J Autism Dev Disord (2007) 37:98–107
DOI 10.1007/s10803-006-0340-6

ORIGINAL PAPER

Predictors of Optimal Outcome in Toddlers Diagnosed with Autism Spectrum Disorders

Saasha Sutera · Juhi Pandey · Emma L. Esser · Michael A. Rosenthal · Leandra B. Wilson · Marianne Barton · James Green · Sarah Hodgson · Diana L. Robins · Thyde Dumont-Mathieu · Deborah Fein

Published online: 6 January 2007

Abstract A diagnosis of autism spectrum disorder (ASD) is usually taken to be permanent. In this study, 13 two-year-old children with ASD lost the diagnosis by age 4, at which time they scored within the normal range on standardized measures of cognitive and adaptive functioning. No differences were found in symptom severity, socialization, or communication between children who lost the ASD diagnosis and children who did not, but children with PDD-NOS were significantly more likely than those with full autistic disorder to move off the spectrum. The clearest distinguishing factor was motor skills at age 2. Results support the idea that some toddlers with ASD can lose their diagnosis and suggest that this is difficult to predict.

Keywords Autism · Autistic spectrum · Optimal outcome · Recovery

Introduction

Autistic spectrum disorders (ASD) can be accurately diagnosed in children as young as 20 months (Cox et al., 1999). Characterized by impairments in the areas of social interaction, communication, and restricted patterns of behavior (APA, 1994), children with ASD show early behavioral signs of ASD which manifest as limitations in joint attention, eye contact, reciprocal smiling (Robins, Fein, Barton, & Green, 2001), play skills, and imitation (Rogers, Hepburn, Stackhouse & Wehner, 2003). Findings of Stone and Hogan (1993), Osterling and Dawson (1994), Baranek (1999), and Lord, Storoschuk, Rutter, and Pickles (1993) confirm that the presence of repetitive behaviors does not differentiate children with autism and children with other developmental delays at this early stage of development.

An early diagnosis of ASD is generally stable over time (Eaves & Ho, 2004; Lord, 1995; Moore & Goodson, 2003; Stone et al., 1999). Using clinical judgment to assign a final diagnosis, Eaves and Ho (2004) found diagnostic stability to be at 79%, with 93% of the children remaining on the autism spectrum and the remaining 7% moving off the spectrum, in a group of 49 children who were evaluated at 30 months and again at 54 months. Lord (1995) found that 14 of 16 children (88%) who received a diagnosis of autism at age 2 received an independent diagnosis of autism at age 3. Stone et al. (1999) found that of 37 children diagnosed with an ASD at 2-years old (25 autism, 12 PDD-NOS), 35 of them retained the diagnosis of an autistic spectrum disorder 1 year later when evaluated by an independent clinician. Moore and Goodson (2003) diagnosed 16 children with an ASD between the ages of 29 and 40 months. When reassessed 2 years later, all of these children remained within the autism spectrum.

Charman et al. (2005) followed a group of 26 children with ASD from the age of 2–7 years. Consis-

S. Sutera (✉) · J. Pandey · E. L. Esser · M. A. Rosenthal · L. B. Wilson · M. Barton · J. Green · S. Hodgson · D. L. Robins · T. Dumont-Mathieu · D. Fein
Department of Psychology, University of Connecticut, 406 Babbidge Road, Unit 1020, Storrs, CT 06269, USA
e-mail: saasha.sutera@uconn.edu

Present Address:
D. L. Robins
Department of Psychology, Georgia State University, Atlanta, GA, USA

tent with previous studies, they found that the ASD diagnosis was stable over time. Twenty-six children diagnosed with childhood autism at age 2 years were reassessed at the ages of 3 years and 7 years. Based on clinical judgment, all but one of these children remained on the autism spectrum at the age of 7 years.

Of particular interest are the children who are diagnosed with an ASD at a young age and eventually lose this diagnosis. The idea of ''recovery'' in children with autism was originally introduced by Lovaas in 1987 when he found that 47% of children who received intensive behavioral treatment achieved normal intellectual and educational functioning. Sigman and Ruskin (1999) followed children from their various research programs, dating as far back as 1979. In their sample of 51 children who were originally diagnosed with an ASD (mean age = 45 months), 17% lost their diagnosis of an ASD over time (at a mean age of 154 months). This finding is not as striking as Lovaas' report; the differences may be due to the fact that not all of the children who Sigman and Ruskin followed received intensive intervention. However, the Sigman and Ruskin study was consistent in finding a number of children who lost their diagnosis of an ASD and achieved an optimal outcome.

Intensive early intervention specialized for children with ASD and their specific learning patterns has been most effective in producing quantifiable gains with this population (Sallows & Graupner, 2005). In general, effective treatments include high numbers of treatment hours per week, direct instruction, an emphasis on attending to others and social skill training, development of imitation skills, a focus on functional language development, and high levels of structure and consistency (Dawson & Osterling, 1997; Rogers, 2001). Examples of positive outcomes noted in previous studies include gains in IQ, speech, social interaction, and improved transition and performance in elementary school (McEachin, Smith, & Lovaas, 1993; Sallows & Graupner, 2005). Sallows and Graupner (2005) replicated Lovaas' findings by employing his method of intervention and comparing a clinic-directed treatment group to a parent-directed treatment group of children. Both groups of children with ASD received intensive intervention during the first year of treatment (with a mean of 39 h per week of intervention in the clinic-directed group and 32 h per week in the parent-directed group). As children were integrated into a school setting, amount of intervention gradually decreased. There were few, if any, differences between the two groups at follow-up, but the authors note that even the parent-directed group received intensive intervention. Across groups, 48% of children were succeeding in a typical classroom, and after 4 years of treatment, these children attained an IQ above 85. All of these children displayed rapid acquisition of new material early in treatment. Pre-treatment characteristics, rather than clinic versus parent-directed intervention and hours per week of intervention, were most likely to predict outcome. Pre-treatment imitation skills, receptive language, IQ, social interaction, and adaptive skills predicted post-treatment IQ, adaptive skills, and language scores. In addition, the authors reported that many of these children had their diagnosis of an ASD eventually removed by their referring psychiatrist. These findings suggest that treatment was necessary but not sufficient in achieving optimal outcome.

