J Autism Dev Disord (2014) 44:1357–1372
DOI 10.1007/s10803-013-1997-2

ORIGINAL PAPER

School-Aged Functioning of Children Diagnosed with Autism Spectrum Disorder Before Age Three: Parent-Reported Diagnostic, Adaptive, Medication, and School Placement Outcomes

Patricia O. Towle · Karyn Vacanti-Shova · Shristi Shah · Ann Higgins-D'alessandro

Published online: 18 December 2013

Abstract Eighty children with early autism spectrum disorder (ASD) diagnoses (under 36 months) were identified using a chart abstraction protocol applied to early intervention charts. Parents filled out questionnaires by mail when the children were school-aged (ages 6–16 years). Similar to previous studies, approximately 20 % no longer had ASD diagnoses; the other participants were assigned to Moderate/Severe versus Mild ASD outcome groups. These three groups were compared across several variables, including diagnostic features and functional features including adaptive behavior, social experiences, medication use, and school placement. The findings expand our knowledge about outcomes in longitudinal studies of children with ASD, as well as provide support for using relatively indirect methods (chart review, parent questionnaire) to gather this type of information.

Keywords Autism · ASD · Early diagnosis · Longitudinal · School-age · School placement · Adaptive behavior · Medication use

P. O. Towle (✉) · K. Vacanti-Shova · S. Shah
Westchester Institute for Human Development, Cedarwood Hall, Valhalla, NY 10538, USA
e-mail: ptowle@wihd.org

P. O. Towle
School of Health Science and Practice, New York Medical College, Valhalla, NY, USA

K. Vacanti-Shova · A. Higgins-D'alessandro
Department of Psychology, Fordham University, Bronx, NY, USA

Present Address:
K. Vacanti-Shova
Stepping Stones Learning Center, Rochester, NY, USA

Present Address:
S. Shah
Department of Psychiatry, Pine Rest Christian Mental Health Services, Grand Rapids, MI, USA

Introduction

The current strong interest in early detection of autism spectrum disorders in young children follows directly from efforts to provide intervention services as early as possible (Boyd et al. 2010). A valid question, however, is whether accurate diagnoses can be made in infants and toddlers. A body of literature has grown up around examining reliability and validity of early diagnoses through demonstration of diagnostic stability from early to later ages.

The studies as a whole are remarkably consistent in demonstrating that, of children diagnosed with an autism spectrum disorder before the age of three, the great majority will remain on the autism spectrum after a follow-up interval for two or more years. Woolfenden et al. (2012) examined the studies from a context of methodological rigor and made a best estimate of 85–89 % of children remaining on the spectrum. This and other reviews (Helt et al. 2008; Kleinman et al. 2008; Rondeau et al. 2011) note that there is consistency in subgroups of subjects/participants whose diagnostic stability is somewhat lower, specifically children with milder early diagnoses [Pervasive Development Disorder-Not Otherwise Specified (PDD-NOS) vs. Autistic Disorder] and children who were among the youngest when first diagnosed (under 24 months of age).

That fact that these studies are a heterogeneous group both helps and hinders conclusions. Such consistent outcomes given the various approaches, samples, diagnostic strategy, focus of inquiry, and follow-up interval does strengthen the conclusion that early ASD diagnoses are in fact stable, and compels researchers to understand why about 20 % go "off the spectrum." On the other hand, examining in detail the various study characteristics reveals areas of weakness in the studies as a body. Woolfenden et al. (2012) raise the perennial issues of small sample size and subject heterogeneity (especially with respect to cognitive level).

One consideration *not* raised in these reviews, however, is that a number of the studies included children diagnosed between the ages of 3 and 5, as well as early school age—yet an ASD diagnosis after age three would not be considered an early diagnosis by current standards. When considering the extant studies that include only children younger than three, their numbers are considerably reduced.

On the other end of the age issue, the studies vary significantly with respect to how much later children are examined. The majority follows their subjects for 1–3 years. It can be argued, however, that longer periods of time provide more definitive information about outcome of a disorder that has such great variability in presentation. For example, distinguishing among milder forms of ASD (specifically, Asperger syndrome, PDD-NOS, and high functioning autism—which is now a moot point since new diagnostic criteria are in effect that eliminate subtypes per se), but also the array of language-based learning disability, nonverbal learning disability, social-communication disorder, and attentional and anxiety disorders, usually requires the child to be somewhat older—at least fully out of the preschool period. For example, the Centers for Disease Control (CDC) targets 8-year-old children for their epidemiologic studies.

Table 1 presents the existing longitudinal studies in terms of their Time 1 (age of first diagnosis) and Time 2 (age of follow-up diagnosis) age points. Here it can be seen that only four studies actually start before age three and extend into school age. While it is very useful to know that children can change the severity of their diagnostic picture and move off the spectrum as early as 3 and 4 years of age, longer periods of follow-up yield more detailed and practical prognostic information.

The majority of these studies focus almost exclusively on diagnostic outcome, meaning whether a diagnosis is present or not, and/or whether the child has moved from one type of diagnosis to another within the spectrum. However, this approach has limits in terms of understanding of prognosis more fully. First, it is well known that the "spectrum" of autism includes great variability in functioning in terms of communication ability, adaptive and academic functioning, and need for support across the lifespan. It is important, therefore, to broaden the view of "outcomes" past diagnosis alone to that which includes a greater variety of meaningful behaviors.

It is likely that there are more studies of shorter duration and of more narrow outcome focus because of the cost and effort involved in gathering extensive, longer-term data. Following subjects over a long period of time requires considerable resources, since families may have moved, and/or their motivation to stay involved with the research project may have diminished. A related source of variability in studies that has a direct impact on resources required is the strategy used to recruit and examine subjects, and these strategies run along a continuum of how directly versus indirectly individuals are examined. On the one end of the continuum is the very resource-heavy strategy of seeing children and families at the clinic and laboratory, using time-intensive tasks, standardized instruments and observations, and clinical examinations by more than one clinician to ensure reliability. In these studies, sample size tends to be modest. On the other end of the continuum, internet-based surveys (e.g., Rosenberg et al. 2011) and phone surveys (Kogan et al. 2009) require much less ongoing cost to the researcher and result in much

Table 1 Longitudinal studies of ASD diagnostic stability grouped by Time 1 to Time 2 ages

>3 years with f/u to grade school (n = 6)	≤3 years with 1–2 years f/u (n = 7)	≤3 years with f/u at age 4–5 years (n = 3)	<3 years with f/u to grade school (n = 4)
Moore and Goodson (2003)	Gillberg et al. (1990)	Szatmari et al.(2000)	Michelotti et al. (2002)
Eaves and Ho (2004)	Lord (1995)	Moss et al. (2008)	Charman et al. (2005)
Scambler et al. (2006)	Cox et al. (1999)	van Daalen et al. (2009)	Turner et al. (2006)
Sutera et al. (2007)	Stone et al. (1999)		Lord et al. (2006)
Turner and Stone (2007)	Jonsdottir et al. (2007)		
Kleinman et al. (2008)	Chawarska et al. (2009)		
	Malhi and Singhi (2011)		

f/u follow-up

larger sample sizes because of ease of participation; however, less control over procedural features (e.g., objective measurement checks on the veridicality of parent report) may impact reliability and validity of measurement.

In between these two ends of the continuum are studies that use archival materials: medical and educational charts and records. In this way children are not examined directly, but medical, clinical, and academic records may represent them very well depending on the comprehensiveness of the material in the charts. Perhaps the most prominent group of studies utilizing chart review is that conducted by the CDC to determine prevalence of ASD in 8-year-olds (CDC 2007, 2009, 2012). The investigators developed a chart abstraction method to determine presence of ASD using clinic charts and school records. Avchen et al. (2011) examined the validity of this method, which they refer to as *multisource records-review of health and education records*. The researchers used a detailed chart abstraction guideline based on symptoms as delineated in demonstrably valid ASD diagnostic instruments, as applied to a variety of health and educational documents in the records of a cohort of 8-year-olds in a given catchment area. In addition to noting explicit diagnoses applied at any time in the child's development, they also documented all behavioral descriptions that support the diagnosis of an autism spectrum disorder. They then recruited a sample of children who were presumed to have ASD based on the record review method in addition to a sample presumed to not have ASD and conducted direct clinical evaluations on 177 children. The results showed estimates of specificity, positive predictive value, and negative predictive value to be fairly high (.96, .79, and .91, respectively), while sensitivity was lower (.60). The authors concluded that these rates were acceptable for surveillance purposes. In another study, which examined the utility of general practitioner records to diagnose ASD in the UK, Fombonne (2002) demonstrated good inter-rater reliability and high positive predictive value of autism symptoms noted in the developmental record for eventual confirmation of pervasive developmental disorder from consultant letters.

