Published in final edited form as:

J Autism Dev Disord. 2012 December ; 42(12): 2636–2647. doi:10.1007/s10803-012-1521-0.

Predicting Developmental Status from 12 to 24 Months in Infants at Risk for Autism Spectrum Disorder: A Preliminary Report

Suzanne L. Macari,

Child Study Center, Yale University School of Medicine, New Haven, CT, USA. Toddler Developmental Disabilities Clinic, Yale Child Study Center, 40 Temple St, Suite 7D, New Haven, CT 06510, USA

Daniel Campbell,

Child Study Center, Yale University School of Medicine, New Haven, CT, USA

Grace W. Gengoux,

Child Study Center, Yale University School of Medicine, New Haven, CT, USA

Celine A. Saulnier,

Child Study Center, Yale University School of Medicine, New Haven, CT, USA

Ami J. Klin, and

Child Study Center, Yale University School of Medicine, New Haven, CT, USA

Katarzyna Chawarska

Child Study Center, Yale University School of Medicine, New Haven, CT, USA

Suzanne L. Macari: Suzanne.macari@yale.edu

Abstract

The study examined whether performance profiles on individual items of the Toddler Module of the Autism Diagnostic Observation Schedule at 12 months are associated with developmental status at 24 months in infants at high and low risk for developing Autism Spectrum Disorder (ASD). A nonparametric decision-tree learning algorithm identified sets of 12-month predictors of developmental status at 24 months. Results suggest that identification of infants who are likely to exhibit symptoms of ASD at 24 months is complicated by variable patterns of symptom emergence. Fine-grained analyses linking specific profiles of strengths and deficits with specific patterns of symptom emergence will be necessary for further refinement of screening and diagnostic instruments for ASD in infancy.

Keywords

Autism; Infancy; High risk studies; Longitudinal studies; Pervasive developmental disorder

Introduction

Autism Spectrum Disorders (ASD) represent a class of highly heritable (e.g., Bailey et al. 1996; Muhle et al. 2004) neurodevelopmental disorders characterized by marked deficits in social and communicative development as well as the presence of atypical sensory and motor behaviors and restricted interests (APA 2000). Because of the strong genetic basis of ASD, siblings of children with ASD are at higher risk for developing the syndrome. While older estimates suggested 2–10 % recurrence rates amongst siblings of children affected by ASD (see Bailey et al. 1996), more recent studies suggest that these rates may have been underestimated due to multiple factors including changes in diagnostic criteria (see Volkmar et al. 2007) and underappreciated ASD symptomatology in siblings (Virkud et al. 2009). In a recent prospective study of a large sample of younger siblings of children with autism, 18.7 % of the siblings were diagnosed with ASD when one child was already affected in the family, and that risk increased to over 32 % in families with two or more affected children (Ozonoff et al. 2011a). In addition to ASD, increased rates of various difficulties including cognitive and language delays as well as impairments in social interaction and play skills have been noted in an additional 18–30 % of siblings of children with ASD (Gamliel et al. 2007; Landa et al. 2007; Toth et al. 2007; Yoder et al. 2009).

Considering the high rates of enduring developmental delays and abnormalities amongst siblings of children with ASD, it has become increasingly imperative to develop strategies to identify high-risk infants who have the greatest likelihood of developing ASD and related disorders warranting clinical attention as early as possible. Recent studies of high-risk infant siblings suggest that while the full-blown syndrome might not be apparent until later in the second year of life, as a group, high-risk infants later diagnosed with ASD begin to manifest overt behavioral symptoms around 12 months of age. These include atypicalities in the domains of social communication and responsivity (Bryson et al. 2007a; Landa et al. 2007; Ozonoff et al. 2010; Rozga et al. 2010; Sullivan et al. 2007; Zwaigenbaum et al. 2005), attention to objects and object use (Ozonoff et al. 2008b), speech-like vocalizations (Paul et al. 2011), and language (Landa and Garrett-Mayer 2006; Mitchell et al. 2006; Zwaigenbaum et al. 2005). Although these differences are evident at the *group level*, prediction at the *individual level* is essential. And yet, identification of symptoms that would allow predicting with reasonable certainty which infants are likely to develop ASD or related disorders still represents an under-examined area of inquiry.

A number of factors contribute to the difficulty in identifying a set of predictors at 12 months that can inform about increased risk for ASD by the end of infancy. Research based largely on parental report suggests that emergence of behavioral symptoms of ASD might follow several patterns including “early onset”, where symptoms become apparent shortly after birth (Kanner 1943), “plateau”, where after a period of more or less typical development, rate of skill acquisition slows down and the child’s development appears to stagnate (Bryson et al. 2007; Ozonoff et al. 2008b; Siperstein and Volkmar 2004), “regression”, when a child loses previously acquired social and communication skills in the second year of life (Bryson et al. 2007a; Eisenberg and Kanner 1956; Ozonoff et al. 2010; Werner and Dawson 2005); and “delays plus loss”, with the presence of early delays combined with a loss of social-communicative skills in the second year of life (Landa et al. 2007; Ozonoff et al. 2005; Werner et al. 2005). Assuming the accuracy of the patterns described above, it would be only around 24 months or later when *all* affected children could be reliably identified and diagnosed. Furthermore, given the nature of high-risk samples, an additional challenge stems from the fact that the infants are not only at risk for developing ASD but also a range of other developmental problems including language delays, learning difficulties, and social problems, which may not lead to an ASD diagnosis, but reach levels requiring clinical attention in infancy and early childhood. Differentiating

infants who are likely to develop ASD from those who experience some problems but do not develop the full-blown syndrome is likely to be challenging, especially in the second year of life. In the same vein, identifying factors that predict typical outcomes in this cohort might be difficult as well.