It might be expected that residual symptoms, especially in the language, attention, and social domains, would persist in children who are no longer diagnosable with an ASD. A recent study by our group (Fein, Dixon, Paul, & Levin, 2005) documented 11 children who moved from diagnoses of ASD to diagnoses of ADHD, by a mean age of 7 years. All had significant attentional impairments, but social functioning varied from excellent to significantly impaired; however, when social functioning was impaired, it tended to have an ADHD quality (immature, clumsy, impulsive) rather than an autistic quality (odd or aloof). Another study by our group (Kelley, Paul, Fein, & Naigles, 2006) followed 14 children who had moved off the autistic spectrum. Employing an extensive language evaluation, we found that standard scores on language tasks were all in the normal range, but that probing with more complex language and social cognitive tasks such as comprehension of second order theory of mind and mental state verbs, ability to construct narratives, and ability to reason inductively about animate things still showed residual difficulties. Ongoing follow-up of these children, however, is showing that these residual deficits tend to subside. Both of these papers (Fein et al., 2005; Kelley et al., 2006) document optimal outcome in children with firm ASD diagnoses, but the children in these studies were all school age; in the present paper, we report on a similar outcome in an independent sample of preschool aged children.

Although movement off of the autism spectrum has not yet been the focus of many studies, some studies have examined the relationship between treatment and positive outcome, as defined by improvement in autistic symptoms. However, intervention did not lead independently to a positive outcome. While intervention was highly important, prognosis depended upon the relationship between child characteristics, such as initial language and joint attention skills, and amount of intervention (Bono, Daley, & Sigman,

2004). Koegel, Koegel, and McNerney (2001) hypothesized that there were pivotal areas that influenced outcome. Pivotal skills, such as spontaneous initiations of social behavior, were related to a positive response to treatment. Further, children could be taught these pivotal skills, thereby leading to favorable treatment outcomes (Koegel, Koegel, Shoshan, & McNerney, 1999). Specifically, the authors found that self-initiated social communication at pre-treatment was associated with successful treatment outcome.

Cognitive level has also been shown to have a significant effect on outcome (Fein et al., 1999; Gabriels, Hill, Pierce, Rogers, & Wehner, 2001), and it has been found to be the best single predictor of developmental trajectory (Fein et al., 1999). Fein and colleagues found that children with high nonverbal intelligence were more likely to display symbolic play skills, communicative skills, and social improvement over time, as compared to children with lower nonverbal IQ. In addition to cognitive level, improvement in language skills is associated with response to bids for joint attention (Bono, Daley, & Sigman, 2004). Similarly, Charman et al. (2005) found that rate of nonverbal communicative attempts at age 2 years predicted expressive language at age 3 and receptive language at age 7 years. Sigman and McGovern (2005) found that early childhood predictors of adolescent language skills included functional play skills, response to joint attention, and frequency of requests.

Surprisingly, symptom severity appears to have little predictive power in determining outcome (Fein et al., 1999; Stevens et al., 2000; Szatmari, Bryson, Boyle, Streiner, & Duku, 2003). The predictive ability of initial autistic symptoms is weak (Szatmari et al., 2003). Rather, nonverbal cognitive ability and language ability predict outcome more reliably than does early symptom severity.

The focus of the current study is those children with an early diagnosis on the autism spectrum who achieved an optimal outcome. Optimal outcome in the present study is defined by (1) initially meeting DSM-IV criteria for PDD-NOS or Autistic Disorder, *and* at follow-up (2) no longer meeting criteria for any ASD, and (3) functioning in the average range on standardized measures of cognition, language and adaptive skills. Residual problems, such as pragmatic language difficulties or social awkwardness, may still be present.

The above studies suggest that individual child characteristics are important in predicting relatively positive outcome in children who retain the ASD diagnosis. These same findings regarding positive outcome might apply to the children who move off the autism spectrum and achieve an optimal outcome. In the present study we examine whether communication skills, daily living skills, social skills, motor skills, cognitive ability, and symptom severity at the age of 2 years predict loss of an ASD diagnosis at the age of 4 years.

Methods

Participants

All children were part of a larger study aimed at developing and validating an effective screening tool for autism spectrum disorders, the Modified Checklist for Autism in Toddlers (M-CHAT; Robins et al., 2001), for toddler age children in an unselected population and in high-risk populations. Children were screened between the ages of 16 and 30 months at (1) well-child visits with their primary-care provider, (2) intake visits with an early intervention agency, or (3) if they were the younger sibling of a child diagnosed with an ASD. Groups 2 and 3 comprised the high risk sample. A total of 90 children were evaluated after screening positive on the M-CHAT at the age of approximately 2 years (Time 1) and re-evaluated at the age of approximately 4 years (Time 2). Of these 90 children, 73 were diagnosed with an ASD at their initial evaluation, and 17 were diagnosed with a non-autistic spectrum disorder, such as language or developmental delay. At re-evaluation, all of the non-ASD spectrum children remained non-ASD (NON–NON), while 13 of the 73 children originally diagnosed with an ASD no longer met criteria for a diagnosis of an ASD (ASD–NON), and the remaining 60 children remained on the spectrum (ASD–ASD).