Chart abstraction for case finding or confirmation of early diagnosis has begun to be applied reliably to early intervention charts as well. Towle et al. (2009) used the chart abstraction method with public early intervention records to establish ASD diagnoses for very young children. In addition to explicit diagnoses recorded, the investigators also used documented DSM-IV-TR symptoms from multiple sources (evaluation reports, service coordination notes, documents from other community resources, and provider progress notes), attendance in programs specifically for ASD, and patterns and type of services (e.g., applied behavior analysis) as coding criteria. They demonstrated that this approach could be used to reliably determine the presence of ASD, and that this method of case ascertainment held promise as a research strategy for epidemiologic as well as other types of studies. A recent study that used examination of early intervention records to confirm early presence of ASD also demonstrated reliability in using the charts to determine early diagnosis (Fein et al. 2013). Finally, Wiggins et al. (2013) determined that adding early intervention chart review to school-aged records was useful for case ascertainment in epidemiologic studies of ASD.

The present study expands the current literature on longitudinal examination of children determined to have an ASD from a young age, with three major goals. The first is to replicate previous studies in establishing a rate of diagnostic stability for school-aged children diagnosed with ASD early (specifically before age three, with many subjects before age two), thus contributing information about these children starting at an earlier age and over a greater period of time than is typical for similar studies. The second goal is to expand the type of information about the children's outcome from diagnostic (including community diagnoses) to include a number of additional features to examine specific developmental and behavioral functioning, adaptive and social functioning, medication use, and school placement. A final, overarching goal is methodological: to examine the utility of relatively indirect methods (chart review at Time 1 and parent questionnaire at Time 2) to gather this type of information.

Methods

Participants and Setting

Participants were 80 children who were identified with ASD before the age of 3 years (N = 80) and whose parents provided information about their functioning at grade school age (ages 6–16 years), as well as experiences their children had in between (N = 57). All families were involved with a University Center for Excellence in Developmental Disabilities in a large county outside of New York City. The children's charts were from two different sources within the center: the Early Childhood Services program, wherein multidisciplinary and supplemental evaluations are performed through contract with the Department of Health's public early intervention program and a Service Coordination service, wherein families with children receiving early intervention services are provided guidance through their time in the early intervention system before transitioning to the school district system for the ages of 3–5 years.

This study was approved by the New York Medical College's Institutional Review Board Committee on Human Subjects.

Early Intervention Charts

The charts maintained in these programs have extensive sets of reports and records that constitute a rich repository of developmental and behavioral information about the children. They are created as part of the public early intervention system guided by federal law such that each state and municipality follows the same procedures regarding evaluations, service procedures, and documentation. Specifically, each child enters the system by receiving a multidisciplinary evaluation that must examine the five domains of cognition, communication, social-emotional functioning, daily living skills/adaptive behavior, and physical functioning (including gross and fine motor skills). The evaluation results need to be detailed enough to determine if the child qualifies for the early intervention program (in other words, generating standardized scores and/or developmental age equivalents) and provide enough functional description to create an initial intervention plan. If the child receives services, then the record contains systematic information about services received as well as quarterly progress notes, which provide goals, behavioral descriptions, and developmental age levels, by multiple service providers.

Although the same basic assessment and reporting strategies were used for each child, there is variability in the specific instruments used by clinicians performing the evaluations. The evaluators themselves must meet qualification set forth by the federal and state guidelines in terms of discipline, certification, and licensure. They are also instructed to use standardized tests if possible, but can complement this with criterion-referenced and curriculum-based instruments, as well as "professional judgment" in the case where a child's behaviors rule out standardized assessment. Therefore, across the charts examined, a variety of different instruments were used to measure cognition, communication, adaptive skills, and autism symptoms.

Instruments and Measurements

Measurements Obtained from the Early Intervention Charts (Time 1)

Autism Spectrum Disorder Decision-Making Protocol for EI Charts (ASD-DMP, Towle et al. 2009). A coding system to determine the presence of ASD from the materials in EI charts was used; this is described in Towle et al. (2009) but is summarized briefly here.

The coding system is based on the DSM-IV-TR criteria (American Psychiatric Association (APA) 2000). It translates easily, however, to DSM-V criteria (APA 2013) because the behaviors that are social or communication-based in nature now fit under the Social-Communication symptom domain. It is necessary for the coder to be familiar with local programs, so that if a child attended a particular program that was specifically for children with ASD, this figured heavily in the decision-making process. The coding categories are as follows:

0—No ASD present (no evidence or insufficient evidence for ASD);
1—ASD Possible (there is some evidence to suspect the child might have ASD, but the evidence is not quite sufficient either because the chart was too thin or because of a lack of the details in the behavioral evidence);
2—ASD Highly Likely (there is sufficient to abundant descriptive, behavioral, and treatment evidence to support the diagnosis, but no diagnosis *per se* appears in the chart);
3—ASD Confirmed (a diagnosis is given by a qualified clinician, and there is sufficient or better evidence from behavioral descriptions to support this).

Only charts coded as "2" or "3" were used for the participant pool. As described in Towle et al. (2009) interrater agreement for the ASD-DMP was good, with a weighted kappa of .86 (95 % CI .33–.94; $p < .0001$).

''Sufficient evidence'' consisted of converging information, pinpointing symptoms from all three DSM symptom domains, from several sources in the chart—background information from evaluation reports; stated concerns of caregivers, teachers, or service providers; the body of evaluation reports; provider progress notes; and Individualized Family Service Plan descriptions and goals. If a child had an independent diagnosis of ASD, or received one as a result of the early intervention evaluation, this information would be in these sources. Behavioral descriptions noted included: descriptions of or concerns about eye contact, social relatedness and interaction, self-stimulatory behaviors, atypical language or language delay with significant pragmatic delays, delayed and/or unusual play patterns, and absence of many descriptions that contraindicated ASD.

Another source of evidence was the services that the child had been receiving or subsequently received after the evaluation. Services that are typical of children with ASD are: (a) several or many hours of Applied Behavior Analysis, or other intervention associated with ASD such as FloorTime™, (b) services from many different disciplines combined with above, (c) attendance in programs that have a high proportion of children with ASD.

There was no attempt to distinguish Autistic Disorder from milder presentations, such as PDD-NOS. In a few charts, Asperger syndrome was suspected or designated at Time 1 but not differentiated for this study and was included as ASD.

Child Characteristics: Gender and Age. The chronological age of the child at Time 1 was determined as follows. We reasoned it was most meaningful for the purposes of the study to determine the age at which the child started services that were associated with the recognition of the presence of an ASD. This was not always when the child first entered the early intervention system, since a few started very young (e.g. 4 months) due to early motor delays and others started speech services but after several months were re-evaluated for ASD and then began to receive much heavier services. Nor was the age the child received a formal diagnosis chosen, because this sometimes did not occur at all, and sometimes it was confirmed well after the child began to receive services for autism symptoms. Therefore, a chronology of evaluation and service was determined for each child, and the age that the evaluation occurred that resulted in autism-specific services was used for the most part. In a few instances, in the absence of a definitive evaluation for this purpose, service coordination and other notes were used to estimate a date when planning or implementing such services occurred that clearly signaled that the providers recognized the child to have ASD.

Vineland Adaptive Behavior Scales (VABS; Sparrow et al. 1984). The Vineland is a standardized, structured parent interview of the child's everyday functioning. It yields domain scores in the areas of communication, daily living skills, socialization, and motor development, and an overall Adaptive Behavior Composite (ABC). Items are scored on a scale of 2 (yes, usually), 1 (sometimes or partially), and 0 (No, never). Sixty-seven percent of the charts had VABS scores in them, although not all gave the subdomain scores in addition to the ABC; a few gave subdomain scores and not the ABC.