Despite these difficulties, the question remains, given recent advances in the development of behavioral methods for quantifying symptoms of social impairment in infancy, how far down we can push the age at which to assess the likelihood of ASD in individual infants with reasonable certainty? To address this question we examined whether developmental status at 24 months can be predicted based on a constellation of behavioral features as measured at 12 months in infants at high and low risk for ASD. Rather than focusing on more elementary measures of social cognition and attention (e.g., Chawarska and Shic 2009; Chawarska et al. 2012; Jones et al. 2008; Shic et al. 2011), this study focused on behaviors that can be elicited, quantified, and analyzed in a context of a standard play session. The behaviors of interest were captured using the Autism Diagnostic Observation Schedule-Toddler (ADOS-T) Module (Lord et al. in press; Luyster et al. 2009). The ADOS-T is a standardized assessment measure designed for quantifying a range of social communicative and other behaviors relevant to a diagnosis of ASD in 12- to 30-month-old infants. The measure constitutes a downward extension of the Autism Diagnostic Observation Schedule-Generic (ADOS-G) Module 1 (Lord et al. 2000), and was designed to improve specificity and sensitivity of the ADOS-G in very young and nonverbal children. The Toddler Module follows the same structure as the other modules, but offers several new activities particularly suitable for infants in the second year of life, a number of new items, and an expanded scoring scale. Though developed as a diagnostic tool, the ADOS also provides detailed and graded information on a variety of communicative, social, and repetitive behaviors that are elicited and coded in a standardized manner, which can be used for quantifying changes over time as well as evaluating the effects of intervention (Carr et al. 2011; Colombi et al. 2011).

In this paper we target three questions: (1) Can we predict with reasonable certainty whether high-risk infants are likely to develop ASD-like symptoms by 24 months based on their presentations at 12 months? (2) If so, what types of behavioral markers would be helpful in assessing such risk? (3) What types of behaviors at 12 months predict a *typical* course of development in the same population? The questions have great clinical and research significance, with the first and second addressing the aim of identifying infants at risk who are likely to require intervention in the second year of life, and the third addresses the issue of behaviors that can be considered positive prognostic signs at 12 months with regard to *any* developmental problems an infant can encounter in the 2nd year. While in clinic-referred samples, clinical best estimate (CBE) diagnosis of ASD in children under the age of 2 years is relatively stable (Charman et al. 2005; Chawarska et al. 2009; Lord et al. 2006), the developmental trajectories of high-risk infants with regard to CBE diagnosis are as yet unknown. Given that some infants in our study were too young for a confirmatory diagnosis at 36 months, we focused our analysis on their developmental status at 24 months defined as either clinically significant levels of social-communicative impairment specific to ASD (ASD), typical development (TYP), as well as status that is marked by a variety of developmental delays and atypical features (ATYP). Even though the ASD presentation at 24 months might not be stable over time, the presence of significant clinical concerns should constitute a call for treating the presenting symptoms and intense monitoring of diagnostic status. Considering that in high-risk infants, symptoms of social and communicative disability are likely to be distributed along a continuum from full-blown ASD to broader autism phenotype (BAP) features, it is essential to identify and treat children who are likely to exhibit marked deficits in the second year, regardless of what the eventual diagnostic classification might be.

We hypothesized that infants who show clear symptoms of ASD at 24 months will exhibit behavioral profiles at 12 months that enable clinicians to detect enhanced risk and confirm the importance of efforts to monitor their development closely in the second year. We also hypothesized that those who will follow a typical developmental trajectory will exhibit performance profiles at 12 months that help to reassure both parents and clinicians of the low likelihood of them developing any clinically significant symptoms in the second year of life.

Methods

Participants

A total of 84 infants (54 males, 64 %) participated in a prospective study of social-cognitive development and were evaluated at 12, 18 (not included in analyses), and 24 months. Fifty-three were at high risk for ASD (HR-ASD) due to ASD in an older sibling. The older sibling's diagnosis was confirmed via clinical best estimate diagnosis, and ADI (Autism Diagnostic Interview; Lord et al. 1994) and/or ADOS. Thirty-one infants had no history of ASD in 1st or 2nd degree relatives and were considered low risk (LR) for ASD. All infants were enrolled in the study by the age of 6 months and were recruited via our existing research programs, our website, advertising, and word of mouth. Exclusionary criteria were gestational age below 34 weeks, any hearing or visual impairment, non-febrile seizure disorders, or known genetic syndrome. Given the requirements of the ADOS-T (Luyster et al. 2009), infants with nonverbal mental ages below 12 months were excluded from this analysis (4 LR, 4 HR-ASD). All parents provided permission for their infants to participate and the study was approved by the Human Investigation Committee. The vast majority of parents (93 %) identified their child's race as Caucasian, with the remaining 7 % identifying as African-American, Asian, or mixed racial heritage, with no racial differences across the groups ($\chi^2(1) = 0.31$, ns). Hispanic infants comprised 5.5 % of the sample, with a trend toward more Hispanic infants in the HR-ASD group than in the LR group (6 vs. 0, respectively; $\chi^2(1) = 3.78$, $p = .052$). The level of educational attainment of both mothers and fathers was high, corresponding to slightly more than a 4-year college degree on average, and did not significantly differ across groups ($p = .19$ for mothers; $p = .83$ for fathers).

Clinical best estimate (CBE) diagnosis—At 24 months, a team of expert clinicians specialized in early diagnosis of developmental disorders assigned the provisional CBE diagnosis. The clinicians were blind to the infants' risk status. The classification was based on results of developmental, language, and diagnostic (ADOS-G Module 1) assessments, review of medical and developmental history, DSM-IV criteria, and clinical judgment (see Chawarska et al. 2009 for details). Infants were classified as having ASD ($n = 13$; 12 HR-ASD, 1 LR) if they exhibited marked delays and abnormalities in social interaction and communication skills, evident across instruments and contexts in conjunction with parental report of similar atypical behavioral patterns. Infants in the TYP ($n = 34$; 12 HR-ASD, 22 LR) group exhibited no developmental concerns either at 18 or 24 months. Given the aims of the study, we grouped the infants such that the ASD group included only those who exhibited unambiguous symptoms at 24 months and those with TYP classification exhibited consistent typical developmental profiles at 18 and 24 months. Thus the 'catch-all' ATYP category ($n = 37$, 29 HR-ASD, 8 LR) was broadly defined and included infants who showed marked and persistent developmental delays, any atypical features, or mild and transient delays either at 18 or 24 months or both. Based on these considerations, the ATYP group included those who had a score more than 1.5 SDs below the mean on one or more subscales of the Mullen or exhibited behavior problems, e.g., disruptive behavior ($n = 20$). In addition, infants with social-communicative delays (regardless of Mullen scores) or isolated atypical

language features (e.g., echolalia), or unusual sensory and repetitive behaviors who did not meet criteria for ASD were included in this category ($n = 17$). The majority of infants in the ATYP group showed some kind of difficulties at both 18 and 24 months ($n = 29$, 78 %), with relatively few ($n = 8$, 22 %) having difficulties at 18 but not at 24 months. Infants identified as symptomatic for ASD or ATYP were more likely to be male (77 and 81 %, respectively) than those in the TYP group (41 %), $\chi^2(2) = 13.36$, $p = .001$. See Table 1 for sample characterization.