The mean age of the children in the ASD–ASD group was 27.6 months (SD = 4.7) at Time 1, and 52.2 months (SD = 6.6) at Time 2. For the children in the ASD–NON group, the mean age was 26.5 months (SD = 4.9) at Time 1, and 54.4 (SD = 10.1) at Time 2. The mean age of the children NON–NON group was 28.0 months (SD = 3.9) at Time 1, and 57.2 months (SD = 9.4) at Time 2. There were no significant differences among groups in age at Time 1 or Time 2. There were eight females (13.3%) in the ASD–ASD group, one female (7.6%) in the ASD–NON group, and five females (29.4%) in the NON–NON group. Within the ASD–ASD group, 49 children were referred from early intervention sites, 8 were referred from a pediatrician's office, and 1 was a younger sibling of a child with an ASD; within the ASD–NON group, 11 children were referred from early intervention sites, and 2 were younger siblings; and

within the NON–NON group, 12 children were referred from early intervention sites, and 5 were referred from a pediatrician's office (see Table 1).

Children were excluded from the present study if they had already received a diagnosis of ASD or other disorder (e.g., global developmental delay), if they were already receiving early intervention services of more than one to 2 h per week, if they were older than 30 months or younger than 16 months when their caregiver completed the screener, or if they had a severe impairment that prevented the use of standardized evaluation instruments (e.g., blind, deaf, unable to sit independently). Additionally, four children were excluded from the present study because they originally had a diagnosis of an ASD at Time 1 and were diagnosed with non-autistic mental retardation at Time 2. The reason for their exclusion was that we were interested in children with the most optimal trajectory, not children where what initially appeared to be autism resolved into mental retardation.

Instruments

The M-CHAT (Robins et al., 2001) is a 23-item yes–no parent report screening cheklist for ASD, which includes items related to socialization, communication, and atypical behaviors. Failure (screening positive) on the checklist is defined as any three items failed, or any two critical items failed, which is then followed by a telephone interview that asks caregivers to elaborate upon failed items. A free developmental and diagnostic evaluation was offered to those whose responses continued to indicate failure after the follow-up phone interview.

Clinical judgment by experienced clinicians is considered to be the ''gold standard'' for autism diagnosis (Volkmar, Chawarska, & Klin, 2005). In the current study, experienced clinicians used the DSM-IV criteria for Autistic Disorder (APA, 1994), and each symptom for ASD was marked as either present or absent based on clinical judgment using all available information. Clinical judgment diagnosis was used as the final diagnosis. A child could be diagnosed with Autistic Disorder, PDD-NOS, or as nonautistic.

The Autism Diagnostic Interview-Revised (ADI-R; Lord, Rutter, & Le Couteur, 1994) is a semi-structured clinician-based interview for caregivers that evaluates the child's communication, social development, play, and restricted, repetitive, and stereotyped behaviors. A child can be classified as having Autistic Disorder or as non-autistic.

The Autism Diagnostic Observation Schedule-Generic (ADOS; Lord, Rutter, DiLavore, & Risi, 1999) is a semi-structured assessment of communication, social interactions and relatedness, play, imagination, and stereotyped or repetitive behaviors. This measure yields scores in the social domain, communication domain, and a combined score. A child can be classified as having Autistic Disorder, PDD-NOS, or as nonautistic.

The Childhood Autism Rating Scale (CARS, Schopler, Reichler, & Renner, 1988) consists of 15 items intended to measure the presence and severity of symptoms found in pervasive developmental disorders. The CARS includes items on socialization, communication, emotional responses, and sensory sensitivities. A child can be classified with mild, moderate, or severe autism, or as non-autistic. The cutoff for autism is 30 or higher.

Vineland Adaptive Behavior Scales (VABS; Sparrow, Balla, & Cicchetti, 1984) is a widely used parent interview scale that assesses adaptive functioning in the areas of communication, daily living, socialization, and motor skills, as well as yielding an adaptive behavior composite score.

Mullen Scales of Early Learning (Mullen, 1995) is an instrument that measures cognitive functioning. The subscales of fine motor, receptive and expressive language, and visual problem solving were administered. It also yields a composite score based on these four scales.

Bayley Scales of Infant Development, Second Edition (Bayley, 1993) is an instrument that measures

Table 1 Participant characteristics

	ASD–NON (n = 13)	ASD–ASD (n = 60)	NON–NON (n = 17)
Age at Time 1 (M, SD)	26.72, 5.10	27.6, 4.7	27.91, 4.07
Age at Time 2 (M, SD)	55.58, 9.7	52.2, 6.6	57.41, 9.52
Gender			
Males	n = 12 (92.3%)	n = 52 (86.7%)	n = 12 (70.6%)
Females	n = 1 (7.7%)	n = 8 (13.3%)	n = 5 (29.4%)
Site			
Early intervention	n = 11 (84.6%)	n = 49 (81.7%)	n = 12 (70.6%)
Pediatrician	n = 0	n = 8 (13.3%)	n = 5 (29.4%)
Younger sibling	n = 2 (15.4%)	n = 1 (1.7%)	n = 0

mental and psychomotor development. It yields a developmental index score of the child's overall development.