Childhood Autism Rating Scale (CARS; Schopler et al. 1988). The CARS is an objective, judgment-based (with behavioral anchors) rating system used for diagnostic purposes and to assess changes in symptom severity. It is comprised of 15 items on which a child receives a score ranging from 1 (age appropriate with no abnormality) to 4 (severely abnormal behavior for that age). The items are related to one of the following areas: relationships with others; imitation; emotional expression; body use; peculiarity in object use; resistance to change; visual, auditory, and tactile responsiveness; anxiety; verbal and nonverbal communication; activity level; and intellectual ability. The final item is the rater's overall evaluation of severity of autism symptoms. A global score is obtained by summing the item scores. Scores less than 29.5 are considered "non-autistic," scores between 30 and 36.5 are considered "mild to moderately autistic," and scores 37 and above are considered "severely autistic." The CARS had been used in 56 % of the Time 1 evaluations.

Measurements Obtained from Parent-Completed Questionnaires at School-Age (Time 2)

An extensive parent-completed Background Questionnaire had the following components:

1. *Demographics* Parents provided their dates of birth, highest educational degrees obtained, occupations, and place of residence.
2. *Current ASD diagnosis/es* Parent was asked: if the child had received an ASD diagnosis when young (and if so, when), if they agreed with it, which diagnosis(es) had been assigned to the child, what were the current (Time 2) diagnoses, and if the parent considered their child to be "on the spectrum" currently. In actuality, very few children had received recent diagnostic evaluations, so current diagnostic status was derived through a procedure described below.
3. *Child medical history and diagnoses* Any positive results for genetics or neurologic consultations, medical diagnoses, and any additional neurodevelopmental/psychiatric diagnoses such as ASD, Attention Deficit Hyperactivity Disorder (ADHD), anxiety, or learning disability.
4. *Medication status* Parents were asked to report on what medicines their child was currently taken, and for what type of problem.
5. *Current Behavioral, Physical, and Social Functioning Questionnaire* This questionnaire, parts I–IV, was created for the purposes of this study since no published instruments were deemed to fulfill the purposes of the questions posed.
    (a) *Symptom Checklist* A checklist of items related to the three symptom domains from the DSM-IV criteria had a "yes" versus "no" endorsement. This checklist was intended to tap into school-aged behaviors that may be relevant to children with milder presentations of ASD, and as a check on the diagnostic information the parent might provide and the Gilliam Autism Rating Scale (GARS) scores (see below). Examples from the Socialization domain were: Outgoing (reversed item); Awkward around others; Mixes easily with other children (reversed); Plays well only if he/she chooses the activity; Doesn't always look you in the eye when s/he should; Is very affectionate (reversed), Is more comfortable with adults than

same-aged peers; Can be flexible when playing and lets others take the lead (reversed). Examples from the Communication Domain: Can only talk about one or two favorite subjects; Speaks out when not appropriate; Is quite good at holding a conversation and including the other person in it (reversed); Talks too much about favorite subject; Has trouble carrying on a two-sided conversation. Examples from the Repetitive Behaviors Domain: Gets obsessed with topics or ideas; More interested in things than people; Has some odd habits, such as twirling hair, spinning things, repetitive physical movements, staring at lights or out the window, arranging things in a pattern.

(b) *Problem Checklist* Parents were asked to check off the areas in which their child still experienced challenges. They were presented with the following list: learning, language, attention, non-compliance or other challenging behavior, and social problems. They were asked: In which of these areas does your child still have problems?

(c) *Physical and Social Functioning* Parent endorsed a general rating of physical coordination and involvement in team sports. A general rating of social functioning was requested (Good, fair, poor); as well, three specific social functioning questions were posed regarding number of close friends, and, compared to other children his/her age, the extent to which the child participates in birthday parties and sleepovers.

(d) *Current school placement and services received* This part of the questionnaire inquired about current grade and school placement, specifying the type of classroom (regular classroom in a public school, regular classroom with services. Integrated or team-taught classroom, special classroom in a public school, segregated school building for special education, private special education school, and residential placement). In addition, the general degree of support was characterized by asking if their child had an aide in school, a behavior plan, social skills group, or received 504 accommodations.

Gilliam Autism Rating Scale (GARS, Gilliam 1995). The GARS contains 56 items and is divided into four subscales of Stereotyped Behaviors, Communication, Social Interaction, and Developmental Disturbance. A 4-point Likert scale ranging from 0 (Never Observed) to 3 (Frequently Observed) is used to rate the items on the first three subscales. A dichotomous scale (yes or no) is used to score the items on the Developmental Disturbance subscale, which addresses behaviors and milestones in the first 36 months of life. The Developmental Disturbance subscale is a retrospective parent report of early autism symptoms. Each subscale can be computed into a standard score and then tabulated for a total score, the autism quotient (AQ). The AQ has an average of 100 and a standard deviation of 15. It provides a measure of the likelihood that a child has autism. A score of 100 indicates that a child has symptoms similar to the average child with autism and a lower score indicates fewer symptoms than the average child with autism.

Vineland Adaptive Behavior Scales, Second Edition (VABS-II, Sparrow et al. 2005). The Vineland-II differs from the Vineland-I in several ways. First, new items have been added to improve measurement for very young children. Items have been added to the Communication domain that assesses the development of spoken language and the ability to initiate and sustain conversations, increasing the usefulness in assessing qualitative impairments in communication generally associated with autism spectrum disorders (Sparrow et al. 2005). The Daily Living Skills domain includes more items to measure independent living. The Socialization domains includes more items to measure how the individual is able to understand and use nonverbal means to regulate and maintain social interactions and relationships. This makes the Socialization domain more sensitive to the social interaction difficulties characteristic of individuals with ASD (Sparrow et al. 2005). Of the 57 returned packets, 39 or 60.4 % of parents completed the Time 2 VABS-II.

Procedure

Participant Pool, Recruitment, and Data Collection

Stored service coordination and evaluation charts for children with birthdates from 1995 to 2005 were reviewed. The ASD-DMP Protocol was applied by the first author (PT) to the charts to identify children who fit the "ASD Confirmed" and "ASD Highly Likely" categories. This resulted in an initial pool of 214 potential participants. Contact information was retrieved first from the chart; if the information was found to be out-of-date, then efforts were made to obtain contact information using standard internet search capabilities such as WhitePages.com™. After multiple efforts to get contact information, 98 charts were determined to be "Unlocatable." In some cases, addresses and phone numbers appeared functional but phone calls were never returned. These potential participants were sent recruitment letters. If no response was ever received, these charts also were deemed Unlocatable. However, basic demographic information and Time 1 data (child age and gender, town or city of residence, also early

CARS and Vineland scores if available) were recorded from the chart in a de-identified manner so that comparisons could be made between included and excluded participants.

When contact information was functional, parents were called and recruited into the study by the first author, whose position at the agency allowed her access to these records. If parents consented to participate, a consent form and the set of questionnaires and measures were mailed to them with a stamped, addressed return envelope. Thirteen parents declined participation and these charts were labeled "Declined;" their de-identified Time 1 data was also entered. Of the 13, nine declined because of the time commitment, two declined because of a negative attitude toward participating in studies, and two were unknown as to reason for declining.

Follow-up phone calls were made in 2 weeks if the questionnaires were not returned. These calls continued until either the materials were returned or it was deemed that the parents were in fact not going to participate. The latter charts/potential participants were called "Packet Not Returned," and their de-identified Time 1 data was also recorded. Parents who returned the packets ("Packet Returned") were then called to complete the VABS-II over the phone. A proportion of the parents did not complete the VABS-II in this way because they did not respond to efforts to reach them and have them call back. Another group was parents who agreed to participate and gave enough information during the initial phone call to identify the Time 2 diagnosis, but who did not end up returning packets. This information did supply data for some analyses, as their Time 1 data was also recorded in a de-identified manner; this group was called "Packet Not Returned/Time 2 Dx."

Determining Time 2 Diagnosis

Before the study began, we made the incorrect assumption that most families would have had up-to-date diagnostic evaluations for their school-age children. In fact only two families reported that this was the case. Other participating children had seen psychiatrists or developmental pediatricians recently, but for medication management and/or school services recommendations. Many parents considered the ASD designation their child received when very young as still applying, and their efforts had subsequently been directed toward educational and behavioral management. Another group of parents believed that their child no longer fit the diagnosis, and still others described their child as being confusing diagnostically and they (the caregivers) paid attention primarily to management concerns rather than obtaining a definitive diagnosis.