Developmental skills were assessed with the *Mullen Scales of Early Learning* (Mullen 1995). Infants in all three groups scored similarly on the nonverbal scales of the Mullen (Visual Reception and Fine Motor) at 12 months (see Table 1). However, there were significant differences in the Receptive Language and Expressive Language scales [$F(2,81) = 7.5$, $p < .001$, $\eta^2 = 0.16$, and $F(2,81) = 18.2$, $p < .001$, $\eta^2 = 0.31$, respectively]. Post hoc comparisons indicated that the TYP group scored higher on both verbal scales, but the ASD and ATYP group did not differ from each other on either scale.

Social-communicative behaviors were assessed using the *ADOS-Toddler* (ADOS-T; Lord et al. in press). Ph.D.-level clinicians administered the ADOS-T blind to each infant's risk status. All had extensive experience assessing young children with and without ASD using the ADOS-1 prior to training on the ADOS-T. Two members of the team (SM and GG) attended the ADOS-T formal training and subsequently trained the rest of the clinicians to reliability on individual items, with ongoing reliability conducted at several points throughout the study. The three groups differed significantly, $F(2,81) = 11.2$, $p < .001$, $\eta^2 = 0.22$ on the total algorithm scores at 12 months, such that the ASD group obtained the highest scores, and the TYP group the lowest (see Table 1). There was, however, marked variability with regard to the ADOS-T total scores in all three groups, with substantial overlap in the distributions (see Fig. 1).

Procedure

To examine what behavioral features as measured by the individual items of the ADOS-T at 12 months are predictive of diagnostic classification at 24 months, we conducted an item-level analysis of the ADOS-T. Due to the limited sample size relative to the number of items on the ADOS-T (84 participants vs. 41 items), traditional classification models such as discriminant function analysis and linear/logistic regression, with their associated variable selection techniques, were not feasible for investigating interactions among items. Even restricting to only second-order interactions, there would be more variables than observations, making these techniques inapplicable. Furthermore, we anticipated that different combinations of ADOS-T items might predict different patterns of symptoms emergence and such variability would complicate the use of standard regression techniques. To solve these problems, we employed classification trees, a nonparametric decision-tree learning technique (Breiman et al. 1984; Hastie et al. 2009) that uses successive, nested partitions of the data into binary categories in order to predict group membership. The classification tree algorithm finds the optimal split among all input variables and constants in terms of misclassification error, partitions the data into two subsets, and then repeats this process by finding a split on each subset until some stopping condition is reached, such as a maximum number of splits in the tree. The end result is a transparent model that naturally incorporates interactions among variables and automatically performs variable selection.

Once the classification tree is built, the fitted value for any subject is obtained as follows. Beginning at the top (the "root") of the tree, move either right or left based on the subject's value for the variable indicated at the split. Continue moving right or left according to the selected variables, in the order that they appear, until no splits remain and a "leaf" is reached. The consecutive steps down the tree specify to which subset of the data the subject

belongs, as defined by the values of all the variables above it simultaneously. Final predicted group membership for each subject is then decided by majority vote of the class labels among all subjects in each leaf of the tree.

To avoid overfitting the data, whereby the model would predict class labels of the data extremely well but would not generalize to new, out-of-sample data, we used five-fold cross-validation to estimate out-of-sample misclassification error. The dataset was randomly partitioned into five subsets of approximately equal size, but separately for each diagnostic group so that each subset contained roughly equal group representation. Each subset was treated in turn as a test set, with the remaining four subsets as the corresponding training set on which the tree model was fit. The tree was then pruned to a specific tree size varying from 2 to 15 splits, and group prediction on the test set was obtained. After repeating this process for each test set and each specified tree size, misclassification errors were averaged across all five test sets, and the tree size that minimized the average out-of-sample misclassification error was selected.

Results

TYP Versus Non-TYP Groups

Classification tree modeling predicting TYP ($n = 34$) versus non-TYP (*ASD and ATYP*; $n = 50$) developmental status at 24 months correctly classified 70.6 % ($n = 24$) of the TYP and 86.3 % ($n = 44$) of the non-TYP infants based on 3 ADOS-T items: (1) *showing (B12)*, *overactivity (E1)*, and *initiating joint attention (B13)* (see Fig. 2). Infants with typical developmental trajectories were very likely to (1) engage in at least one clear, spontaneous, and robust instance of showing objects (e.g., toys) to others during the ADOS-T play session (B12), (2) engage in at least one clear instance of spontaneous initiation of joint attention (i.e., looking at an interesting object, at an adult, and at the object again as if trying to share interest and/or attract the adult's attention to the object) (B13), and (3) have the age-appropriate ability to regulate attention and activity level (E1). Atypical performance on these items at 12 months signaled an increased risk for broadly defined developmental challenges in the second year of life, as the great majority of infants without these three strengths experienced various degrees of problems, ranging from subthreshold delays to more significant and persistent delays and ASD. There were, however, 10 (29.4 %) infants who had a typical course of development but were misclassified in this analysis. These infants showed difficulties with one or sometimes two of the three behaviors above. Furthermore, seven (13.7 %) infants who were not typically developing were classified as typical in this analysis, but these children had either relatively mild problems at 24 months (e.g., mild speech delay), or a history of such delays at 18 months that resolved by 24 months. Notably, none of the children with ASD were misclassified as TYP in this analysis. Positive predictive value (PPV) was 0.77 and negative predictive value (NPV) was 0.82, indicating a low number of false positives and false negatives. With a sensitivity (se) of 0.71 and specificity (sp) of 0.86, this combination of items and scores classified the infants into 24-month categories of TYP and non-TYP with relatively good precision.

