of ability to spontaneously engage in pretend play (C2) (leaf F, score 2 or 3). The relative risk for ASD in this leaf was 1.79, suggesting an almost 2-fold



**FIGURE 1** Classification and Regression Tree (CART) diagram predicting groups with autism spectrum disorder (ASD), atypical development (ATYP), and typical development (TD) based on 18-month Autism Diagnostic Observation Schedule (ADOS) item data. Note: Bars reflect the proportion of toddlers from each diagnostic group assigned into each leaf based on their presentation during ADOS evaluation. ADOS items used in the tree are B1 (Unusual Eye Contact), D4 (Unusually Repetitive Interests or Stereotyped Behaviors [RSB]), A3 (Intonation of Vocalizations or Verbalizations), B8 (Giving), C2 (Imagination/Creativity), and A8 (Gestures). Numbers displayed to the left and right of the item label reflect the actual scores on the given item of the ADOS.



increased likelihood of ASD in this leaf. Toddlers in leaf F had moderate calibrated symptom severity scores that remained stable over time and displayed below-average verbal and average nonverbal skills (Table 3). Finally, a third group of siblings with ASD classified into leaf D (6%) displayed appropriate eye contact meshed with other communicative behaviors (B1, score 0), but presented with repetitive and stereotyped behaviors ranging from relatively mild to severe (D4, score 1, 2, or 3) and rarely or never gave objects to get help or to share (B8, score 2) during the ADOS assessment. Relative risk for ASD in this leaf was 3.24. Children in this leaf had calibrated severity scores at 18 months in the borderline range; however, their symptoms worsened over time (Table 3). They also exhibited below-average verbal skills and average non-verbal skills at both time points.

*Comparison of HR siblings with ASD classified correctly versus misclassified.* Subsequently, we compared siblings with ASD outcomes who were at 18 months “Classified Correctly” (ASD: n = 69) or “Misclassified” (i.e., classified as ATYP or TD;

n = 53). There were no significant differences between correctly classified and misclassified participants with ASD in regard to gender, age at recruitment or assessment, race or ethnicity, and maternal or paternal age and education (Table S2, available online). However, there was a significant effect of age ( $F_{1,119} = 112.7, p < .001$ ), classification group ( $F_{1,113} = 69.7, p < .001$ ), and age  $\times$  classification interaction ( $F_{1,119} = 24.96, p < .001$ ) on autism severity score (Figure 2). The misclassified children had significantly lower severity scores than those classified correctly at both time points, and severity scores increased over time in both groups (all post hoc comparisons are significant at least at  $p < .01$  level with Tukey–Kramer correction). There was no significant effect of site ( $p = .28$ ) or gender ( $p = .23$ ). For verbal DQ there were significant effects of age ( $F_{1,77} = 9.84, p = .003$ ) and classification ( $F_{1,113} = 32.49, p < .001$ ), but no interaction effect ( $p = .10$ ). The misclassified siblings had higher VDQs at both time points than those classified correctly, and in both groups the VDQ scores increased over time. There was no effect of site ( $p = .15$ ) or