Therefore a post hoc method was developed to designate a Time 2 diagnosis for each child study participant so that they fell into one of three categories: Moderate/Severe ASD, Mild ASD, and No ASD/Learning Disability (LD) (this last category including those who no longer appeared to have an ASD, but still could have a learning disability). The materials reviewed to place a child participant in one of these three categories were: statements made by the parent when first contacted about the study; responses on the *Current Behavioral and Social Functioning Questionnaire;* current school placement and services being received, and GARS scores.

To determine the reliability of this tripartite diagnostic classification system, the first two authors reviewed 25 charts independently and classified them into one of the three categories. Agreement between the raters was assessed using the kappa statistic, κ, which adjusts total agreement for the proportion of agreement arising from chance (Fleiss 1981). For each of these assessments, values of $\geq$.75 represent excellent agreement, whereas values below .40 represent poor agreement (Rosner 1995; Fleiss 1981). A test of significance determines whether kappa differs from zero, and 95 % confidence intervals give the degree of precision of the estimate (Donner and Eliasziw 1992; Fleiss 1981). In further analyses, the RATCAT program (Cicchetti and Heavens 1981) was used to account for agreement among adjacent categories for the three-level classification. The specific weighting scheme, suggested by Cicchetti and Sparrow (1981) was applied to classification pairs as follows: (a) 1.00 for pairs scored identically, (b) .80 for pairs scored one level apart that do not involve "present" versus "absent" (e.g., Moderate/Severe vs. Mild ASD) and, (c) .60 for pairs scored one level apart that do involve "present" versus "absent" (e.g., mild vs. no diagnosis). Weighted kappa statistics were computed for overall agreement and for agreement at each level.

Overall, there was good agreement between raters (see Table 2). Overall agreement was 84 %. The largest proportion of disagreement occurred in distinguishing between Mild ASD and Moderate/Severe. When the weighted kappa was computed according to the method of Cicchetti et al. (2006), overall agreement was very good ($\kappa_w = .834$). After reliability was established, all of the subsequent data charts were reviewed by the two evaluators and if a disagreement occurred, it was resolved through discussion.

Analyses

Missing Data

There was opportunity for uneven sets of data as a result of both the archival nature of the data (e.g., early charts not having CARS or Vineland scores by virtue of choices made

Table 2 Inter-rater agreement for autism three-category classification scheme: Severe/Moderate ASD, Mild ASD, No ASD/LD

Rater 2	Rater 1			
	Mod/Sev ASD	Mild ASD	No ASD/LD	Total
Mod/Sev ASD	9	0	0	9 (36 %)
Mild ASD	3	9	1	13 (52 %)
No ASD/LD	0	0	3	3 (12 %)
Total	12 (48 %)	9 (36 %)	4 (16 %)	25 (84 %)

ASD autism spectrum disorder, *LD* learning disability

by clinical evaluators) and by the different level of participation of the recruited parents—primarily whether the VABS-II could be administered at Time 2. In addition, as described above, diagnostic outcome could be determined for some children even if the packet was not returned. For some analyses, all participant data that had a Time 2 diagnosis was used regardless of whether a packet was returned. For all analyses, the sources of missing data are identified.

Results

Participant Characteristics

Of the 214 charts identifying potential participants, 126 could be located and contacted (50.1 %). Of these, 13 families declined participation (10.3 %). Packets were sent out to 113 parents, of which 57 were returned, making a return rate of 52.2 %. Of those who did not return their research packet, 23 had Time 1 data and could be given a Time 2 diagnosis.

Comparison Among Unlocatable, Declined, Packet Returned, and Packet Not Returned Groups

Basic Time 1 data were compared across Declined, Packet Returned, Packet Not Returned, and Unlocatable (for 25 of the Unlocatable charts) groups to determine if there were any differences that suggested bias in the data for those who did return the packet. For example, it is clear that for the most part this is a highly educated sample (see below); this is probably due in part to the overall demographic of the county, as well as to the sampling method—those with a lower SES are more likely to move around more or not have a current phone/contact information. However, for all variables compared—Age of ASD Recognition, CARS, and VABS ABC—there were no significant differences across groups.

Participant Characteristics

Of the final sample, 66 of the children were male and 14 were female, (male:female ratio of 4.7:1). Table 3A shows child age (at time of ASD Recognition) at Time 1 and then at Time 2. One of the important features of this sample is that a large number of children had been identified at relatively young age. To make this more clear, Table 3B shows the distribution of Age at ASD Recognition. Here it can be seen that almost 54 % were 24 months or younger when identified, after which relatively intensive services began for the majority.

Of the 80 participants, 81 % had an ASD diagnosis recorded in the chart from the early intervention evaluation and/or a community professional (as well as behavioral descriptors and intervention experiences that corroborated ASD). The remainder was judged to have ASD based on the behavioral and treatment evidence as noted through applying the ASD-DMP.

Table 4 presents parent demographics including age and education, showing that this was a very highly educated group as a whole.

Time 2 Diagnostic Outcomes

Table 5 shows a number of features related to diagnostic outcomes including, first, Time 2 diagnostic status distribution. About half (51.9 %) the participants fell in the Moderate/Severe ASD category, about 30 % (28.4 %) in the Mild ASD category, and about 20 % (19.8 %) in the No ASD/LD category. Second, male:female distribution is shown for each of the three diagnostic outcome groups. Although there appears to be a differential male-to-female

Table 3 (A) Mean participant ages at Time 1 (T1) and Time 2 (T2) (N = 80), and (B) age range (in months) at Time 1

(A) Mean participant ages at Time 1 (T1) and Time 2 (T2) (N = 80)[a]

	M	*SD*	Range
T1 age	24.9 months	5.0 months	16–36 months
T2 age	10 years, 6 months	26 months	7–16 years

(B) Age range (in months) at Time 1[a]

n	Range
8	16–18
15	19–21
16	22–24
13	25–27
14	28–30
9	31–33
2	34–36

[a] Includes packet returned plus packet not returned/Time 2 Dx

ratio depending on the diagnostic outcome group, nonparametric statistical analysis showed that differences approached significance but did not reach it at the accepted levels.

As shown in Table 5, autism-related diagnoses had been given to all the participants in the Moderate/Severe group, the majority of the Mild group, and none of the No ASD/LD group (past the early intervention/preschool period).

Next, parents were asked what other clinical diagnoses assigned by community professionals applied to their child. Recall that few children had recent evaluations, so these are diagnoses that had been given at any point up to the present, but past the early intervention period when most had has an ASD diagnosis given them. It can be seen that certain diagnoses were more prevalent for the Moderate/Severe versus Mild versus No ASD/LD groups. In particular, Asperger syndrome had been applied to the Mild ASD group for slightly less than half the children. ADHD and learning disability diagnoses occurred for all groups, but were particularly prevalent for both the Mild and No ASD/LD groups.

Based on reporting on the GARS, about 20 % of the Moderate/Severe group only was nonverbal or barely verbal.

The next part of Table 5 shows parent-reported difficulties that their children were currently experiencing. There was a high rate of parent-reported learning problems reported across all groups, ranging from 75 % for the No ASD/LD group to 100 % for the Moderate/Severe group. Language problems had rates almost as high across the groups. Attentional problems were reported for about one-third of No ASD/LD children, while they were reported to occur in over 70 % of children for both ASD groups. The ASD groups were also similar in that about 95 % of parents reported social problems, compared to 12 % (one out of eight children) in the No ASD/LD group. Behavior problems were defined as aggression and noncompliance, and these were reported to be present in over half the Moderate/Severe ASD group, about one-fifth of the Mild ASD group, and very infrequently (12 %) in the No ASD/LD group.