ASD Versus Non-ASD Groups

An analogous analysis aimed at identifying items that would differentiate ASD ($n = 13$) versus non-ASD ($n = 71$; TYP and ATYP) cases, correctly classified 84.6 % ($n = 11$) of the ASD group and 95.8 % ($n = 68$) of the non-ASD infants. The items found critical to this classification were: (1) *level of engagement (B17)*, (2) *amount of requesting (B10)*, (3) *imitation (C3)*, (4) *fussiness (E2)*, (5) *showing (B12)*, (6) *gestures (A8)*, and (7) *intonation (A3)* (see Fig. 3).

The classification process for the ASD group was more complex than that for the TYP versus non-TYP groups, reflecting the heterogeneity of developmental trajectories of infants exhibiting ASD symptoms (Fig. 3). Infants with ASD symptoms appear in several leaves on the tree, indicating that different combinations of features at 12 months might be associated with an ASD presentation at 24 months in different children. We considered the four leaves that contained infants with a CBE of ASD at 24 months as four possible “subgroups” within the autism spectrum, and explored differences among the subgroups in their patterns of phenotypic features. Almost 40 % ($n = 5$) of infants with ASD were correctly classified at 12 months based on a single qualitative summary item, Level of Engagement (item B17). These infants (ASD₁) exhibited particularly poor engagement with the majority of activities and required extensive effort on the part of the clinician throughout the assessment to maintain their interest.¹ Examination of their ADOS-T total algorithm scores suggests that these infants had consistently high algorithm scores both at 12 and 24 months (see Fig. 4). Their symptoms at 12 months were expressed in the context of average nonverbal skills and moderately delayed expressive and receptive language skills (Fig. 5). Only one high-risk non-ASD infant scored similarly on item B17 at 12 months; this child exhibited significant social-communication difficulties in the first and second year, including mild global delays with some ASD symptoms at 18 months, with some residual behavioral and speech difficulties at 24 months.

While infants in ASD₁ exhibited a highly concerning level of symptoms throughout the second year of life, the remaining subgroups showed more variable developmental trends in symptom emergence. Two out of four groups showed an increase in ADOS-T algorithm scores over time (Fig. 4). Infants in ASD₂ ($n = 4$, 31 %) were classified as likely to develop ASD based on a combination of 6 items. Despite exhibiting relatively good engagement with the activities (B17) and a relatively good repertoire of gestures (A8) as well as low fussiness (E2), they made few communicative bids for requesting (B10) and showing (B12) purposes, and had poor imitation skills (C3). Interestingly, infants in ASD₃ ($n = 2$, 15 %) whose symptom severity also increased over time were misclassified by the analysis. They were classified into a small group of high-risk infants who displayed a variety of social-communicative difficulties at both 12 and 24 months (Fig. 4). Their profile of ADOS-T item scores was identical to that of the previous group except that these infants used very few gestures for communicative purposes. These children had the highest verbal skills of all infants with ASD at 12 months, which might have contributed to their misclassification (Fig. 5). By 24 months, their ADOS-T algorithm scores increased, with some improvements in language and a trend suggesting decline in non-verbal standard scores. Finally, ASD₄ ($n = 2$, 15 %) exhibited relatively good engagement with the activities, but had poor requesting and imitation skills. They were also poorly regulated (fussy) and their vocalizations had unusual intonation. The infants in this group had relatively high ADOS-T algorithm scores at 12 months, alongside impaired verbal abilities. However, as their language skills improved in a dramatic fashion in the second year (Fig. 5), their social-communicative symptoms declined rapidly (Fig. 4). Two infants without ASD were also classified into this group: both were from the high-risk group and experienced receptive and expressive language delays throughout the second year; one child exhibited subthreshold ASD features at 24 months. The ASD versus non-ASD classification method resulted in relatively high sensitivity (0.85) and high specificity (0.96), with few false positives (PPV = 0.79) and very few false negatives (NPV = 0.97).

¹ A score of 2 on this item indicates: “Engaged only when the examiner works hard to get and keep the child’s interest OR child is only engaged during snack.” A score of 3 is coded when the child is: “Not engaged at all, even when examiner makes efforts to attract the child’s interest OR the child is only engaged during games involving physical contact.”.

Discussion

The main motivation behind the study was to examine the predictive relationship between social-communicative skills at 12 months and developmental status at 24 months in a group of high- and low-risk infants. To address this complex question we selected a non-parametric decision-tree learning algorithm that allowed us to select relevant ADOS-T items and examine predictive models of developmental status. The study suggests that: (1) About two-thirds of infants at high risk for ASD experience some kind of developmental difficulties in the second year of life, ranging from mild to severe; (2) Identification of infants who are likely to exhibit clinically significant symptoms of ASD at 24 months is complex due to variable patterns in the emergence of symptoms; (3) Fine-grained analyses linking specific profiles of strengths and deficits with specific patterns of symptom emergence will provide information necessary for further refining screening and diagnostic instruments for ASD in infancy.

At 24 months, 23 % of infants in our high-risk sample exhibited frank symptoms of ASD. While it is not clear if they will retain the diagnosis at the age of 3 or 4 years, the severity of concerns warranted comprehensive intervention. This recurrence rate is consistent with that reported recently in the largest diagnostic outcome study on high-risk infants (Ozonoff et al. 2011a). In addition, many high-risk infants showed some developmental delays or other atypical patterns of development in the second year of life. For 9 % of infants, these were transient difficulties which resolved by 24 months; for 45 % of infants, the difficulties were more enduring. These findings highlight the fact that genetic vulnerability for ASD can be expressed in a highly variable manner and that development of infants at risk is likely to be marked by atypical features early on, even though in many cases, these difficulties may resolve as children reach preschool age. This certainly complicates diagnostic considerations in high-risk samples, as clinicians must develop tools for differentiating those with milder or transient difficulties from those with ASD or broader phenotype features requiring treatment.