**TABLE 2** Composition of the Subgroups Based on Classification and Regression Tree (CART) Analysis at 18 Months

ADOS-G Items Predictive of ASD	Leaf	Leaf Label	ASD <sup>a</sup>		ATYP <sup>a</sup>		TD <sup>a</sup>		Total Leaf Size	Relative Risk for ASD (95% CI)		
			n	%	n	%	n	%				
B1: Appropriate eye contact	A	TD	15	(12)	7	[16]	169	(55)	220	[100]	0.32 (0.19–0.53)	
	A3: Normal intonation										1.26 (0.48–3.36)	
	B	ATYP	3	(2)	[27]	6	(4)	[55]	2	(1)	[18]	0.69 (0.45–1.07)
	A3: Somewhat odd intonation										3.24 (2.10–5.01)	
B4: Some RSB	C	TD	20	(16)	[15]	43	(31)	[32]	71	(23)	[53]	0.20 (0.05–0.79)
	B8: Good or limited giving										1.79 (1.12–2.87)	
B1: Poor eye contact	D	ASD	7	(6)	[70]	2	(1)	[20]	1	(0)	[10]	1.51 (0.94–2.42)
	B8: No giving										3.17 (2.54–3.96)	
B8: Good giving	E	TD	2	(2)	[4]	15	(11)	[33]	29	(10)	[63]	1.00 (0.96–1.04)
	C2: Spontaneous pretend play											
B8: Limited Giving	F	ASD	12	(10)	[39]	7	(5)	[23]	12	(4)	[39]	
	C2: No spontaneous pretend play											
B8: Limited Giving	G	TD	13	(11)	[33]	12	(9)	[30]	15	(5)	[38]	
	A8: Variety of gestures											
B8: Limited Giving	H	ASD	50	(41)	[69]	17	(12)	[23]	6	(2)	[8]	
	A8: Limited use of gestures											
Total outcome group			122	(100)		138	(100)		305	(100)		565

Note: n = 565; 20% of original sample removed for the validation set. Relative risk estimates for autism spectrum disorder (ASD) are based on prevalence of ASD among all highrisk (HR) siblings in this sample (risk compared to the general population is 14.7 times greater). ADOS = Autism Diagnostic Observation Schedule; ATYP = atypical development; RSB = repetitive and stereotyped behaviors; TD = typical development.  
<sup>a</sup>Column marginals (proportion of all children of a particular outcome at 36 months, as given by the column header, found in each leaf) are given in parentheses, whereas row marginals (proportion of children within that leaf with an ultimate diagnosis given by the column header) are given in square brackets.

**TABLE 3** Characterization of Subgroups of High-Risk Siblings Based on Classification and Regression Tree (CART) Analysis at 18 Months

Leaf (Label)	n	% Boys	Age of recruitment (mo)	Age of assessment (mo)	Severity 18 mo	Severity 36 mo	VDQ 18 mo	VDQ 36 mo	NVDQ 18 mo	NVDQ 36 mo
A (TD)	220	44.1	7.4 (4.6)	18.3 (6)	<b>1.5 (.9)</b>	<b>2.1 (1.7)***</b>	100 (28)	104 (16)	105 (11)	109 (16)
B (ATYP)	11	72.7	7.2 (2.3)	18.2 (4)	<b>1.6 (1.2)</b>	<b>4.6 (1.9)**</b>	108 (7)	106 (15)	117 (12)	105 (13)
C (TD)	134	63.9	7.5 (3.7)	18.3 (5)	<b>1.9 (1.1)</b>	<b>2.9 (2.4)***</b>	97 (20)	101 (17)	105 (12)	103 (16)
D (ASD)	10	90.0	9.2 (4.7)	18.2 (6)	<b>3.2 (1.5)</b>	<b>6.4 (2.8)**</b>	83 (25)	84 (28)	93 (19)	90 (29)
E (TD)	46	58.7	7.8 (3.4)	18.4 (5)	3.7 (1.4)	2.7 (2.2)	<b>85 (17)</b>	<b>104 (14)***</b>	102 (8)	105 (16)
F (ASD)	31	71.0	7.4 (3.2)	18.5 (6)	4.8 (1.7)	4.1 (2.7)	85 (24)	93 (28)	102 (10)	97 (19)
G (TD)	40	57.9	7.9 (3.3)	18.4 (5)	4.6 (1.3)	4.1 (2.4)	<b>84 (16)</b>	<b>99 (19)**</b>	105 (12)	98 (17)
H (ASD)	73	74.0	8.1 (3.8)	18.5 (7)	6.1 (1.9)	6.1 (2.6)	<b>62 (21)</b>	<b>77 (26)**</b>	92 (13)	84 (21)

Note: Post hoc comparisons between scores at 18 and 36 months with Bonferroni correction for multiple comparisons. Variables associated with significantly different comparisons are boldfaced within each row. Excluding n and % Boys, variables are reported as means and standard deviations: mean (SD). ASD = autism spectrum disorder; ATYP = atypical development; NVDQ = nonverbal developmental quotient; TD = typical development; VDQ = verbal developmental quotient.  
 \*\*\*p < .0001; \*\*p < .001.