Finally, parents reported the presence of sensory sensitivities of any type in a little over half of the Moderate/Severe ASD group, about 40 % of the Mild ASD group, and rarely (1 in 8, or 12 %) in the No ASD/LD group. Auditory sensitivities were most commonly reported for

Table 4 Parent demographics

	Mother			Father		
	M	*SD*	Range	*M*	*SD*	Range
Age at child's birth	34.2 years	4.1	25–42	37.0 years	6.3	26–53

Education level completed	Mother		Father	
	n	%	*m*	%
High school or less	3	5.1	10	16.7
2–4 years college	28	47.4	22	36.7
Master's degree or higher	29	47.5	28	46.6

Includes packet returned participants and packet not returned/time 2 Dx for which demographic information could be obtained from the early intervention chart

Table 5 Distribution of Time 2 diagnosis, gender, community professional diagnoses, and parent-reported problem areas for participants at school age

	Time 2 diagnostic category[a]						Total
	Moderate/Severe ASD		Mild ASD		No ASD/LD		
	n	%	*n*	%	*n*	%	
Diagnosis	42	51.9	23	28.4	16	19.8	80
Gender							
Male	35	85.4	21	90.9	10	61.1	66
Female	7	14.6	2	9.1	5	38.8	14
M:F ratio	5:1		10.5:1		1.6:1		
ASD diagnosis from community professional[b]	$n = 27$		$n = 19$		$n = 11$		57
Autism, ASD, PDD, PDD-NOS, Autistic Disorder	27	100	13	68	0	0	
Asperger syndrome	0	0	8	42	0	0	
ADHD	2	7	4	21	2	20	
Learning disability[d]	3	3	5	26	5	45	
Parent reported problems[e]	$n = 26$		$n = 18$		$n = 8$		52
Learning	26	100	16	89	6	75	
Language	25	96	13	72	5	62	
Attention deficit	19	73	13	72	3	37	
Social	25	96	17	94	2	25	
Behavior	16	61	4	22	1	12	
Sensory sensitivities (any)	15	58	7	39	1	12	
Auditory	12	46	6	33	1	12	
Visual	9	35	0	0	1	12	
Tactile	6	23	4	22	1	12	
Motion/vestibular	3	10	0	0	0	0	

ASD autism spectrum disorder, *LD* learning disability

[a] Includes packet returned plus packet not returned/Time 2 diagnosis

[b] Includes packet returned only

[c] Includes those who have no words and those with a few words only

[d] Includes the following reported diagnoses: non-verbal learning disability, language processing disorder, dyslexia, apraxia, expressive—receptive language disorder, language-based learning disability, executive function disorder, sensory processing disorder

[e] Includes packet returned only minus 5 packets where the problem list was deemed not adequately filled out

the ASD groups; visual sensitivities were reported at a rate of 35 % in the Moderate/Severe group, compared to virtually none for the other groups, and tactile sensitivities occurred in about one-fifth of both ASD groups. The endorsement for all different sensitivities was accounted for by one particular child in the No ASD/LD group.

Table 6 shows the GARS scores for the same groups. Although GARS scores were used in making the decision regarding diagnosis group, and therefore do not serve as independent verification of the method, it is noteworthy that the overall scores for each group fall into line with the severity level assumed for the groups. For each subscale and the total score, a one-way ANOVA was highly significant; a post hoc Scheffe test showed that each ASD group was significantly different from the other on all GARS scores.

Time 2 Nondiagnostic Outcomes

Adaptive Behavior

Table 7 shows the Vineland Adaptive Behavior Scales-II scores for each of the groups. The VABS-II scores were *not* used in the judgment-based Time 2 diagnosis process. A post hoc Scheffe test confirmed that for all VABS-II domain and total scores, each diagnostic group was significantly different from the two others.

Since a well-known issue affecting daily life for families with a child on the autism spectrum is that of food intake, a question on the Background Questionnaire inquired about whether their child was a "picky eater." (This information was not considered when making judgments about which diagnostic category to place the participant). Parents endorsed "Yes" for 54 % (n = 14) of the Moderate/Severe ASD group, for 22 % (n = 4) of the Mild ASD group, and 12 % (n = 1) of the No ASD/LD group.

Four specific questions on the parent questionnaire (C. Physical and Social Functioning section) were used to further describe Time 2 child functioning. The first had to do with physical functioning/motor coordination, which parents rated as Excellent, Good, Fair, or Has Significant Difficulty. Table 8 shows that physical functioning was rated as problematic for about half of each of the ASD groups and not problematic for half; comparatively more of the No ASD/LD group was reported as having better

Table 6 GARS means and standard deviations for Time 2 diagnosis groups and ANOVA results

	Diagnostic category						*F* (2,35)	Post-hoc
	Mod/Sev ASD ($n = 27$)		Mild ASD ($n = 19$)		No ASD/LD ($n = 11$)			
	M	*SD*	*M*	*SD*	*M*	*SD*		
Stereotypic behavior	9.33	3.19	5.26	2.37	1.73	1.42	34.55***	Mod/Sev < Mild < No
Communication	9.48[a]	2.87	4.11	1.85	1.27	1.68	54.38***	Mod/Sev < Mild < No
Social interaction	8.58	2.94	5.06	2.82	1.55	1.37	29.05***	Mod/Sev < Mild < No
Autism quotient	94.1	14.54	70.8	12.33	53.2	7.73	45.12***	Mod/Sev < Mild < No

Includes packet returned participants only; two of the returned GARS could not be scored because they were incompletely filled out

GARS Gilliam Autism Rating Scale, *Mod/Sev* Moderate/Severe, *ASD* autism spectrum disorder, *LD* learning disability

*** $p < .001$

[a] For this cell, n = 22 because if the child is nonverbal, the subscale is not filled out

Table 7 VABS-II means and standard deviations for Time 2 diagnostic groups and ANOVA results

	Diagnostic category						*F* (2,35)	Post hoc
	Mod/Sev ASD ($n = 18$)		Mild ASD ($n = 12$)		No ASD/LD ($n = 9$)			
	M	*SD*	*M*	*SD*	*M*	*SD*		
Communication	64.7	12.9	81.2	8.0	107.1	14.8	36.83***	Mod/Sev < Mild < No
Daily living skills	68.3	14.5	86.9	11.0	107.4	9.4	30.17***	Mod/Sev < Mild < No
Socialization	62.1	14.6	77.8	13.7	106.2	9.3	33.03***	Mod/Sev < Mild < No
Adaptive behavior composite	63.9	13.2	80.2	8.7	107.1	9.8	44.68***	Mod/Sev < Mild < No

Includes Packet Returned only; of the 57 who returned packets, 18 could not be reached for the VABS-II administration

VABS-II Vineland Adaptive Behavior Scales, Second Edition, *Mod/Sev* Moderate/Severe, *ASD* autism spectrum disorder, *LD* learning disability

*** $p < .001$

physical functioning. The issue of participating in team sports, which could be a function of either or both physical or social functioning, showed a more dramatic difference between the two ASD groups and the No ASD/LD group.

Parent's judgment about their child's social functioning, both in general and as defined by participation on sleepovers and play dates, showed a pattern that was heavily weighted toward less active social functioning for children with ASD. No child in the No ASD/LD group was judged by their parent to have these experiences "very little in comparison to others," whereas most children in the two ASD groups were judged in this way.

The next table, Table 9, shows the medications that parents reported their children currently taking. For the entire group, 26 or 45.6 % were on medication of some sort, and of those, 13 or 22 % were on more than one medication (range 2–3 medications). However, the table shows that patterns of use were different depending on the severity of ASD. Medication for ADHD was most prevalent for the Mild ASD group, and neuroleptics, which include Risperdal and Abilify, were represented most heavily in the Moderate/Severe ASD group.

School Placement

Table 10 presents the Time 2 placements for the three diagnostic groups. The placements were categorized as shown in the table. The more service-intensive, specialized, and segregated settings were favored for the two ASD groups. Three of the children in the Moderate/Severe ASD group were now in residential placements. The parents of all three participants reported that behavior problems contributed to this decision.

The classroom supports reported by parents showed that about half the children in the Moderate/Severe ASD group had aides with them, compared to about a third of the Mild ASD group, and none of the No ASD/LD. Behavior plans were more common for the more severe group, but American Disability Act (ADA) Section 504 accommodations were reported for over 80 % of the mild group, and close to a third of the No ASD/LD group.