Results of a standard developmental test at 12 months were not particularly informative with regard to predicting developmental status a year later. Scores on the ADOS-T, while statistically different between groups, were highly overlapping. Thus, to address the questions posited at the beginning of the study, we conducted item-level analyses of the ADOS-T using a classification tree approach. The analysis that identified indices of positive outcome suggested that the presence of robust communicative behaviors aimed at sharing experiences with others, including showing and using gaze to initiate joint attention, along with age-appropriate ability to regulate activity level and attention, bode well for a typical developmental outcome by the end of infancy. Nearly 71 % of high- and low-risk infants combined who followed a strictly defined typical trajectory in the second year exhibited these three behaviors around their first birthday. In comparison, 86 % of those with later delays or atypicalities ranging from mild to more serious delays or ASD did not show all three behaviors at 12 months. Consequently, lack of early-emerging attention sharing behaviors should be considered a red flag with regard to risk of any kind of developmental problems in the second year (including language delays) both in high- and low-risk infants, indicating a need for closer monitoring. These results are consistent with previous work (e.g., Bates et al. 1979; Carpenter et al. 1998; Wetherby et al. 2002) and highlight the critical role of pre-linguistic social-communication behaviors in later development, particularly in language outcomes. Although developmentally- and language-delayed children 'catch up' with these social-communication behaviors in the second year of life, research on toddlers with ASD suggests that deficits in showing and initiating joint attention tend to persist and become some of the defining features of ASD in the 2nd and 3rd year (Chawarska et al. 2007; Landa et al. 2007; Stone et al. 2004; Wetherby et al. 2004).

Predicting who is likely to develop clinically relevant symptoms of ASD at 24 months from 12-month data proved more complicated, largely due to a rather variable pattern of symptom onset (see also Bryson et al. 2007; Landa et al. 2007; Ozonoff et al. 2011b). Although the analysis classified ASD and non-ASD groups with a high level of accuracy ($se = 0.85$, $sp = 0.96$), it was achieved through identification of several subgroups of infants characterized by a different combination of features at 12 months. The first subgroup (almost 40 % of the affected infants) represents an early onset and severe impairment group, with very high total scores on the ADOS-T algorithm at 12 months, despite the fact that they displayed average nonverbal skills and only moderately delayed verbal skills. These infants displayed significant and pervasive difficulties engaging in dyadic play-based activities, leading to the global rating of their spontaneous engagement as extremely limited. However, nearly 60 % of infants with later symptoms of ASD were more able to engage with the ADOS-T activities at 12 months, and their performance profiles represented a combination of features including limited requesting, imitation, and showing. Behaviors that further defined the subgroups included fussiness, intonation, and gestures. Various combinations of these features (impaired or intact) classified correctly 85 % of infants with ASD and 96 % of those without ASD. It is important to note that classification tree analysis returns solutions based on *interactions* between items rather than simply a sum of scores on specific items as in a standard algorithm. This work empirically identified developmental trajectories that were previously hypothesized based on parental report (e.g., Ozonoff et al. 2008a).

Most of the behaviors at 12 months associated with ASD classification at 24 months were consistent with the broader literature on the impairments of young children with ASD including imitation, showing, gestures, and frequency of requesting (Landa et al. 2007; Rozga et al. 2011; Wetherby et al. 2004). A new feature, 'level of engagement' is a global rating signifying the infant's ability to engage with the examiner's activities. This rating is distinct from other items such as quality of rapport or shared enjoyment, reflecting, simply, the child's interest in and engagement with the various toys and activities throughout the assessment. Interestingly, while at 12 months, low frequency of requesting in combination with other features was associated with later ASD status, in the 2nd and 3rd year of life, this feature is likely to turn into a relative strength in toddlers with ASD. Once the toddlers acquire the ability to communicate, they communicate primarily for the purpose of requesting rather than sharing experiences, which represents one of the key diagnostic features of ASD (Mundy et al. 1994). A number of behaviors that have been found to be important diagnostic markers in toddlers with ASD did not surface in this analysis including deficits in pointing, response to joint attention, functional play, shared enjoyment, or eye contact. It is likely that some of the deficits at this age are not ASD-specific (e.g., limited response to joint attention or pointing), whereas others might not be robust even in typically developing infants (e.g., functional play). It is also likely that some other behaviors such as eye contact or sharing enjoyment are still intact at 12 months in some of the infants; intensification of ASD symptoms later in the second year then lead to the decline of these important dyadic behaviors. These findings illustrate an important point regarding the developmental character of the disorder, whose symptoms are likely to be expressed differently depending on the individual's mental and chronological age. This is particularly true in the first 5 years of life when children undergo a myriad of qualitative and quantitative changes in their attentional, representational, and social development, and as such needs to be considered in discussions regarding early markers and underlying mechanisms.

These results underscore the necessity of conceptualizing the earliest symptoms of ASD within a developmental framework and provide direct evidence that patterns of emergence of behavioral symptoms of ASD are indeed variable (see also Ozonoff et al. 2008a; Ozonoff et al. 2011b). Above all, the findings illustrate the complexities involved in identifying early signs of ASD in high-risk infants at 12 months and suggest that it may not be possible to

establish a universal set of behaviors that predict developmental status and distinguish high-risk infants with ASD from those with other delays. Instead, our work highlights the possibility that at this early age, there may be clusters of symptoms that correspond to specific patterns of symptom emergence. The clinical significance of these clusters will have to be interpreted in the context of other developmental dimensions such as verbal and nonverbal skills. Just as it is critical to consider developmental skill profiles when making diagnostic differentiations in toddlers (e.g., Chawarska et al. 2009), we will need to consider the *constellation* of delays, abnormalities, and strengths in 12-month-old infants at risk for ASD in attempts to predict their developmental trajectories. The fine-grained behavioral data provided by the ADOS-T are ideal for further exploration of classification issues with larger samples.

Limitations

This study represents the first step in the direction of identifying behavioral markers of ASD at 12 months based on an assessment measure that can be administered readily in most clinical settings. Limitations of the present study include the relatively small sample size as well as the lack of a confirmatory diagnosis at 36 months of age. Furthermore, the study attracted a large number of families with high socio-economic and educational status. Extension of this work to a broader socio-economic spectrum will be necessary in the future. Lastly, the results of classification trees modeled on a given sample, particularly if the sample is small, are necessarily tentative and may not be generalizable to other populations. We strived to attenuate overfitting of the models with the use of fivefold cross-validation, but our findings remain exploratory and need to be replicated with larger samples of infants at risk for ASD.