Differential Abilities Scale (DAS; Elliot, 1990) is a standardized assessment instrument that measures general cognitive functioning and generates sub-scores in the areas of nonverbal and verbal abilities.

Procedure

All children at Time 1 were administered the Vineland Adaptive Behavior Scales (VABS; Sparrow et al., 1984) and a cognitive measure, which was either the Bayley Scales of Infant Development, Second Edition (Bayley, 1993), or the Mullen Scales of Early Learning (Mullen, 1995), which replaced the Bayley later in the study. Thirty-nine children were administered the Bayley Scales of Infant Development, and 43 were administered the Mullen Scales of Early Learning. A cognitive measure was not successfully administered to 10 children, because noncompliance was so significant that the results were felt to be invalid. The Childhood Autism Rating Scale (CARS; Schopler et al., 1988) was administered as a measure of symptom severity. In addition to the above-mentioned measures, a full history of the child was obtained during an interview with the parents, part of which was based on DSM-IV criteria for Autistic Disorder (APA, 1994). Children enrolled later in the study received the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 1999; $n = 47$), and the Autism Diagnostic Interview-Revised (ADI-R; Lord et al., 1994; $n = 41$).

At Time 2, children were administered the VABS, the Mullen Scales of Early Learning or the DAS, the ADI, the ADOS, the CARS, and a clinical interview based on DSM-IV criteria for Autistic Disorder.

Families were recruited from one of the three sources listed previously. The caregiver of a child who was between the ages of 16 and 30 months completed the M-CHAT. If the child screened positive on the M-CHAT and subsequent phone follow-up interview, they were invited to receive a free developmental and diagnostic evaluation. All children who received an initial evaluation were invited to receive a follow-up evaluation between the ages of 42–54 months. Some children ($n = 36$) did not receive the follow-up evaluation until after 54 months due to difficulties in contacting and scheduling.

One licensed clinical psychologist or developmental pediatrician and one graduate student performed the evaluation. One of the team members collected history information and completed the caregiver interviews while the other member of the team engaged the child in the cognitive and play-based portion of the evaluation; in almost all cases, everyone stayed in the same room. A blind assessment was not possible at Time 1, as all children presenting for an evaluation had failed the M-CHAT, indicating some developmental concern. Attempts were made to keep the person testing the child at Time 2 blind to prior diagnosis (in some cases, contact between the tester and the parents before the evaluation and information offered spontaneously by the parent made this impossible). Caregivers were provided with verbal feedback on the day of the evaluation and received a written report several weeks later, including appropriate recommendations for intervention services.

Results

First examined was the proportion of children with Autistic Disorder versus PDD-NOS who moved off the spectrum. In the ASD–ASD group, 18% ($n = 11$) were diagnosed with PDD-NOS, and 82% ($n = 49$) were diagnosed with Autistic Disorder at Time 1. In the ASD–NON group, 54% ($n = 7$) were diagnosed with PDD-NOS, and 46% ($n = 6$) were diagnosed with Autistic Disorder at Time 1. Thus, 39% of the children initially diagnosed with PDD-NOS moved off the spectrum, while only 11% of the children diagnosed with Autistic Disorder did so. A chi square analysis showed that there was a significant difference in these proportions ($p = .004$).

Means and standard deviations for the three groups on the main measures are presented in Tables 2 and 3, as are effect sizes of the magnitude of group differences. Independent samples *t*-tests showed significant differences at Time 1 between the ASD–ASD group and the ASD–NON group in VABS Motor Domain ($t(67) = 2.977$, $p = .004$) (effect size = 1.27) scores and Mullen Fine Motor ($t(28) = 2.271$, $p = .031$) (effect size = .99) scores, with the ASD–NON group scoring higher on both scales than the ASD–ASD group. In addition, there was a strong trend toward higher VABS Daily Living Skills ($t(67) = 1.928$, $p = .058$) (effect size = .64) in the ASD–NON group. There were no statistically significant differences in VABS Communication, VABS Socialization, Mullen Visual Reception, Mullen Receptive Language, Mullen Expressive Language, IQ, CARS scores, or total number of DSM-IV symptoms. With the exception of IQ, with a medium effect size of .58, all sizes of these group differences were in the small range. At Time 2, the children in the ASD–NON group were functioning significantly better (at the $p < .05$ significance level) on

Table 2 Mean scores on measures at Time 1

Measure	ASD–NON		ASD–ASD		NON–NON		ES
	M	SD	M	SD	M	SD	
VABS Motor***	88.60	4.19	77.76	11.32	87.35	12.70	1.27
VABS Daily Living*	70.50	6.77	66.25	6.38	75.64	11.96	.64
VABS Communication	65.90	8.35	63.74	6.25	77.47	13.22	.29
VABS Socialization	66.90	11.21	66.27	7.85	81.52	8.80	.06
Mullen VR	29.25	12.73	25.03	7.37	35.75	18.40	.40
Mullen FM**	38.50	15.26	25.73	9.73	31.25	18.13	.99
Mullen RL	23.50	7.00	20.85	3.87	28.00	13.85	.46
Mullen EL	26.00	10.70	23.18	5.81	24.66	5.03	.32
IQ	62.81	18.13	54.66	8.14	72.00	24.63	.58
CARS	33.85	5.17	34.53	5.07	22.40	6.78	.13
Total DSM items failed	6.22	1.56	5.95	2.73	1.40	1.64	.12