Table 8 Parent-reported physical and social functioning

	Time 2 diagnostic category					
	Mod/Sev ASD ($n = 26$)		Mild ASD ($n = 19$)		No ASD/LD ($n = 11$)	
	n	%	*n*	%	*n*	%
Physical functioning (coordination)						
Significant difficulty or fair	16	59	10	50	3	33
Good or excellent	10	41	9	50	8	67
Participates in team sports						
No	18	71	16	83	4	42
Yes	8	29	3	17	7	58
Social functioning						
Significant difficulty or fair	24	92	17	94	2	29
Good or excellent	2	8	2	6	9	71
Sleepovers and play dates						
Very little in comparison	25	96	12	63	0	0
Somewhat less than others	1	4	4	21	3	36
About the same as others	0	0	2	5	8	64

Includes only packets returned, minus one where the parent skipped the question

Mod/Sev Moderate/Severe, *ASD* autism spectrum disorder, *LD* learning disability

Discussion

Diagnostic Outcomes

The first goal of this study was to follow children who were diagnosed or identified with autism spectrum disorder early (before the age of three, and many before and by 24 months) into school age. The results showed that the diagnostic or symptom level outcomes were very similar to that of other longitudinal studies: approximately 20 % went "off the spectrum" (Woolfenden et al. 2012). Further examination demonstrated, however, that this "optimal outcome" group retained important learning challenges. Over half had been given a diagnosis of ADHD and/or some type of learning disability by a community

Table 9 Time 2 medication use across ASD diagnostic categories

	Time 2 diagnostic category[a]					
	Mod/Sev ASD ($n = 27$)		Mild ASD ($n = 19$)		No ASD/LD (n = 11)	
	n	%	*n*	%	*n*	%
Uses 1 or more medication	16	59	9	47	1	9
Uses 2–4 meds	9	33	4	21	0	0
Type of medication	$n = 24$[b]		$n = 19$		$n = 11$	
Stimulants and non-stimulants[c] used for ADHD	5	21	8	42	1	9
Neuroleptics	9	37	2	10	0	0
SSRIs	5	21	3	16	0	0
Alpha adrenergic agonist	2	8	0	0	0	0
Anti-epileptic (Depakote)	1	4	0	0	0	0

Mod/Sev Moderate/Severe, *ASD* autism spectrum disorder, *LD* learning disability

[a] Returned packets only

[b] Three parents reported that medication was taken but did not report a drug name, so those cases were excluded from the analysis

[c] Nonstimulant is atomoxetine (Strattera)

Table 10 School-age placements and service supports for Time 2 diagnostic groups

School placement	Time 2 diagnostic category[a]					
	Moderate/Severe ASD ($n = 31$)		Mild ASD ($n = 18$)		No ASD/LD ($n = 12$)	
	n	%	*n*	%	*n*	%
Regular public school class with no or one service	0	0	1	5	6	50
Regular private school	0	0	0	0	2	17
Regular public school program, range of services	3	10	9	50	4	33
Special education class in public school system	20	64	7	39	0	0
Private special education school	5	16	1	5	0	0
Residential placement	3	10	0	0	0	0
Service supports[b]	$n = 27$		$n = 19$		$n = 11$	
Aide	15	55	7	37	0	0
Behavior plan	8	30	3	16	0	0
504 Accommodations	6	22	16	84	3	27

[a] Includes Packet Returned plus the Packet Not Returned/Time 2 Dx when this information was known(n = 61)

[b] Includes Packet Returned only

professional. In addition, 75 % of parents of the No ASD/LD group characterized their child as have a learning problem and 62 % a language problem. A smaller but notable percent endorsed social and behavioral functioning as areas of difficulty. This is consistent with previous research examining the social, learning, and language profiles of children who had lost their ASD diagnoses from early childhood (Fein et al. 2005; Kelley et al. 2006). Three out of eight parents reported attentional problems for this group; previous research has suggested that "recovered" (from ASD) children may retain ADHD symptoms in particular (Helt et al. 2008). Additional confirmation of continuing problems was seen in the school placement and supports results in the current study, in that about one-third of the No ASD/LD group received related services in a regular classroom and/or 504 accommodations.

Nondiagnostic Outcomes Across Symptom Level Groups

This study, however, endeavored to distinguish between more and less severe outcomes for those who continued on the spectrum, as well as to examine a number of behaviors that provide a more functional, detailed understanding of the ramifications of differential outcome severity. Each participant was placed in one of the three categories by the researchers (Moderate/Severe ASD, Mild ASD, and No ASD/learning disability) based on parent statements during the initial phone call, results of the study questionnaires—which inquired about many features of child behavior in different settings—and the Gilliam Autism Rating Scales (GARS) results. Inter-rater reliability proved to be good for this system. The first nondiagnostic variable to be examined was adaptive behavior as measured by the Vineland Adaptive Behavior Scales-Second Edition (VABS-II). The Moderate/Severe group had means in the 60s (over two standard deviations low) across subdomains as well as the ABC; clinically, these levels signal the need for substantial support in daily living and a significant delay in skills levels compared to same-aged peers. In contrast, the Mild ASD group mean scores were in the 80s, suggesting below average skills but a less intensive need for supports on an ongoing basis. The No ASD/LD group had mean scores in the Average range.

Another way to inquire into specific outcomes was to ask parents to rate areas of developmental functioning both as general constructs as well as operationalized in terms of specific behaviors. With respect to motor coordination, more than half the children in both ASD groups were rated as problematic while the minority were so rated in the No ASD/LD group; similarly, participation in team sports was not common for the ASD groups, while it was more so for the No ASD/LD participants. Level of social functioning was even more dramatically contrasted between the ASD and non-ASD groups, as was participation in birthday parties and sleepovers. Results showed that children with ASD symptoms are having these experiences very little. The No ASD/LD participants appeared to approach typical levels for these activities.

Another feature of daily life for many children on the autism spectrum is the use of medication to address a variety of challenges such as attentional, mood, and behavioral problems. Logan et al. (2012) found that of 8- to 15-year olds with ASD using Medicaid in South Carolina, 40 % used psychotropic medication and 20 % used multiple psychotropic classes. Arnold et al. (2006) reported a usage rate of 46.7 % for children and adolescents with ASD in Ohio, with 11.9 % taking more than one psychotropic drug. In the current study, the rate of psychotropic drug use was 59 and 47 % for the Moderate/Severe and Mild ASD groups, respectively. One child in the No ASD group took a stimulant for ADHD. The rate of polypharmacy, or using more than one psychotropic drug, was about one-third for the Moderate/Severe group and about one-fifth for the Mild group.

Previous surveys have shown the association between greater use of medication and increased autism severity (Aman et al. 2003); however, no study to date has examined the use of different classes of medications in relation to level of ASD severity groups in school-aged children. In

this study, the most common type of psychotropic medications used by the Moderate/Severe group was neuroleptics, including risperadone and aripiprazole—both used to treat irritability and associated behavior problems such as aggression and tantrums. They were used at a rate of 37 % in the Moderate/Severe ASD group and 10 % in the Mild ASD group. In contrast, the most-used drug type for the Mild ASD group was stimulants and nonstimulants that are prescribed to treat ADHD (42 %). This again appears to reflect the prominence of attentional symptoms when the overall autism symptom picture is milder. One study that did investigate higher functioning children, adolescents, and adults found that 55 % of participants were taking psychotropic medications, with 29.3 % on more than one simultaneously; the most common drug class for this group, however, was SSRIs, followed by stimulants, followed by neuroleptics (Martin et al. 1999). It is possible that inclusion of older aged individuals compared to the current study shifted the proportion of use among psychotropic classes. A limitation of our study is that numbers are small compared to other surveys of medication usage in children and adolescents with ASD; nonetheless; these preliminary results alert future researchers to this issue.

The final set of nondiagnostic outcome variables was related to school placement. Very few studies have examined this feature of outcome; the great majority use school or classroom placement as an outcome variable for early intervention effects (Akshoomoff et al. 2010; Cohen et al. 2006; Eaves and Ho 1997; Harris and Handleman 2000; Smith et al. 2000; White et al. 2007), rather than conducting a survey of a sample of children. One study that did take such an approach surveyed 76 children with ASD at a mean age of 11 years (Eaves and Ho 1997). Results showed that none was in institutions or segregated schools, 35 % were in regular classes with an aide, and 16 % were in regular classes without an aide. Children who were older, more delayed, and with more severe symptoms were more likely to be in special classes. The current study, contrasting different levels of ASD severity, predictably found that the Moderate/Severe children had a much greater chance of being in special and segregated settings in that 90 % were in either in special classes with a range of services (the majority, 64 %, was in this setting), special education schools (all in this subgroup were in a particular school in the county that specializes in Applied Behavior Analysis), or in residential settings (three children). The remaining 10 % were in non-special education classes with relatively heavy services, including fulltime aides. In contrast, the Mild ASD group had 55 % in regular classes but with a range of services, and 45 % in segregated, special education classes, including one child in a private school specifically for language-based learning disabilities. None of the children in the No ASD group was in specialized settings; two-thirds were in regular classrooms with no or one service (e.g., for speech articulation), while one-third were in regular classrooms with more than one related service (e.g., speech or occupational therapy, reading/writing support, or resource room part of the day).