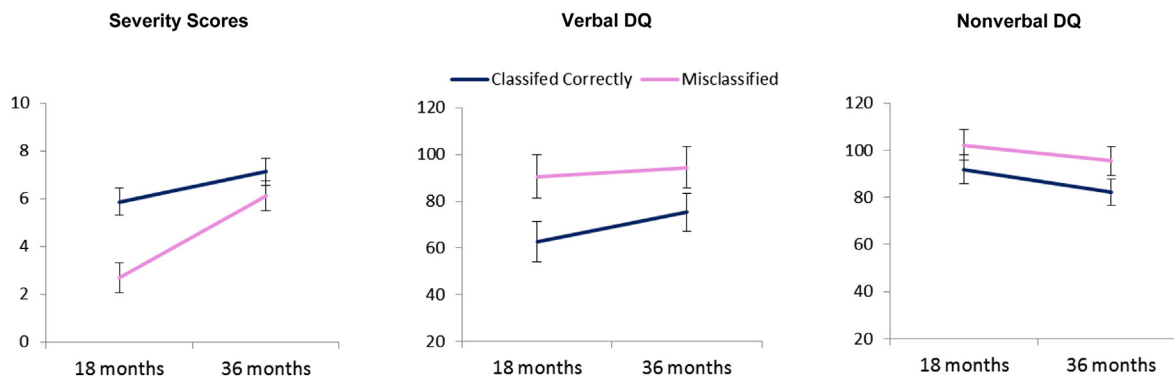
gender ( $p = .94$ ) on VDQ. Finally, for NVDQ, there was a significant effect of age ( $F_{1,78} = 11.91, p < .001$ ) and classification ( $F_{1,113} = 19.22, p < .001$ ), but no interaction effect ( $p = .55$ ). Misclassified children had higher NVDQ scores than those who were classified correctly, and in both groups, the scores declined slightly over time. There was no site effect ( $p = .17$ ) or gender effect ( $p = .54$ ) for NVDQ.

### DISCUSSION

This is the first large-scale, multi-site study aimed at identifying specific behavioral features that distinguish HR siblings with ASD from their typically and atypically developing high-risk peers as early as 18 months of age. Almost half of the HR siblings exhibited clinically relevant concerns at 36 months in the form of ASD (21.8%) or other atypical developmental outcomes (24.8%). Compared to recent population prevalence rates of ASD,<sup>2</sup> the HR siblings in our study were 14.7 times more likely to develop ASD. Boys were 2.5 times more likely to be affected by ASD,<sup>4,33</sup> and 1.5 times more likely to have ATYP outcomes than girls.<sup>6,7</sup> Despite these differences in gender distribution with regard to outcome, boys and girls were equally likely to be classified correctly by the CART analysis, suggesting that this analytic approach did not favor detection of signs of ASD in one gender over another.

To identify indices of diagnostic outcome at the age of 3 years, we applied a nonparametric decision-tree learning algorithm to an array of behavioral ratings collected during a standardized social interaction at 18 months. By considering combinations of only 6 behavioral features at 18 months, it was possible to identify ASD cases with approximately 83% overall accuracy, 57% sensitivity, and 90% specificity in the training sample. This is a relatively high level of accuracy, given that a large proportion of HR siblings showed broader phenotype features already at 18 months, and that siblings missed by the classification procedure were higher functioning with an increase in symptom severity after 18 months and therefore were unlikely to trigger clinical concerns at this age. Classification based on social-communicative features and other autism-specific behaviors at 18 months was not particularly successful at isolating siblings with non-ASD atypical outcomes. Almost 20% of this group showed characteristics seen in the ASD leaves, whereas more than 70% were classified

**FIGURE 2** Marginal means ( $\pm 2$  standard errors) of severity scores, verbal developmental quotient (DQ) scores, and nonverbal DQ scores in children with autism spectrum disorder (ASD) classified correctly as ASD and misclassified as typically or atypically developing at 18 months.



together with TD siblings at 18 months. Follow-up work will be necessary to better understand the distribution of ASD-related traits at 18 months in siblings without ASD and their association with outcomes.