Note: VABS = Vineland Adaptive Behavior Scales; VR = Visual Reception; FM = Fine Motor; RL = Receptive Language; EL = Expressive Language; ES = Effect size of difference means between ASD–NON and ASD–ASD groups

*$p < .1$. **$p < .05$. ***$p < .01$. Group difference between ASD–NON and ASD–ASD

Table 3 Mean score on measures at Time 2

Measure	ASD–NON		ASD–ASD		NON–NON		ES
	M	SD	M	SD	M	SD	
VABS Motor*	87.00	10.25	69.17	15.59	86.07	19.73	1.35
VABS Daily Living*	79.80	10.50	58.22	8.78	82.06	21.48	2.23
VABS Communication*	96.00	20.43	63.56	16.12	87.26	16.35	1.76
VABS Socialization *	89.70	8.26	62.50	10.66	84.20	18.74	2.85
Mullen VR*	62.00	6.40	29.54	14.61	50.00	17.38	2.87
Mullen FM*	49.25	18.86	29.97	13.78	41.87	13.29	1.16
Mullen RL*	45.87	12.11	27.22	11.68	37.87	10.64	1.57
Mullen EL*	39.75	9.61	26.77	10.58	40.25	6.92	1.28
IQ*	92.30	33.59	62.54	20.01	90.14	18.82	1.08
CARS*	18.77	3.07	33.17	5.56	18.57	2.67	3.21
Total DSM items failed*	1.62	1.30	6.80	2.11	1.41	1.31	2.96

Note: VABS = Vineland Adaptive Behavior Scales; VR = Visual Reception; FM = Fine Motor; RL = Receptive Language; EL = Expressive Language; ES = Effect size of difference means between ASD–NON and ASD–ASD groups

*$p < .01$. Group difference between ASD–NON and ASD–ASD

all measures than the children in the ASD–ASD group, and there was no difference on any of the measures of functioning compared to the NON–NON group; in fact, the ASD–NON group scored higher (nonsignificantly) on most of the measures.

Figure 1 shows the growth of Vineland Socialization Standard Score for the three groups, as an illustration of how the ASD–NON group starts at the same level as the ASD–ASD group and shows rapid growth, to end in the normal range and comparable to the NON-NON group. Most of the other scores show the same pattern.

A discriminant function analysis (DFA) was performed to determine which variables at Time 1, if any, could predict which children diagnosed with ASD at Time 1 would remain on the autism spectrum at Time 2, and which would move off the spectrum. Three children in the ASD–NON group and seven children in the ASD–ASD group were missing individual data points and therefore could not be included in the DFA. Results were interpreted according to the guidelines suggested by Tatsuoka (1988). The DFA was significant ($p = .05$), with a Wilks' Lambda of .804. Of the 63 cases with ASD at Time 1, 85.7% were correctly classified when diagnosis at time 1 (autistic disorder versus PDD-NOS), IQ, VABS Socialization, VABS Communication, VABS Motor, and VABS Daily Living Skills Domain scores were entered into the analysis. Diagnosis, VABS Motor Domain score, and IQ made a meaningful contribution to the discrimination (even though IQ was not significantly different ($p < .17$) between the ASD–ASD and ASD–NON groups). VABS Communication, Daily Living Skills, and Socialization Domain scores did not contribute to the discrimination between groups. Although the DFA was statistically significant, it did not produce a highly accurate classification of children who remained on the

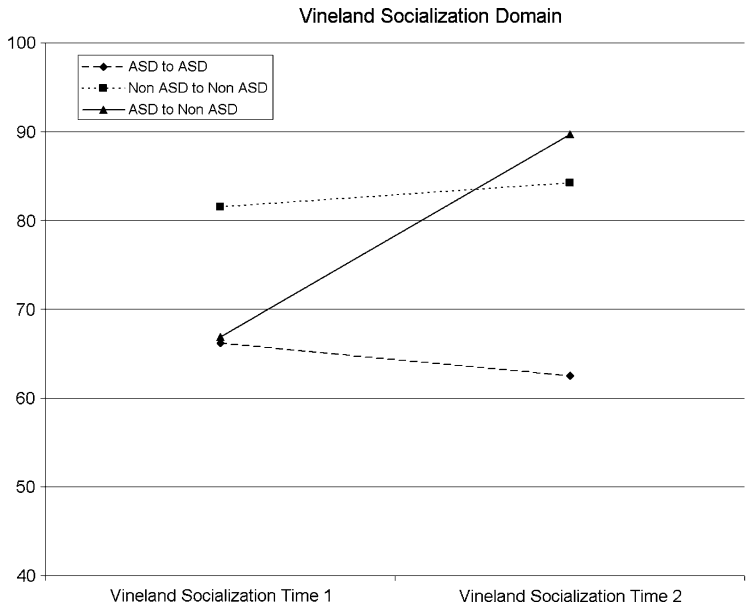

Fig. 1 Vineland Socialization scores for the three groups over time

autism spectrum and those who moved off the spectrum. Only 3 of the 10 children who actually lost their diagnoses were accurately predicted. (See Tables 4 and 5.)