It seems evident that distribution of school setting for any group of children with disabilities is going to vary not only with the students' learning profile, but also with availability of resources and school district policy. For example, there are a number of private, specialized educational resources available in the current study's geographic area. In addition, higher-resource school districts do pay out-of-district or residential tuition if it is found that they cannot meet the student's needs in-district. (This policy is reversing, however, with more demand to create programs that keep students in the district). A related issue is that two districts in this study's catchment area have a "full inclusion" policy, and direct their resources toward in-classroom supports rather than out-of-district placement. Because of this, some students whose cognitive and behavioral functioning might result in a segregated placement in another geographic area, are instead in an inclusion or regular class with heavy supports.

A future report will address the school services of this sample in greater detail, including the progression of the distribution of settings over three points in time (preschool, Kindergarten, and grade school).

Methodological Considerations

A last but overarching goal of this study was to examine the utility of relatively indirect methods of determining early diagnostic status (chart review and abstraction) as well as follow-up diagnostic status (parent questionnaire). Since reliability must underlie any claim to validity, this issue should first be evaluated for both strategies. A small literature has developed around chart abstraction for identifying young children with ASD that supports its reliability and validity (Towle et al. 2009; Wiggins et al. 2013; Fein et al. 2013). A much larger literature exists that examines the validity of early diagnoses (Time 1 diagnostic status determined primarily through direct examination) based on the stability of those diagnoses over time (see Woolfenden et al. 2012). One method of evaluating the validity of the current study's early diagnosis approach, therefore, is to compare the results with what has been established in the literature. The fact that approximately 20 % of the current subjects went "off the spectrum," which is the prevailing result across the studies, is positive evidence that the records-based judgments are valid. In addition, 80 % of the charts chosen to be included did specify ASD or one its diagnostic iterations (not the exact subjects that stayed on the spectrum later, however), and

thus only 20 % of the charts relied on behavioral descriptions and service patterns to be designated as ASD.

The reliability of the Time 2 diagnostic status or symptom level assigned to each child was first tested directly through an inter-rater reliability substudy that demonstrated good reliability. The rest of the evidence rests primarily on the logic of the various outcomes across Time 2 variables as well as agreement with previous studies examining similar outcomes. The GARS scores were reviewed for Time 2 group placement and thus not independent verification of the method; however, they were at appropriate mean levels for the groups, with each group significantly different from the other. School placement was also reviewed when placing children in outcome categories, but the entire pattern of the groups in terms of special education placements, services, and educational supports such as aides, was entirely consistent with the severity grouping levels. Variables that were *not* a source for decision-making are as follows. The Problem List and the profiles of problems parents endorsed were consistent with the groupings, as were diagnoses that had been assigned by community professionals the families encountered. Medication use showed logical and differential patterns across the groups. Finally, the VABS-II averages of the groups were clinically consistent, as discussed in the first section above.

The limitations of the current study were related to the methodological issues being investigated; that is, that direct clinical evaluations were not performed to confirm the diagnostic status of children at Time 2. Thus, the majority of Time 2 data relied on parent perception and report. Future investigations of this sample will endeavor to have an independent check on diagnosis through clinical examination. As well, future reports will focus on predictive features of early developmental status and ASD symptoms, taking into account early intervention experiences of the children.

Acknowledgments This work was done while the first author was supported by Grant # T73MC04320 from Department of Health and Human Services Health Resources and Services Administration. It was also part of a doctoral dissertation completed by Karyn Vacanti-Shova, Fordham University. We are grateful for the assistance of graduate students participating on the first author's research team while completing their Leadership Education in Neurodevelopmental Disabilities fellowship at the Westchester Institute for Human Development.

References

Akshoomoff, N. A., Stahmer, A. C., Corsello, C., & Mahrer, N. E. (2010). What happens next? Follow-up from the children's toddler school program. *Journal of Positive Behavior Interventions, 12*, 245–253.

Aman, M. G., Lam, K. L., & Collier-Crespin, A. (2003). Prevalence and patterns of use of psychoactive medicines among individuals with autism in the Autism Society of Ohio. *Journal of Autism and Developmental Disorders, 33*, 527–534. doi:10.1023/A:1025883612879.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (text revision, 4th ed.). Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Arnold, L., Aman, M. G., Cook, A. M., Witwer, A. N., Hall, K. L., Thompson, S., et al. (2006). Atomoxetine for hyperactivity in autism spectrum disorders: Placebo-controlled crossover pilot trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 45*, 1196–1205. doi:10.1097/01.chi.0000231976.28719.2a.

Avchen, R., Wiggins, L. D., Devine, O., Van Naarden Braun, K., Rice, C., Hobson, N. C., et al. (2011). Evaluation of a records-review surveillance system used to determine the prevalence of autism spectrum disorders. *Journal of Autism and Developmental Disorders, 41*, 227–236. doi:10.1007/s10803-010-1050-7.

Boyd, B. A., Odom, S. L., Humphreys, B. P., & Sam, A. M. (2010). Infants and toddlers with autism spectrum disorder: Early identification and early intervention. *Journal of Early Intervention, 32*(2), 75–98. doi:10.1177/1053815110362690.

Centers for Disease Control and Prevention. (2007). Prevalence of autism spectrum disorders—Autism and Developmental Disabilities Monitoring Network, 14 sites, United States, 2002. *Morbidity and Mortality Weekly Report, 56*, 12–28.

Centers for Disease Control and Prevention. (2009). Prevalence of autism spectrum disorders—Autism and Developmental Disabilities Monitoring Network, United States, 2006. *Morbidity and Mortality Weekly Report, 58*, 1–24.

Centers for Disease Control and Prevention. (2012). Prevalence of autism spectrum disorders—Autism and Developmental Disabilities Monitoring Network, 14 sites, United States, 2008. *Morbidity and Mortality Weekly Report, 61*, 1–19.

Charman, T., Taylor, E., Drew, A., Cockerill, R., 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. doi:10.1111/j.1469-7610.2004.0037.x.

Chawarska, K., Klin, A., Paul, R., Macari, S., & Volkmar, F. (2009). A prospective study of toddlers with ASD: Short-term diagnostic and cognitive outcomes. *The Journal of Child Psychology and Psychiatry, 50*(10), 1235–1245. doi:10.1111/j.1469-7610.2009.02101.x.

Cicchetti, D., Bronen, R., Spencer, S., Haut, S., Berg, A., Oliver, P., et al. (2006). Rating scales, scales of measurement, issues of reliability: Resolving some critical issues for clinicians and researchers. *Journal of Nervous and Mental Disease, 194*, 557–564. doi:10.1097/01.nmd.0000230392.83607.c5.

Cicchetti, D. V., & Heavens, R. (1981). A computer program for determining the significance of the difference between pairs of independently derived values of kappa or weighted kappa. *Educational and Psychological Measurement, 41*, 189–193. doi:10.1177/001316448104100119.

Cicchetti, D. V., & Sparrow, S. A. (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. *American Journal of Mental Deficiency, 86*, 127–137.

Cohen, H., Amerine-Dickens, M., & Smith, T. (2006). Early intensive behavioral treatment: Replication of the UCLA model in a community setting. *Journal of Developmental and Behavioral Pediatrics, 27*, S145–S162.

Cox, A., Klein, K., Charman, T., Baird, G., Baron-Cohen, S., Swettenham, J., et al. (1999). Autism spectrum disorders at 20

and 42 months of age: Stability of clinical and ADI-R diagnosis. *Journal of Child Psychology and Psychiatry, 40*, 719–732.

Donner, A., & Eliasziw, M. (1992). A goodness-of-fit approach to inference procedures for the kappa statistic: Confidence interval construction, significance testing and sample size estimation. *Statistics in Medicine, 11*, 1151–1519.

Eaves, L., & Ho, H. (1997). School placement and academic achievement in children with autistic spectrum disorders. *Journal of Developmental and Physical Disabilities, 9*, 277–291.

Eaves, L. C., & Ho, H. H. (2004). The very early identification of autism: Outcome to age 4 ½–5. *Journal of Autism and Developmental Disorders, 34*, 367–378.

Fein, D., Barton, M., Eigsti, I., Kelley, E., Naigles, L., Schultz, R. T., et al. (2013). Optimal outcome in individuals with a history of autism. *Journal of Child Psychology and Psychiatry, 54*, 195–205.