Conclusions

High levels of concern amongst parents of high-risk infants (Gengoux et al. 2010; Ozonoff et al. 2009) as well as encouraging reports regarding effects of early intervention (Dawson et al. 2010), have made efforts aimed at extending downward the age of the ASD diagnosis one of the key priorities in the field of autism research and clinical practice. Our work addresses two fundamental issues, one clinical and one methodological. First, there is a need to establish reliable diagnostic criteria for ASD as early as possible and, at the same time, find behavioral indicators that will help reassure parents who are concerned about their infants' developmental outcomes. The results suggest that while it is possible to identify infants at highest risk for developing ASD and other developmental problems at 12 months, the task is complex and has to consider diverse symptom onset patterns. Second, given the perennial problems in autism research related to small sample sizes, marked heterogeneity of symptom expression, as well as multiplicity of measures and subsequent issues of variable selection, we need to explore new approaches to generating predictive models of diagnostic classification. This study represents a step in this direction, although, undoubtedly, a tremendous amount of work remains to be done.

Acknowledgments

This study was supported by National Institute of Child Health and Human Development P01 HD003008, Project 1 (PI: KC); National Institutes of Mental Health R01 MH 087554-01 (PI: KC); Simons Foundation 187398 (PI: AK). Preliminary data from this study were presented at the International Meeting for Autism Research, May 2011, San Diego. We would like to thank Karen Bearss, Amanda Steiner, Tina Goldsmith, Anne Snow, Rhea Paul, Elizabeth Schoen Simmons, Megan Lyons, and Sarita Austin for their contribution to subject characterization; Karyn Bailey, Daniela Blum, and Amy Carney for their clinical support of the families; and Amy Margolis, Martha Dye, Kerry O'Loughlin, Jessi Garzarek, Amanda Smith, Deanna Simeone, Mairin Meltvedt, Marika Coffman, Grace Chen, Jessica Bradshaw, Brittany Butler, and Jessa Reed for their assistance with data collection. We wish to express our sincere appreciation to the families and their infants for their time and participation.

References

- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4. Washington, DC: Author; 2000. text revision
- Bailey A, Phillips W, Rutter M. Autism: Towards an integration of clinical, genetic, neuropsychological, and neurobiological perspectives. *Journal of Child Psychology and Psychiatry*. 1996; 37(1):89–126. [PubMed: 8655659]
- Bates, E.; Benigni, L.; Bretherton, L.; Camaioni, L.; Volterra, V. The emergence of symbols: Cognition and communication in infancy. New York: Academic Press; 1979.
- Breiman, L.; Friedman, J.; Olshen, R.; Stone, C. Classification and regression trees. New York, NY: Wadsworth; 1984.
- Bryson SE, Zwaigenbaum L, Brian J, Roberts W, Szatmari P, Rombough V, et al. A prospective case series of high-risk infants who developed autism. *Journal of Autism and Developmental Disorders*. 2007; 37(1):12–24. [PubMed: 17211728]
- Carpenter M, Nagell K, Tomasello M, Butterworth G, Moore C. Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the society for research in child development*. 1998; 63(4):1–143.
- Carr, T.; Colombi, C.; MacDonald, M.; Lord, C. Measuring response to intervention with the autism diagnostic observation schedule-change (ADOS-C). Presented at the Society for Research in Child Development; Montreal, CA. 2011.
- Charman T, Taylor E, Drew A, Cockerill H, Brown JA, Baird G. Outcome at 7 years of children diagnosed with autism at age 2: Predictive validity of assessments conducted at 2 and 3 years of age and pattern of symptom change over time. *Journal of Child Psychology and Psychiatry*. 2005; 46(5): 500–513. [PubMed: 15845130]
- Chawarska K, Klin A, Paul R, Macari S, Volkmar F. A prospective study of toddlers with ASD: Short-term diagnostic and cognitive outcomes. *Journal of Child Psychology and Psychiatry*. 2009; 50(10): 1235–1245. [PubMed: 19594835]
- Chawarska K, Klin A, Paul R, Volkmar F. Autism Spectrum Disorder in the second year: Stability and change in syndrome expression. *Journal of Child Psychology and Psychiatry*. 2007; 48(2):128–138. [PubMed: 17300551]
- Chawarska K, Macari S, Shic F. Context modulates attention to social scenes in toddlers with autism. [Electronic version]. *Journal of Child Psychology and Psychiatry*. 2012
- Chawarska K, Shic F. Looking but not seeing: Abnormal visual scanning and recognition of faces in 2 and 4-year old children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*. 2009; 39(12):1663–1672. [PubMed: 19590943]
- Colombi, C.; Carr, T.; MacDonald, M.; Lord, C. Developing a measure of treatment outcome: The Autism Diagnostic Observation Schedule-Change. Presented at the Society for Research in Child Development; Montreal, CA. 2011.
- Dawson G, Rogers S, Munson J, Smith M, Winter J, Greenson J, et al. Randomized, controlled trial of an intervention for toddlers with autism: The early start denver model. *Pediatrics*. 2010; 125(1):e17–e23. [PubMed: 19948568]
- Eisenberg L, Kanner L. Early infantile autism, 1943–55. *American Journal of Orthopsychiatry*. 1956; 26:556–566. [PubMed: 13339939]
- Gamliel I, Yirmiya N, Sigman M. The development of young siblings of children with autism from 4 to 54 Months. *Journal of Autism and Developmental Disorders*. 2007; 37(1):171–183. [PubMed: 17203244]
- Gengoux, GW.; Steiner, AM.; Chawarska, K. Parent stress related to infants at-risk for Autism Spectrum Disorder. Presented at the International Meeting for Autism Research; Philadelphia, PA. 2010.
- Hastie, T.; Tibshirani, R.; Friedman, J. The elements of statistical learning: Data mining, inference, and prediction. 2. New York, NY: Springer; 2009.
- Jones W, Carr K, Klin A. Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-olds with autism spectrum disorder. *Archives of General Psychiatry*. 2008; 65:946–954. [PubMed: 18678799]