Given the notable *lack* of significant differences on major variables between the children who moved off the spectrum and those who did not, individual items from the Vineland and DSM-IV autism symptoms were examined, in exploratory fashion, to determine whether any particular type of skill or symptom differentiated the groups. A chi-square test was used to examine whether individual items within the VABS domains were unevenly distributed between the ASD–NON and ASD–ASD groups. Items were recoded so that a score of ''1'' or ''2'' represented one construct of ''sometimes or usually'' demonstrates the skill, whereas a score of ''0'' continued to mean that the child never demonstrated the skill. Items within the range of functioning of most of the children were examined. In other words, only those items between the basal and the ceiling level for children at Time 1, demonstrating variability, were examined. Twenty-one items within the Communication Domain, 22 items within the Daily Living Skills Domain, 20 items within the Socialization Domain, and 20 items within the Motor Skills Domain at Time 1 were analyzed. Within the Communication Domain, children who eventually moved off the autism spectrum were significantly more likely to listen to a story for at least 5 min ($p = .014$). Within the Daily Living Skills Domain, children who moved off the spectrum were more likely to be able to bathe themselves with some assistance (as compared to needing full assistance; $p = .027$) and help with extra chores when asked ($p = .002$). Within the Socialization Domain, children who moved off the autism spectrum were more likely to show a desire to please their caregiver ($p = .038$). Within the Motor Skills Domain, children who moved off the spectrum were more likely to run smoothly with changes in speed and direction ($p = .048$), open doors by turning and pulling the doorknob ($p = .005$), and pedal a tricycle or other three-wheeled vehicle for at least six feet ($p = .041$). It is striking that of 95 comparisons, there were so few significant at $p < .05$, and were a more stringent alpha level of .01 to be used to guard against Type I errors, only two items (opening

Table 4 Standard canonical discriminant function coefficients

	Function 1
IQ	.304
VABS Communication Domain	–.104
VABS Daily Living Skills Domain	–.025
VABS Socialization Domain	–.023
VABS Motor Domain	.796
Time 1 Diagnosis	.364

Wilks' Lambda = .804, $p = .05$

Note: VABS = Vineland Adaptive Behavior Scales

Table 5 Classification matrix for discriminant function analysis of the IQ and Vineland domains

Predicted group membership			
Group	ASD–NON	ASD–ASD	Total
ASD–NON	3	7	10
ASD–ASD	2	51	53

Note: Of the original cases, 69.8% were correctly classified

doors by turning doorknobs, and helping with chores) remained significant.

A chi-square test was used to look also at the presence or absence of each DSM-IV criterion; there was no difference between the ASD–NON and ASD–ASD groups at Time 1 on any individual criterion.

Discussion

In this study, children who receive an early diagnosis of an ASD and eventually achieve a truly optimal outcome, or "recover," from the ASD diagnosis, display a unique developmental trajectory. In most areas of functioning, when they are 2 years of age they are functioning at the level of the children diagnosed with an ASD who remain on the spectrum. However, by 4 years of age, these children who no longer have a diagnosis of an ASD are functioning at least at the same level as the children who were never diagnosed with an ASD (who generally had a language or global developmental delay), and in most cases are functioning within the average range on standardized tests. When the children who eventually achieved a non-autistic outcome were 2 years old, they were just as impaired as the children who did not recover in terms of socialization skills, communication skills, language ability, individual symptoms and overall symptom severity, which makes optimal outcome extremely difficult to predict at a young age.

Interestingly, there were no differences in symptom severity between the two groups at Time 1, as measured by CARS score or number of DSM symptoms checked. This suggests the idea that the optimal outcome group was accurately diagnosed at Time 1, and that the children who reach an optimal outcome are not merely on the cusp of an ASD diagnosis when diagnosed initially. This is consistent with previous studies in which predictive validity of autistic symptoms and symptom severity has been a weak predictor of outcome (Fein et al., 1999; Stevens et al., 2000; Szatmari et al., 2003). However, the children with an initial PDD-NOS diagnosis were more likely (39%) to move off the spectrum than those with the full picture of Autistic Disorder (11%). How is this consistent with lack of differences in number of DSM symptoms or CARS score? At this early age, the appearance of some symptoms (e.g., scripted language) may actually be a positive sign, in that only children above a certain developmental level can display that symptom. Therefore, the total number of DSM symptoms may be a weak marker of severity at this age. However, the CARS severity score as well as the Vineland Socialization score were not different between the ASD–NON and ASD–ASD groups, suggesting that the severity of social and communication impairment is not in fact a good marker of the potential for recovery. Most of the children who received a PDD-NOS rather than Autistic Disorder diagnosis did so because of a lack of symptoms in the repetitive behavior domain, which symptoms often do not appear until later in development (Stone et al., 1999). It seems that although overall severity of symptomatology, especially in social interaction and communication, is a poor predictor of the potential for recovery, showing the full picture of autism, including repetitive behavior or resistance to change, may be relatively more predictive. However, it should be noted that half of the children who lost their diagnosis did in fact demonstrate some repetitive behavior at Time 1. Firm conclusions about this issue must await studies of larger samples of children.

There were other significant differences between children who were diagnosed with an ASD at the age of 2 years and no longer met criteria for an ASD diagnosis at the age of 4 years as compared to children who continued to meet criteria for an ASD diagnosis, specifically in motor skills (both by parent report and direct testing) and overall cognitive level (although overall cognitive level was not significantly different between the two groups, it did marginally contribute to the DFA). Although motor skills and cognitive level are not diagnostic of an ASD diagnosis, it seems to be the case that children with these skills intact, especially motor skills, are more likely to achieve an optimal outcome. However, there are some children who have intact motor functioning and do not achieve an optimal outcome. In addition, there are a few children who do not have IQ or motor skills in the normal range, yet still reach an optimal outcome. Therefore, adequate cognitive and motor skills are perhaps signs of positive prognosis but appear to be neither necessary nor sufficient for optimal outcome.