Fein, D., Dixon, P., Paul, J., & Levin, H. (2005). Brief report: Pervasive developmental disorder can evolve into ADHD: Case illustrations. *Journal of Autism and Developmental Disorders, 35*, 525–534. doi:10.1007/s10803-005-5066-3.

Fleiss, J. L. (1981). Balanced incomplete block designs for inter-rater reliability studies. *Applied Psychological Measurement, 5*, 105–112. doi:10.1177/014662168100500115.

Fombonne, E. E. (2002). Epidemiological trends in rates of autism. *Molecular Psychiatry, 7*, S4–S6. doi:10.1038/sj.mp.4001162.

Gillberg, C., Ehlers, S., Schaumann, H., Jakobsson, G., Dahlgren, S., Lindblom, R., et al. (1990). Autism under age 3 years: A clinical study of 28 cases referred for autistic symptoms in infancy. *Journal of Child Psychology and Psychiatry, 31*(6), 921–934. doi:10.1111/j.1469-7610.1990.tb00834.x.

Gilliam, J. E. (1995). *Gilliam Autism Rating Scale*. Austin, TX: Pro-Ed.

Harris, S. L., & Handleman, J. S. (2000). Age and IQ at intake as predictors of placement for young children with autism: A four-to-six-year follow-up. *Journal of Autism and Developmental Disorders, 30*, 137–142.

Helt, M., Kelley, E., Kinsbourne, M., Pandey, J., Boorstein, H., Herbert, M., et al. (2008). Can children with autism recover? If so, how? *Neuropsychology Review, 18*, 339–366. doi:10.1007/s11065-008-9075-9.

Jonsdottir, S. L., Saemundsen, E., Asmundsdottir, G., Hjartardottir, S., Asgeirsdottir, B. B., Smaradottir, H. H., et al. (2007). Follow-up of children diagnosed with pervasive developmental disorders: Stability and change during the preschool years. *Journal of Autism and Developmental Disorders, 37*, 1361–1374. doi:10.1007/s10803-006-0282-z.

Kelley, E., Paul, J. J., Fein, D., & Naigles, L. R. (2006). Residual language deficits in optimal outcome children with a history of autism. *Journal of Autism and Developmental Disorders, 36*, 807–828. doi:10.1007/s10803-006-0111-4.

Kleinman, J. M., Ventola, P. E., Pandey, J., Verbalis, A. D., Barton, M., Hodgson, S., et al. (2008). Diagnostic stability in very young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38*, 606–615. doi:10.1007/s10803-007-0427-8.

Kogan, M. D., Blumberg, S. J., Schieve, L. A., Boyle, C. A., Perrin, J. M., Ghandour, R. M., et al. (2009). Prevalence of parent-reported diagnosis of autism spectrum disorder among children in the US, 2007. *Pediatrics, 124*, 1395–1403. doi:10.1542/peds.2009-1522.

Logan, S. L., Nicholas, J. S., Carpenter, L. A., King, L. B., Garrett-Mayer, E., & Charles, J. M. (2012). High prescription drug use and associated costs among Medicaid-eligible children with autism spectrum disorders identified by a population-based surveillance network. *Annals of Epidemiology, 22*, 1–8. doi:10.1016/j.annepidem.2011.10.007.

Lord, C. (1995). Follow-up of two-year-olds referred for possible autism. *Journal of Child Psychology and Psychiatry, 36*, 1365–1382. doi:10.1111/j.1469-7610.1995.tb01669.x.

Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., & Pickles, A. (2006). Autism from 2 to 9 years of age. *Archives of General Psychiatry, 63*, 694–701.

Malhi, P., & Singhi, P. (2011). Follow up of children with autism spectrum disorders: Stability and change in diagnosis. *Indian Journal of Pediatrics, 78*, 941–945.

Martin, A., Scahill, L., Klin, A., & Volkmar, F. R. (1999). Higher-functioning pervasive developmental disorders: Rates and patterns of psychotropic drug use. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*, 923–931. doi:10.1097/00004583-199907000-00024.

Michelotti, J., Charman, T., Slonims, V., & Baird, G. (2002). Follow-up of children with language delay and features of autism from preschool years to middle childhood. *Developmental Medicine and Child Neurology, 44*, 812–819. doi:10.1017/S0012162201002985.

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. doi:10.1177/1362361303007001005.

Moss, J., Magiati, I., Charman, T., & Howlin, P. (2008). Stability of the autism diagnostic interview-revised from preschool to elementary age in children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38*, 1081–1091. doi:10.1007/s10803-007-0487-9.

Rondeau, E., Klein, L. S., Masse, A., Bodeau, N., Cohen, D., & Guilé, J. (2011). Is pervasive developmental disorder not otherwise specified less stable than autistic disorder? A meta-analysis. *Journal of Autism and Developmental Disorders, 41*, 1267–1276. doi:10.1007/s10803-010-1155-z.

Rosenberg, R. E., Kaufmann, W. E., Law, J., & Law, P. A. (2011). Parent report of community psychiatric comorbid diagnoses in autism spectrum disorders. *Autism Research & Treatment*, 1–10. doi:10.1155/2011/405849.

Rosner, B. (1995). *Fundamentals of biostatistics*. Belmont, CA: Duxbury Press.

Scambler, D., Hepburn, S., & Rogers, S. (2006). A two-year follow-up on risk status identified by the checklist for autism in toddlers. *Journal of Developmental and Behavioral Pediatrics, 27*, S104–S110.

Schopler, E., Reichler, R. J., & Renner, B. R. (1988). *Childhood Autism Rating Scale*. Los Angeles: Western Psychological Services.

Smith, T., Groen, A., & Wynn, J. (2000). Randomized trial of intensive early intervention for children with pervasive developmental disorder. *American Journal on Mental Retardation, 105*, 269–285.

Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *Vineland Adaptive Behavior Scales*. Circle Pines, MN: American Guidance Services.

Sparrow, S., Cicchetti, D., & Balla, D. (2005). *Vineland Adaptive Behavior Scales-II (Survey Form)*. Circle Pines, MN: American Guidance Services.

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

Sutera, S., Pandey, J., Esser, E. L., Rosenthal, M. A., Wilson, L. B., Barton, M., et al. (2007). Predictors of optimal outcome in toddlers diagnosed with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37*, 98–107. doi:10.1007/s10803-006-0340-6.

Szatmari, P., Bryson, S. E., Streiner, D. L., Wilson, F., Archer, L., & Ryerse, C. (2000). Two-year outcome of preschool children with autism or Asperger's syndrome. *The American Journal of Psychiatry, 157*, 1980–1987. doi:10.1176/appi.ajp.157.12.1980.

Towle, P., Visintainer, P., O'Sullivan, C., Bryant, N., & Busby, S. (2009). Detecting autism spectrum disorder from early intervention charts: Methodology and preliminary findings. *Journal of Autism and Developmental Disorders, 39*, 444–452. doi:10.1007/s10803-008-0643-x.

Turner, L. M., & Stone, W. L. (2007). Variability in outcome for children with an ASD diagnosis at age 2. *Journal of Child Psychology and Psychiatry, 48*, 793–802. doi:10.1111/j.1469-7610.2007.01744.x.

Turner, L. M., Stone, W. L., Pozdol, S. L., & Coonrod, E. E. (2006). Follow-up of children with autism spectrum disorders from age 2 to age 9. *Autism, 10*, 243–265. doi:10.1177/1362361306063296.

van Daalen, E., Kemner, C., Dietz, C., Swinkels, S. N., Buitelaar, J. K., & van Engeland, H. (2009). Inter-rater reliability and stability of diagnoses of autism spectrum disorder in children identified through screening at a very young age. *European Child and Adolescent Psychiatry, 18*, 663–674. doi:10.1007/s00787-009-0025-8.

White, S. W., Scahill, L., Klin, A., Koenig, K., & Volkmar, F. R. (2007). Educational placements and service use patterns of individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37*, 1403–1412. doi:10.1007/s10803-006-0281-0.

Wiggins, L. D., Robins, D. L., & Yeargin-Allsopp, M. (2013). Short report: Improving record—Review surveillance of young children with an autism spectrum disorder. *Autism, 17*, 623–629. doi:10.1177/1362361312452161.

Woolfenden, S., Sarkozy, V., Ridley, G., & Williams, K. (2012). A systematic review of the diagnostic stability of autism spectrum disorder. *Research In Autism Spectrum Disorders, 6*, 345–354.