Although the search for invariant biological markers of complex developmental disorders including ASD is ongoing, diagnosis of ASD still relies on the analysis of the behavioral phenotype, a phenotype that is not only highly heterogeneous, but also undergoes important transformations in the first years of life.<sup>34,35</sup> These factors motivated our focus on a very narrow developmental epoch along with the examination of a high number of combinations of behaviors known to be impaired in the early stages of ASD as potential prognostic indicators. The strength of this analytic approach lies in its ability to enhance the predictive value of individual behaviors by taking into consideration how they interact with other behaviors. Specifically, at 18 months, atypical eye contact was noted in 34% of siblings. Yet, only 40% of them had ASD by 3 years, suggesting that, at this age, atypical eye contact by itself is a poor prognostic indicator of ASD among HR siblings. However, if at 18 months the atypical eye contact co-occurred with a paucity of communicative gestures and limited use of giving objects to share, then the likelihood of ASD outcome was more than 3 times greater than in HR siblings in general, and almost all children who displayed these characteristics had clinically significant (ASD or ATYP) outcomes at 3 years. In a small proportion of siblings with ASD, eye contact was not impaired at 18 months, but these siblings displayed emerging repetitive behaviors and limited use of giving objects to others for any

purpose. Children with this combination of features were over 3 times more likely to have ASD than the HR siblings in general, and almost all siblings in this group had ASD or other atypical outcomes at 3 years. Finally, poor eye contact paired with a limited ability to spontaneously engage in pretend play, although not very specific to ASD (relative risk: 1.79), appeared to more generally signal the presence of developmental challenges, as almost two-thirds of siblings in this group had outcomes of either ASD or ATYP at 3 years of age. Importantly, the groups of siblings identified by the 3 combinations of features displayed distinct developmental paths between 18 and 36 months with regard to autism severity, verbal, and nonverbal skills. Taken together, these findings suggest that different constellations of individual behaviors observed at 18 months are associated with a common diagnostic outcome. Although all HR siblings should be monitored throughout the first 3 years of life, the presence of these combinations of characteristics at 18 months—the first characterized by drastically limited nonverbal communication and the second by the presence of repetitive behaviors—should trigger consideration of a comprehensive diagnostic assessment and, if indicated, targeted intervention.

One of the key challenges in the field of autism research is the identification of more phenotypically homogeneous subgroups among individuals affected by the disorder. Identification of such subgroups may facilitate gene finding and discovery of novel treatment targets and strategies. Among the otherwise highly heterogeneous sample of HR siblings with ASD, our analysis identified a large (41%) subgroup of



infants who were highly symptomatic and exhibited marked language delays by 18 months. Importantly, they continued to have pronounced autism symptoms and significant language delays at the age of 3 years. An earlier study identified a group of siblings with similar characteristics around the first birthday using an analogous analytic approach, albeit in a much smaller sample.<sup>11</sup> Taken together, these findings suggest that, in a large minority of HR siblings, clinical symptoms of ASD paired with significant language delays may emerge around 18 months or before and remain relatively stable over time. Although a more long-term follow up would be necessary, this group may represent toddlers at risk for less optimal outcomes among children affected by ASD.

