

Autism

<http://aut.sagepub.com>

Screening for autism in older and younger toddlers with the Modified Checklist for Autism in Toddlers

Juhi Pandey, Alyssa Verbalis, Diana L. Robins, Hilary Boorstein, Ami Klin, Tammy Babitz, Katarzyna Chawarska, Fred Volkmar, James Green, Marianne Barton and Deborah Fein

Autism 2008; 12; 513

DOI: 10.1177/1362361308094503

The online version of this article can be found at:
<http://aut.sagepub.com/cgi/content/abstract/12/5/513>

Published by:



<http://www.sagepublications.com>

On behalf of:



The National Autistic Society

Additional services and information for *Autism* can be found at:

Email Alerts: <http://aut.sagepub.com/cgi/alerts>

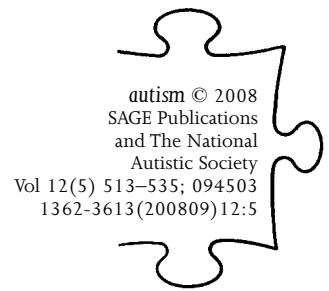
Subscriptions: <http://aut.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.co.uk/journalsPermissions.nav>

Citations (this article cites 44 articles hosted on the SAGE Journals Online and HighWire Press platforms):
<http://aut.sagepub.com/cgi/content/refs/12/5/513>

Screening for autism in older and younger toddlers with the Modified Checklist for Autism in Toddlers



JUHI PANDEY University of Connecticut, Connecticut, USA
ALYSSA VERBALIS University of Connecticut, Connecticut, USA
DIANA L. ROBINS Georgia State University, Georgia, USA
HILARY BOORSTEIN University of Connecticut, Connecticut, USA
AMI KLIN Yale Child Study Center Yale University School of Medicine,
Connecticut, USA
TAMMY BABITZ Yale Child Study Center, Yale University School of
Medicine, Connecticut, USA
KATARZYNA CHAWARSKA Yale Child Study Center, Yale
University School of Medicine, Connecticut, USA
FRED VOLKMAR Yale Child Study Center, Yale University School of
Medicine, Connecticut, USA
JAMES GREEN University of Connecticut, Connecticut, USA
MARIANNE BARTON University of Connecticut, Connecticut, USA
DEBORAH FEIN University of Connecticut, Connecticut, USA

ABSTRACT The Modified Checklist for Autism in Toddlers (M-CHAT) was used to screen younger (16–23 months) versus older (24–30 months) high- and low-risk toddlers. Refusal rates for follow-up interview showed no group differences, but parents of younger/low-risk children were more likely to refuse evaluation than parents of high-risk children. PPP for an ASD diagnosis was: younger/high-risk 0.79, older/high-risk 0.74, younger/low-risk 0.28, and older/low-risk 0.61, with PPP differing by age within the low-risk group. Most of the children in all groups, however, were diagnosed with a developmental disorder. Symptom severity generally did not differ among groups. Cognitive and adaptive measures showed minimal group differences. Therefore, older and younger toddlers had similar symptomatology and developmental delays; PPP for ASD is better at 24 than 18 months

KEYWORDS
autism
spectrum
disorders;
early
detection;
early
identification;
pediatric
screening

for low-risk children; however, these children are still highly likely to show a developmental disorder. Clinical decision making should balance early identification against the lower specificity of M-CHAT screening for the younger/low-risk group.

ADDRESS Correspondence should be addressed to: JUHI PANDEY, University of Connecticut, Storrs, Connecticut 06269, USA. e-mail: juhi_pandey@hotmail.com

Introduction

Empirical studies of toddlers with autism spectrum disorders (ASDs) have found that intensive, specialized early intervention has resulted in quantifiable gains (Horner et al., 2002; McEachin et al., 1993; Sallows and Graupner, 2005; Schreibman, 2000). In order to maximize the opportunity for specialized early intervention, the early identification and diagnosis of ASD are especially important (American Academy of Pediatrics, 2006). Recently, the AAP (Johnson et al., 2007) even suggested that it is important for children suspected of ASD to begin intervention services.

Early identification studies support the feasibility and validity of early diagnosis, even as early as 2 years (Baird et al., 2000; Robins et al., 2001; Stone et al., 2000). Along with screening studies, retrospective studies of infant videotapes (Baranek, 1999; Osterling and Dawson, 1994; Werner et al., 2000), diagnostic stability studies (Charman and Baird, 2002; Gillberg et al., 1990; Kleinman et al., 2008; Lord, 1995; Stone et al., 1999), and inter-rater reliability studies (Klin et al., 2000; Stone et al., 1994) have supported the validity of early diagnosis and have identified symptoms that may be present in the early developmental course of ASD. In addition, prospective studies of ASD have been useful in identifying symptoms present in high-risk infants (such as younger siblings of children with ASD) later diagnosed on the autism spectrum (Landa et al., 2007; Yirmiya and Ozonoff, 2007; Zwaigenbaum et al., 2007). These studies, which have focused on young infants and toddlers, have provided a picture of the symptoms and patterns of behavior observed in the very early course of the disorder.