When looking in detail at the children who achieved an optimal outcome, they seem to be more likely than the children who remain on the spectrum to have some specific skills in various areas of adaptive functioning. Skills such as running smoothly, opening doors, pedaling a tricycle, using scissors, feeding themselves, and bathing with assistance seem to be activities that foster independence and a sense of self-efficacy. Children who later reach optimal outcome might be more motivated at an early age to master specific skills that lead to independence. Mastery of these skills might also reflect higher cognitive ability, or an ability to maximize the benefits of early intervention. It should

be noted that daily living skills (which was approaching significance between the groups) are also heavily dependent on motor ability, and it might be that differences in motor skill underlie some of the differences in daily living skills. Differences in daily living abilities may also reflect parental demands, and perhaps parents who insist on the children learning self-help skills are also more proactive in other areas such as finding effective treatments or continuing to maintain a therapeutic environment within the home.

Skills such as attending to a story and following instructions, which were more likely to be mastered at a young age by the children who reached an optimal outcome, indicate these children may be more likely to attend to language than those who do not achieve an optimal outcome. This interest in and attention to language at such a young age may be necessary for eventual recovery from autism. Perhaps the children who eventually achieve an optimal outcome are more interested in social interactions, yet require structured teaching in order to engage socially with their surroundings.

Previous studies indicate that individual characteristics, such as initial language, communicative attempts, and response to joint attention, predict positive outcome (Bono et al., 2004; Charman et al., 2005; Sallows & Graupner, 2005; Sigman & McGovern, 2005). Children must attend to language before they are able to engage in higher-level communicative interactions. It is therefore likely that attention to language, when exhibited by young children with an ASD, would predict a positive outcome, and perhaps the possibility of an optimal outcome.

Although findings from the present study indicate that differences in a few specific areas do exist between the ASD–NON and ASD–ASD groups, the general conclusion of our study is that these children are still very difficult to differentiate when diagnosed initially. It is probable that there are discoverable, but currently unidentified differences in behavioral development, cognitive development, brain structure, or genetic vulnerability that will eventually allow more accurate prediction of which children may be able to move off the ASD spectrum. Since it is currently difficult to predict which children will benefit most from early intervention, with possibly even the potential for recovery, it is crucial to provide all children intensive, early intervention.

There are limitations to the present study. First, recovery is a relatively low base rate phenomenon. Therefore, the sample of children who lost their diagnosis of an ASD is small, and caution should be taken when generalizing these results to the broader population. Additionally, the power of the current sample size is relatively low. It is possible that with a larger sample there would be more clear predictors. With our sample of 13 ASD–NON and 60 ASD–ASD children, we had a .70 probability of detecting a medium effect size; it is quite possible, therefore, that a larger sample of optimal outcome children would result in other variables, such as IQ, being significant predictors. Second, the children in this study all received early intervention. Although it was recommended that all children with a diagnosis of an ASD receive structured, behaviorally-based intervention that was specifically tailored for each child based on his/her unique strengths and weaknesses, the interventions received were of varying amount, intensity, type, and quality. This information was provided by parents at the time 2 evaluation and will be analyzed in future papers to try to determine the impact of intensity, type, and setting of service on outcome. The relationship between these intervention variables and initial child and family characteristics is likely to be very complex. In addition, conclusions will have to be guarded, since the intervention data were gathered post hoc, no random assignment of children to intervention conditions was attempted, and provision of services was heavily biased by the availability of services in each geographic region of the state.

Children with an ASD probably display varied etiologies. Therefore, research ultimately should explore whether certain pathophysiologies limit the effects of intervention.

Finally, age 4 obviously does not constitute a final developmental outcome. It will be necessary to follow the ASD cohort to see if additional children move off the spectrum, and to determine what happens in future years to the children who move off the spectrum by age 4.

Acknowledgments This research was supported in part by grants from the National Institute of Child Health and Development (5 R01 HD39961), the Maternal and Child Health Bureau (5 R40 MC00270), the National Association for Autism Research, and the UCONN Research Foundation. In addition, the authors would like to thank the many graduate students and undergraduate research assistants who have assisted with this project, and the children and their families who have participated in this project.

References

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Baranek, G. T. (1999). Autism during infancy: A retrospective video analysis of sensory-motor and social behaviors at 9–12 months of age. *Journal of Autism and Developmental Disorders, 29*(3), 213–224.

Bayley, N. (1993). *The Bayley scales of infant development* (2nd ed.). San Antonio, TX: Harcourt Brace.

Bono, M. A., Daley, T., & Sigman, M. (2004). Relations among joint attention, amount of intervention and language gain in autism. *Journal of Autism and Developmental Disorders, 34*, 494–505.

Cox, A., Klein, K., Charman, T., Baird, G., Baron-Cohen, S., Swettenham, J., Drew, A., & Wheelwright, S. (1999). Autism spectrum disorders at 20 and 42 months of age: Stability of clinical and ADI-R diagnosis. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 40*, 719–732.

Charman, T., Taylor, E., Drew, A., Cockerill, H., Brown, J., & Baird, G. (2005). 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, 46*, 500–513.

Dawson, G., & Osterling, J. (1997). Early intervention in autism: Effectiveness, common elements of current approaches. In M. J. Guralnick (Ed.), *The effectiveness of early intervention: Second generation research*. Baltimore, MD: Brookes.