Almost one-half of the siblings later diagnosed with ASD did not trigger an ASD classification at 18 months based on the CART analysis. Neither gender distribution nor parental characteristics differed between the correctly classified and misclassified ASD cases. Importantly, the misclassified siblings appeared to have few developmental delays at 18 months, and their behavioral response to the ADOS probes were largely in the nonclinical range. However, while they maintained average verbal and nonverbal skills over the subsequent months, they evidenced a rapid increase in autism symptom severity between 18 and 36 months. Several, not necessarily mutually exclusive, hypotheses can be advanced to explain this finding. First, it is plausible that the siblings were either asymptomatic or mildly symptomatic at 18 months, and in the subsequent months withdrew socially and either failed to advance or showed a loss of previously acquired social-communicative skills. Second, it is possible that their higher verbal and nonverbal skills masked their social disability in the context of the playful interaction with a highly supportive adult, but as their speech and representational skills advanced, their atypical interactive and communicative behaviors became apparent (e.g., stereotyped speech, all-encompassing interest, and limited conversational skills). Finally, it is possible that we missed these cases because we focused our investigation too narrowly on autism-specific behaviors elicited in a particular context. Considering additional information derived from other social contexts (e.g., peer interactions), developmental domains (e.g., emotion regulation) or levels of analysis (e.g., social perception and attention) may help

to identify those higher-functioning siblings with ASD who were missed at 18 months. Direct investigation of these hypotheses will markedly enhance our understanding of both methodological and theoretical issues surrounding the developmental dynamics of HR siblings with ASD.

It is important to note that 18 months is a developmental time point marked by important transitions in the emergence of representational thought, language, and social interactions. Given the developmental nature of ASD, it is possible that the predictive signs of ASD identified at this time point may not generalize to other developmental epochs. The identification of combinations of features specific to, for example, 6-, 12-, 18-, and 24-month age levels will be crucial for furthering our understanding of processing underlying emergence of social disability and for the advancement of developmentally sensitive ASD screening instruments for high-risk siblings in the general population. Our analytic approach was designed to identify diagnostic features that differentiate toddlers with ASD who have an older sibling on the autism spectrum from HR siblings who do not develop the disorder but may display broader phenotype features. It remains an empirical question whether these findings generalize to a broader population of affected toddlers. Finally, although we have focused our analysis on minimizing misclassification error, alternative analytic strategies could be used if the purpose of the analysis was to develop screening instruments in which greater emphasis is placed on maximizing sensitivity over specificity.<sup>36</sup> The exploration of these variations and techniques, as well as the incorporation of additional data sources in classification and prediction, may result in advancing the development of early screening and diagnostic methods.

This study has several important clinical and theoretical implications. A large minority of HR siblings with ASD display marked symptoms at 18 months, whereas in others, symptoms become pronounced after 18 months, suggesting at least 2 distinct windows of opportunity for identification of the affected cases. Moreover, several combinations of clinical features at 18 months were predictive of ASD outcome, each associated with a different developmental course and clinical profile by the age of 3 years. The clinical implications of these results are amplified as the combinations of predictive features were derived from the ADOS, a well-validated, standardized,

and therefore readily replicable assessment of ASD symptoms. Combined, these findings suggest the presence of different developmental pathways to the common diagnostic ASD outcome, pathways characterized by distinct combinations of early markers. The findings also reinforce the need for repeated diagnostic screening in the first 3 years of life to identify individual cases of ASD as soon as behavioral symptoms become apparent. Such rapid detection will enhance our ability to ameliorate the impact of primary symptoms on the development of social and nonsocial cognition, as well as to prevent secondary symptoms from emerging. Better understanding of the neurobiology of the various underlying pathways to ASD may advance both the identification of novel treatment targets and the design of interventions tailored to specific clinical profiles and their developmental dynamics. &

Accepted October 1, 2014.

This article was reviewed under and accepted by associate editor James J. Hudziak, MD.

Drs. Chawarska, Shic, and Macari are with Yale University School of Medicine, New Haven, CT. Dr. Campbell is with Amgen Inc., Thousand Oaks, CA. Dr. Brian is with University of Toronto. Dr. Landa is with Kennedy Krieger Institute and Johns Hopkins School of Medicine, Baltimore. Drs. Hutman and Johnson are with University of California, Los Angeles. Dr. Nelson is with Harvard Medical School and Boston Children's Hospital. Drs. Ozonoff and Young are with the Medical Investigation of Neurodevelopmental Disorders (MIND) Institute at the University of California, Davis. Dr. Tager-Flusberg is with Boston University. Dr. Zwaigenbaum is with the University of Alberta, Edmonton, Alberta, Canada. Dr. Cohen is with New York State Institute for Basic Research in Developmental Disabilities, Albany, NY. Dr. Charman is with the Institute of Psychiatry, King's College London. Dr. Messinger is with the University of Miami. Dr. Klin is with the Marcus Autism Center,

Children's Healthcare of Atlanta, and Emory University, Atlanta. Dr. Bryson is with Dalhousie University, Halifax, Canada.