- Kanner L. Autistic disturbances of affective contact. *Nervous Child*. 1943; 2:217–250.
- Landa R, Garrett-Mayer E. Development in infants with Autism Spectrum Disorders: A prospective study. *Journal of Child Psychology and Psychiatry*. 2006; 47(6):629–638. [PubMed: 16712640]
- Landa RJ, Holman KC, Garrett-Mayer E. Social and communication development in toddlers with early and later diagnosis of Autism Spectrum Disorders. *Archives of General Psychiatry*. 2007; 64(7):853. [PubMed: 17606819]
- Lord, C.; Luyster, R.; Gotham, K.; Guthrie, W.; Risi, S.; Rutter, M. Autism diagnostic observation schedule—toddler module manual. Los Angeles, CA: Western Psychological Services; (in press)
- Lord C, Risi S, DiLavore PS, Shulman C, Thurm A, Pickles A. Autism from 2 to 9 years of age. *Archives of General Psychiatry*. 2006; 63(6):694–701. [PubMed: 16754843]
- Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, DiLavore PC, et al. The autism diagnostic observation schedule—generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*. 2000; 30(3):205–223. [PubMed: 11055457]
- Lord C, Rutter M, LeCouteur A. Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*. 1994; 24(5):659–685. [PubMed: 7814313]
- Luyster R, Gotham K, Guthrie W, Coffing M, Petrak R, Pierce K, et al. The autism diagnostic observation schedule—toddler module: A new module of a standardized diagnostic measure for Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*. 2009; 39(9):1305–1320. [PubMed: 19415479]
- Mitchell S, Brian J, Zwaigenbaum L, Roberts W, Szatmari P, Smith I, et al. Early language and communication development of infants later diagnosed with Autism Spectrum Disorder. *Journal of Developmental and Behavioral Pediatrics*. 2006; 27(2):S69. [PubMed: 16685188]
- Muhle R, Trentacoste SV, Rapin I. The genetics of autism. *Pediatrics*. 2004; 113(5):e472–e486. [PubMed: 15121991]
- Mullen, EM. Mullen Scales of Early Learning. Circle Pines, MN: American Guidance Service; 1995.
- Mundy P, Sigman M, Kasari C. Joint attention, developmental level, and symptom presentation in autism. *Development and Psychopathology*. 1994; 6(3):389–401.
- Ozonoff S, Heung K, Byrd R, Hansen R, Hertz-Picciotto I. The onset of autism: Patterns of symptom emergence in the first years of life. *Autism Research*. 2008a; 1(6):320–328. [PubMed: 19360687]
- Ozonoff S, Iosif AM, Baguio F, Cook IC, Hill MM, Hutman T, et al. A prospective study of the emergence of early behavioral signs of autism. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2010; 49(3):256–266. [PubMed: 20410715]
- Ozonoff S, Macari S, Young GS, Goldring S, Thompson M, Rogers SJ. Atypical object exploration at 12 months of age is associated with autism in a prospective sample. *Autism*. 2008b; 12(5):457–472. [PubMed: 18805942]
- Ozonoff S, Williams BJ, Landa R. Parental report of the early development of children with regressive autism. *Autism*. 2005; 9(5):461–486. [PubMed: 16287700]
- Ozonoff S, Young GS, Carter A, Messinger D, Yirmiya N, Zwaigenbaum L, Bryson S, et al. Recurrence risk for Autism Spectrum Disorders: A Baby Siblings Research Consortium study. *Pediatrics*. 2011a; 128:2010–2825.
- Ozonoff, S.; Young, GS.; Hutman, T.; Sigman, M.; Rogers, SJ.; Hill, MM.; Steinfeld, MB.; Rozga, A. Stability of ASD in a high risk sample. Paper presented at the Biennial Meeting of the Society for Research in Child Development (SRCD); Montreal. 2011b.
- Ozonoff S, Young GS, Steinfeld MB, Hill MM, Cook I, Hutman T, et al. How early do parent concerns predict later autism diagnosis? *Journal of Developmental and Behavioral Pediatrics*. 2009; 30(5):367. [PubMed: 19827218]
- Paul R, Fuerst Y, Ramsay G, Chawarska K, Klin A. Out of the mouths of babes: Vocal production in infant siblings of children with ASD. *Journal of Child Psychology and Psychiatry*. 2011; 52(5): 588–598. [PubMed: 21039489]
- Rozga A, Hutman T, Young GS, Rogers SJ, Ozonoff S, Dapretto M, et al. Behavioral profiles of affected and unaffected siblings of children with autism: Contribution of measures of mother–

- infant interaction and nonverbal communication. *Journal of Autism and Developmental Disorders*. 2011; 41(3):287–301. [PubMed: 20568002]
- Shic F, Bradshaw J, Klin A, Scassellati B, Chawarska K. Limited activity monitoring in toddlers with autism spectrum disorder. *Brain Research*. 2011; 1380:246–254. [PubMed: 21129365]
- Siperstein R, Volkmar F. Brief report: Parental reporting of regression in children with pervasive developmental disorders. *Journal of Autism and Developmental Disorders*. 2004; 34(6):731–734. [PubMed: 15679193]
- Stone WL, Coonrod EE, Turner LM, Pozdol SL. Psychometric properties of the STAT for early autism screening. *Journal of Autism and Developmental Disorders*. 2004; 34(6):691–701. [PubMed: 15679188]
- Sullivan M, Finelli J, Marvin A, Garrett-Mayer E, Bauman M, Landa R. Response to joint attention in toddlers at risk for Autism Spectrum Disorder: A prospective study. *Journal of Autism and Developmental Disorders*. 2007; 37(1):37–48. [PubMed: 17216332]
- Toth K, Dawson G, Meltzoff AN, Greenson J, Fein D. Early social, imitation, play, and language abilities of young non-autistic siblings of children with autism. *Journal of Autism and Developmental Disorders*. 2007; 37(1):145–157. [PubMed: 17216560]
- Virkud YV, Todd RD, Abbacchi AM, Zhang Y, Constantino JN. Familial aggregation of quantitative autistic traits in multiplex versus simplex autism. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*. 2009; 150B(3):328–334.
- Volkmar, F.; Chawarska, K.; Klin, A. *Autism Spectrum Disorders in infants and toddlers: Diagnosis, assessment, and treatment*. New York, NY: Guilford Press; 2007. *Autism Spectrum Disorders in infants and toddlers: An introduction*; p. 1-22.
- Werner E, Dawson G. Validation of the phenomenon of autistic regression using home videotapes. *Archives of General Psychiatry*. 2005; 62(8):889–895. [PubMed: 16061766]
- Werner E, Dawson G, Munson J, Osterling J. Variation in early developmental course in autism and its relation with behavioral outcome at 3–4 years of age. *Journal of Autism and Developmental Disorders*. 2005; 35(3):337–350. [PubMed: 16119475]
- Wetherby A, Allen L, Cleary J, Kublin K, Goldstein H. Validity and reliability of the Communication and Symbolic Behavior Scales Developmental Profile with very young children. *Journal of Speech, Language and Hearing Research*. 2002; 45(6):1202–1218.
- Wetherby A, Woods J, Allen L, Cleary J, Dickinson H, Lord C. Early indicators of Autism Spectrum Disorders in the second year of life. *Journal of Autism and Developmental Disorders*. 2004; 34:473–493. [PubMed: 15628603]
- Yoder P, Stone WL, Walden T, Malesa E. Predicting social impairment and ASD diagnosis in younger siblings of children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*. 2009; 39(10):1381–1391. [PubMed: 19449096]
- Zwaigenbaum L, Bryson S, Rogers T, Roberts W, Brian J, Szatmari P. Behavioral manifestations of autism in the first year of life. *International Journal of Developmental Neuroscience*. 2005; 23(2–3):143–152. [PubMed: 15749241]