Eaves, L., & Ho, H. (2004). Brief report: stability and change in cognitive and behavioral characteristics of autism through childhood. *Journal of Autism and Developmental Disorders 26*, 557–569.

Elliot, C. D. (1990). *Differential ability scales*. Orlando, Florida: The Psychological Corporation.

Fein, D., Dixon, P., Paul, J., & Levin, H. (2005) Pervasive developmental disorder can evolve into ADHD: Case illustrations. *Journal of Autism and Developmental Disorders, 35*, 525–534.

Fein, D., Stevens, M., Dunn, M., Waterhouse, L., Allen, D., Rapin, I., & Feinstein, C. (1999). Subtypes of pervasive developmental disorder: Clinical characteristics. *Child Neurology, 5*, 1–23.

Gabriels, R. L., Hill, D. E., Pierce, R. A., Rogers, S. J., & Wehner, B. (2001). Predictors of treatment outcome in young children with autism: A retrospective study. *Autism, 5*, 407–429.

Kelley, E., Paul, J., Fein, D., & Naigles, L. (2006). Residual language deficits in optimal outcome children with a history of autism. *Journal of Autism and Developmental Disorders, 36*, 807–828.

Koegel, L. K., Koegel, R. L., Shoshan, Y., & McNerney, E. K. (1999). Pivotal response intervention II: Preliminary long-term outcomes data. *Journal of the Association for Persons with Severe Handicaps, 24*, 186–198.

Koegel, R. L., Koegel, L. K., & McNerney, E. K. (2001). Pivotal areas in intervention for autism. *Journal of Clinical Child Psychology, 30*, 19–32.

Lord, C. (1995). Follow-up of two-year-olds referred for possible autism. *Journal of Child and Adolescent Psychiatry, 36*, 1365–1382.

Lord, C., Rutter, M., & Le Couteur, A. (1994). 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, 24*, 659–685.

Lord, C., Rutter, M., DiLavore, P., & Risi, S. (1999). *Autism diagnostic observation schedule-WPS edition*. Los Angles, California: Western Psychological Services.

Lord, C., Storoschuk, S., Rutter, M., & Pickles, A. (1993). Using the ADI-R to diagnose autism in preschool children. *Infant Mental Health Journal, 14*(3), 234–252.

Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Counseling and Clinical Psychology, 55*, 3–9.

McEachin, J., Smith, T., & Lovaas, I. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal of Mental Retardation, 97*, 359–372.

Moore, V., & Goodson, S. (2003). How well does early diagnosis of autism stand the test of time? Follow-up study of children assessed for autism at age 2 and development of an early diagnostic service. *Autism, 7*, 47–63.

Mullen, E. (1995). *Mullen scales of early learning*. Circle Pines, MN: American Guidance Service.

Osterling, J., & Dawson, G. (1994). Early recognition of children with autism: A study of first birthday home videotapes. *Journal of Autism and Developmental Disorders, 24*(3), 247–257.

Robins, D., Fein, D., Barton, M., & Green, J. (2001). The Modified Checklist for Autism in Toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 31*, 131–144.

Rogers, S. (2001). Diagnosis of autism before the age of 3. *International Review of Research in Mental Retardation, 23*, 1–31.

Rogers S. J., Hepburn S. L., Stackhouse T., & Wehner E. (2003). Imitation performance in toddlers with autism and those with other developmental disorders. *Journal of Child Psychology and Psychiatry, 44*, 763–781.

Sallows, G. O., & Graupner, T. D. (2005). Intensive behavioral treatment for children with autism: Four-year outcome and predictors. *American Journal on Mental Retardation, 6*, 417–438.

Schopler, E., Reichler, R., & Renner, B. (1988). *The childhood autism rating scale*. Los Angeles: Western Psychological Services.

Sigman, M., & Ruskin, E. (1999). Continuity and change in the social competence of children with autism, down syndrome, and developmental delays. *Monographs of the Society for Research in Child Development*.

Sigman, M., & McGovern, C. W. (2005). Improvement in cognitive and language skills from preschool to adolescence in autism. *Journal of Autism and Developmental Disorders, 35*, 15–23.

Sparrow S., Balla, D., & Cicchetti, D. (1984). *The vineland adaptive behavior scales*. Circles Pines, MN: American Guidance Services.

Stevens, M. C., Fein, D. A., Dunn, M., Allen, D., Waterhouse, L. H., Feinstein, C., & Rapin, I. (2000). Subgroups of children with autism by cluster analysis: A longitudinal examination. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 346–352.

Stone, W. L., & Hogan, K. L. (1993). A structured parent interview for identifying young children with autism. *Journal of Autism and Developmental Disorders, 23*(4), 639–652.

Stone, W. L., Lee, E. B., Ashford, L., Brissie, J., Hepburn, S. L., Coonrod, E. E., & Weiss, B. H. (1999). Can autism be diagnosed accurately in children under 3 years? *Journal of Child Psychology and Psychiatry and Allied Disciplines, 40*(2), 219–226.

Szatmari, P., Bryson, S. E., Boyle, M. H., Streiner, D. L., & Duku, E. (2003). Predictors of outcome among high functioning children with autism and Asperger syndrome. *Journal of Child Psychology and Psychiatry, 44*, 520–528.

Tatsuoka, M. M. (1988). *Multivaraite analysis* (2nd ed.). NY: Macmillan.

Volkmar, F., Chawarska, K., & Klin, A. (2005). Autism in infancy and early childhood. *Annual Review of Psychology, 56*, 315–336.