Autism Speaks (AS) provided funding for the creation of the Baby Siblings Research Consortium (BSRC) database. Data collection, analyses, and manuscript preparation were supported by the National Institutes of Health (NIH) grants P01 HD003008, Project 1 (K.C.), R01MH087554 (Principal Investigator [PI]: K.C.), CTSA UL1 RRO24139 (F.S.), MH096961 (T.H.), MH068172 (T.H.), HD055784 (T.H. and S.J.), MH068398 (PI: S.O.), R01DC010290 (Multiple Principal Investigators [MPI]: H.T.F. and C.A.N.), R21DC008637 (PI: H.T.F.), R01HD057284 and R01GM105004 (D.S.M.), P50MH100029 (A.K.), R01MH083727 (A.K.), and R01MH059630 (R.L.), UK Medical Research Council (T.C.), European Science Foundation Cooperation in Science and Technology (COST) Action BM1004 (ESSEA) under FP7 (T.C.), FAR Fund (MPI: K.C., F.S., Fred Volkmar, MD), Canadian Institutes of Health Research (CIHR) (J.B., S.B., L.Z.), NeuroDevNet (J.B., S.B., L.Z.), Simons Foundation (PI: C.A.N.), AS Pilot Grant (PI: H.T.F.), AS Postdoctoral Fellowship (D.J.C.), and AS Canada (J.B., S.B., L.Z.).

Drs. Chawarska, Campbell, and Shic served as the statistical experts for this research.

The authors thank Alycia Halladay, PhD, of Autism Speaks, for organizational support, and Elizabeth S. Kim, PhD, of Yale University, for editorial assistance.

Disclosure: Dr. Shic has received research support from Roche and Janssen. Dr. Campbell is employed by and receives salary and stock options from Amgen Inc. Dr. Zwaigenbaum is the site PI of a study sponsored by SynapDx (receives operating funds but no honoraria) and has received grant or research support from Alberta Innovates – Health Solutions, Stollery Children's Hospital Foundation, and Autism Intervention Research – Physical Health. Dr. Charman has received grant or research support from the National Institute of Health Research, the European Union, Autistica, Research Autism, the Autism Education Trust, Autism Speaks, and the Waterloo Foundation. He has received royalties from Guilford Press and Sage Publications. Drs. Chawarska, Macari, Brian, Landa, Hutman, Nelson, Ozonoff, Tager-Flusberg, Young, Cohen, Messinger, Klin, Johnson, and Bryson report no biomedical financial interests or potential conflicts of interest.

Correspondence to Katarzyna Chawarska, PhD, Yale Child Study Center, 40 Temple, Suite 7D, New Haven, CT 06510; e-mail: Katarzyna.chawarska@yale.edu