Fig. 1. Histogram illustrating the proportion of infants receiving total algorithm scores from 0 to 25 on ADOS-T at 12 months in infants with typical (TYP), atypical (ATYP), and ASD developmental status at 24 months. The vertical lines depict their respective group means

Fig. 2.

TYP versus non-TYP classification tree. Numbers at each junction (0–3) represent the actual ADOS-T scores attained on a given item by each subgroup. The numbers on each branch represent the percentage of the total subjects that fall into each region of the tree. Under each leaf is the percentage of subjects within the leaf that are TYP or non-TYP. For instance, for the leaf defined by scores of 2 or higher item B12, 50 % of total subjects fell into that leaf, and 79 % of subjects in the leaf were non-TYP. Selected variables: showing (B12), overactivity (E1), and initiates joint attention (B13)

Fig. 3.

ASD versus non-ASD classification tree. Numbers at each junction (0–3) represent the actual ADOS-T scores attained on a given item by each subgroup. The *numbers* on each branch represent the percentage of the total subjects that fall into that region of the tree. Under each leaf is the percentage of subjects within the leaf that are ASD or non-ASD. For instance, for the leaf defined by scores of 2 or higher on item B17, 7.1 % of total subjects fell into that leaf, and 83 % of subjects in the leaf were ASD. Selected variables: level of engagement (B17), amount of requesting (B10), imitation (C3), fussiness (E2), showing (B12), intonation (A3), and gestures (A8). Leaves that contained infants with a CBE of ASD at 24 months according to this tree are highlighted with colored circles: ASD₁: severe, stable symptoms; ASD₂, ASD₃: increasing symptoms, ASD₄: decreasing symptoms

Fig. 4.
ADOS-T total algorithm scores at 12–24 months for the four ASD subgroups

Fig. 5.
Mullen domain DQ scores at 12 and 24 months for the four ASD subgroups

Table 1

Sample characterization at 12, 18, and 24 months

Age		ASD (n = 13)	ATYP (n = 37)	TYP (n = 34)
12 months	% Male	77 ^a	81 ^a	42 ^b
	GA (weeks)	39.0 (1.3) ^a	39.0 (1.4) ^a	39.0 (1.8) ^a
	Birth weight (g)	3563.1 (342.0) ^a	3501.7 (385.5) ^a	3468.4 (624.6) ^a
	Chronological age	12.2 (0.4) ^a	12.2 (0.7) ^a	12.2 (0.4) ^a
	Visual reception DQ	104.2 (10.5) ^a	113.9 (16.3) ^a	113.5 (14.7) ^a
	Fine motor DQ	110.5 (15.4) ^a	117.5 (15.4) ^a	117.4 (14.0) ^a
	Receptive language DQ	79.7 (20.8) ^a	84.2 (20.0) ^a	99.4 (18.0) ^b
	Expressive language DQ	67.4 (17.4) ^a	80.1 (17.3) ^a	100.1 (20.4) ^b
18 months	ADOS-T total	14.9 (6.1) ^a	11.1 (4.7) ^b	7.7 (4.6) ^c
	Chronological age	18.4 (.58) ^a	18.5 (.91) ^a	18.2 (.55) ^a
	Visual reception DQ	96.7 (18.8) ^a	98.8 (16.0) ^a	111.5 (17.2) ^b
	Fine motor DQ	98.9 (9.2) ^a	101.6 (14.4) ^a	106.0 (8.2) ^a
	Receptive language DQ	85.2 (34.8) ^a	89.9 (25.2) ^a	118.9 (23.4) ^b
	Expressive language DQ	69.6 (31.3) ^a	91.1 (23.5) ^b	110.7 (25.5) ^c
24 months	ADOS-T total	13.7 (4.4) ^a	8.6 (5.3) ^b	3.9 (2.7) ^c
	Chronological age	24.1 (0.5) ^a	24.8 (1.6) ^a	24.2 (0.5) ^a
	Visual reception DQ	102.0 (26.5) ^a	109.4 (16.7) ^a	124.3 (18.6) ^b
	Fine motor DQ	101.6 (11.1) ^a	101.4 (12.8) ^a	106.1 (10.0) ^a
	Receptive language DQ	98.5 (27.2) ^a	107.5 (20.9) ^a	125.7 (19.3) ^b
	Expressive language DQ	93.0 (25.3) ^a	112.2 (25.5) ^b	128.4 (24.0) ^c
	ADOS-T total	16.0 (5.0) ^a	7.0 (3.8) ^b	3.3 (2.4) ^c

Means with different superscripts were significantly different at $p < .05$ after Bonferroni correction