0890-8567/\$36.00/©2014 American Academy of Child and Adolescent Psychiatry

<http://dx.doi.org/10.1016/j.jaac.2014.09.015>

## REFERENCES

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders: 5th ed. DSM 5. Washington, DC: American Psychiatric Association; 2013.
2. Baio J. Prevalence of Autism spectrum disorders: autism and developmental disabilities monitoring network, 14 sites, United States, 2008. *Morb Mortal Wkly Rep. Surveillance Summaries (Centers for Disease Control and Prevention)*. 2012;61:1-19.
3. Geschwind DH. Genetics of autism spectrum disorders. *Trends Cogn Sci*. 2011;15:409-416.
4. Ozonoff S, Young GS, Carter A, *et al*. Recurrence risk for autism spectrum disorders: a Baby Siblings Research Consortium Study. *Pediatrics*. 2011;128:e488-e495.
5. Georgiades S, Szatmari P, Zwaigenbaum L, *et al*. A prospective study of autistic-like traits in unaffected siblings of probands with autism spectrum disorder: autistic traits in siblings of probands with ASD. *JAMA Psychiatry*. 2013;70:42-48.
6. Ozonoff S, Young GS, Belding A, *et al*. The broader autism phenotype in infancy: when does it emerge? *J Am Acad Child Adolesc Psychiatry*. 2014;53:398-407.
7. Messinger D, Young GS, Ozonoff S, *et al*. Beyond autism: a Baby Siblings Research Consortium Study of high-risk children at three years of age. *J Am Acad Child Adolesc Psychiatry*. 2013;52:300-308.
8. Chawarska K, Macari S, Volkmar F, Kim SH, Shic F. ASD in infants and toddlers. In: Volkmar FR, Paul R, Rogers SJ, Pelphrey KA, eds. *Handbook of Autism and Pervasive Developmental Disorders*. Hoboken, NJ: Wiley; 2014.
9. Zwaigenbaum L, Thurm A, Stone W, *et al*. Studying the emergence of autism spectrum disorders in high-risk infants: methodological and practical issues. *J Autism Dev Disord*. 2007;37:466-480.
10. Rogers SJ, Vismara L. Interventions for infants and toddlers at risk for autism spectrum disorder. *Handbook of Autism and Pervasive Developmental Disorders, Fourth Edition*. Hoboken, NJ: Wiley; 2014.
11. Macari S, Campbell D, Gengoux G, Saulnier C, Klin A, Chawarska K. Predicting developmental status from 12 to 24 months in infants at risk for autism spectrum disorder: a preliminary report. *J Autism Dev Disord*. 2012;42:2636-2647.
12. Landa RJ, Holman KC, Garrett-Mayer E. Social and communication development in toddlers with early and later diagnosis of autism spectrum disorders. *Arch Gen Psychiatry*. 2007;64:853-864.
13. Christensen L, Hutman T, Rozga A, *et al*. Play and developmental outcomes in infant siblings of children with autism. *J Autism Dev Disord*. 2010;40:946-957.
14. Young GS, Rogers SJ, Hutman T, Rozga A, Sigman M, Ozonoff S. Imitation from 12 to 24 months in autism and typical

- development: a longitudinal Rasch analysis. *Dev Psychol.* 2011; 47:1565.
15. Damiano CR, Nahmias A, Hogan-Brown AL, Stone WL. What do repetitive and stereotyped movements mean for infant siblings of children with autism spectrum disorders? *J Autism Dev Disord.* 2013;43:1326-1335.
  16. Eisenberg L, Kanner L. Early infantile autism, 1943-55. *Am J Orthopsychiatry.* 1956;26:556-566.
  17. Siperstein R, Volkmar F. Brief report: parental reporting of regression in children with pervasive developmental disorders. *J Autism Dev Disord.* 2004;34:731-734.
  18. Ozonoff S, Iosif A, Young GS, *et al.* Onset patterns in autism: correspondence between home video and parent report. *J Am Acad Child Adolesc Psychiatry.* 2011;50:796-806.
  19. Ozonoff S, Iosif A, Baguio F, *et al.* A prospective study of the emergence of early behavioral signs of autism. *J Am Acad Child Adolesc Psychiatry.* 2010;49:258-268.
  20. Lord C, Luyster R, Gotham K, Guthrie W. Autism Diagnostic Observation Schedule, (ADOS-2) Manual (Part II): Toddler Module. Los Angeles: Western Psychological Services; 2012.
  21. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* New York: Springer; 2009.
  22. Breiman L, Friedman J, Olshen R, Stone C. *Classification and Regression Trees.* New York: Wadsworth; 1984.
  23. Chawarska K, Paul R, Klin A, Hannigen S, Dichtel L, Volkmar F. Parental recognition of developmental problems in toddlers with autism spectrum disorders. *J Autism Dev Disord.* 2007;37:62-72.
  24. Ozonoff S, Young GS, Steinfeld MB, *et al.* How early do parent concerns predict later autism diagnosis? *J Dev Behav Pediatr.* 2009;30:367-375.
  25. Lord C, Rutter M, DiLavore PC, Risi S. *Autism Diagnostic Observation Schedule-Generic.* Los Angeles, CA: Western Psychological Services; 2000.
  26. Happe F, Ronald A, Plomin R. Time to give up on a single explanation for autism. *Nat Neurosci.* 2006;9:1218-1220.
  27. Mullen E. *Mullen Scales of Early Learning.* AGS Edition. Circle Pines, MN: American Guidance Service; 1995.
  28. Rutter M, Le Couter A, Lord C. *ADI-R: Autism Diagnostic Interview-Revised.* Los Angeles: Western Psychological Services; 2003.
  29. Rutter M, Bailey A, Lord C. *SCQ: Social Communication Questionnaire.* Los Angeles: Western Psychological Services; 2003.
  30. Gotham K, Pickles A, Lord C. Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *J Autism Dev Disord.* 2009;39:693-705.
  31. Ripley B. *Tree: Classification and Regression Trees.* R package version 1.0-29; 2011.
  32. R Core Team. *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing; 2013.
  33. Zwaigenbaum L, Bryson SE, Szatmari P, *et al.* Sex differences in children with autism spectrum disorder identified within a high-risk infant cohort. *J Autism Dev Disord.* 2012;42: 2585-2596.
  34. Lord C, Luyster R, Guthrie W, Pickles A. Patterns of developmental trajectories in toddlers with autism spectrum disorder. *J Consult Clin Psychol.* 2012;80:477-489.
  35. Chawarska K, Klin A, Paul R, Volkmar F. Autism spectrum disorder in the second year: stability and change in syndrome expression. *J Child Psychol Psychiatry.* 2007;48:128-138.
  36. Ritschard G, Zighed DA. Implication Strength of Classification Rules. *Foundations of Intelligent Systems: Springer;* 2006:463-472.

**TABLE S1** Classification Algorithm Used at 36 Months

Diagnostic Classification	Algorithmic Criteria	Sample Size N = 719
ASD	CBE diagnosis of ASD and ADOS severity score $\geq 4$	157 (21.8%)
ATYP	No CBE diagnosis of ASD above and at least 1 of the following: ADOS severity score $\geq 3$ 2 or more MSEL subtest t scores $< 35$ 1 or more MSEL subtest t scores $< 30$	178 (24.8%)
TYP	No CBE diagnosis of ASD above and all of the following: ADOS severity $\leq 2$ 0 or 1 MSEL subtest t scores $< 35$ 0 MSEL subtest t scores $< 30$	384 (53.4%)

Note: Mullen Scales of Early Learning (MSEL) subtests considered are Fine Motor, Visual Reception, Receptive Language, and Expressive Language. ASD = autism spectrum disorder; ATYP = atypical development; CBE = clinical best estimate; TYP = typical development.

**TABLE S2** Characteristics of Children With Autism Spectrum Disorder (ASD) Who Were Correctly Classified and Misclassified in the Classification and Regression Tree (CART) Analysis

Variable	Classified Correctly		Misclassified		p Value
n (%)	69	(56.6)	53	(43.4)	
Male (%)	52	(55.9)	41	(44.1)	.79
Female (%)	17	(58.6)	12	(41.4)	
Age at recruitment, mo	8.3	(3.8)	8.1	(3.4)	.84
Age at assessment, mo	18.5	(.69)	18.5	(.61)	.95
White, %	83.9		78.7		.66
Hispanic, %	12.1		9.1		.93
Maternal age, y	33.7	(4.9)	33.8	(4.4)	.98
Paternal age, y	36.0	(5.4)	36.2	(5.4)	.83
Maternal education: college degree or higher, %	64.9		72.5		.57
Paternal education: college degree or higher, %	61.0		70.7		.43

Note: n = 93 male, 29 female participants in this analysis. SD given in parentheses except where noted.