While individual variability exists in the specific age by which toddlers exhibit ASD symptoms, researchers have found that symptoms are often present by the age of 18 months (Bryson et al., 2007; Kleinman et al., 2008; Landa et al., 2007), except in some cases of later regression. For instance, toddlers with ASD show reduced bids for joint attention, do not respond consistently to their name being called, demonstrate less social engagement driven by eye contact, and show decreased imitation of facial expressions, vocal imitations, and object-oriented imitation (Bryson et al., 2007). These children also examine objects more than typical, same-age peers and do not shift gaze or attention as easily within their environment (Zwaigenbaum

et al., 2005). In addition, by this age, children on the autism spectrum are less motorically active and have documented delays in fine and gross motor functioning compared to typical peers (Landa and Garrett-Mayer, 2006).

One important and consistent finding from the studies mentioned above has been that 2-year-old children with ASD, and especially those below 2, often present with more negative symptoms (decreases in the frequency of or lack of social and communicative behaviors) than positive symptoms (higher rates of unusual behaviors, such as stereotyped language, adherence to routines, and preoccupations) (Filipek et al., 1999; Rogers, 2001; Stone et al., 1999).

In order to maximize the opportunity for specialized early intervention to target the areas of concern noted above, autism screening instruments have been developed to help identify symptoms of autism at earlier ages. While some studies have found accelerated head growth in infants later diagnosed with ASD (Courchesne et al., 2003; Dementieva et al., 2005), this finding is not universal or specific enough to be reliable as a screening marker for early ASD, at least at present; therefore, existing screeners focus on specific behaviors to help identify children at risk for autism. However, the somewhat different clinical presentations noted above for very young children raise questions about whether the same screening instruments (as well as diagnostic criteria) can be used for children above and below the age of 2.

In addition, most of the evidence concerning the efficacy of early screening applies to children above the age of 2. Risks associated with screening include the possibility of false positives (Charwarska et al., 2007), unnecessarily alarming families (Williams and Brayne, 2006), and a possible lack of specificity in screening (Eaves et al., 2006). Consequently, researchers interested in the earliest detection of ASD have been considering the question: are there risks associated with screening even earlier?

One risk of screening before the age of 2 is that children who regress after the age of screening will be not picked up by the screening because behaviors have not manifested at the time of screening (Baird et al., 2001). This concern stems from the reported rates of regression in autism, which range from approximately 10 to 50 percent (Goldberg et al., 2003; Lord et al., 2004). While the age of regression can range between 12 and 36 months, a large number of children regress before their second birthday, with some studies reporting a mean age between 15 and 21 months (Baird et al., 2001; Bryson et al., 2007; Landa and Garrett-Mayer, 2006; Lord et al., 2004; Lyuster, 2005). Recent data have estimated that first parent concerns occur at an even earlier age (average of 14–15 months), with a significant number below 11 months (Chawarska et al., 2007). There also appears to be a significant number of children who experience regression between

ages 2 and 3 (Tuchman and Rapin, 1997). Thus, screening before the age of 2 might miss these children with later onsets, but the evidence presented above suggests that many children would already be demonstrating behaviors associated with autism.

An additional difficulty with early screening is the possibility of a higher false-positive rate for the youngest children and for those children provided diagnoses of PDD-NOS (Charwarska et al., 2007). While few data have been reported on this group of families, parents of children inaccurately suspected of a developmental disorder may suffer unnecessary distress (Williams and Brayne, 2006). Although screening data on very young children are lacking, the fact that some children may show significant developmental concerns at a very young age which then resolve suggests that false-positive rates on screening may be higher for this very young group (Dietz et al., 2006).

The M-CHAT was described by Robins et al. (2001), who reported on a screening of 1293 children (see 'Methods' for description). Fine et al. (2005) used the M-CHAT successfully to screen children with 22q11.2 deletions for autism. Eaves et al. (2006) examined the performance of the M-CHAT with a group of 84 children aged 24–48 months (mean age 37 months) referred for possible autism to a specialty clinic, of whom 64 percent were then diagnosed with ASD. The majority of the remaining children had more than one diagnosis, including intellectual delay and language disorder. Sensitivity was good: for the 2/6 critical item score the sensitivity was 77 percent, and for the 3/23 item score it was 92 percent. However, specificity was low (43% and 27% for the two scores). It should be noted, however, that the follow-up interview to reduce false positives was not available at the time of this study.

The M-CHAT (and the CHAT) were translated into Chinese, and tested on a sample of 212 children with mental ages 18–24 months, about half of whom were diagnosed with ASD (Wong et al., 2004). The seven most discriminating items were largely overlapping with, but not identical to, the six critical items on the M-CHAT identified by Robins et al. (2001). Using a cutoff of failing two of these seven items produced a sensitivity of 0.93 and a specificity of 0.77, whereas failing any six of the 23 items produced a sensitivity of 0.84 and a specificity of 0.85. Mawle and Griffiths (2006) reviewed the available data and suggest that the M-CHAT has promising sensitivity for population screening but that additional follow-up data are needed on the initial sample.

Such follow-up data were recently published (Kleinman et al., 2007). Kleinman and colleagues reported on a replication with a new 2-year-old sample, as well as a follow-up to age 4, where prediction continued to be good. However, as might be expected, positive predictive power (PPP) for

the M-CHAT was found to be higher for children already suspected of a developmental disorder but not yet diagnosed (high risk) than for a low-risk, general pediatric screening sample (Kleinman et al., 2007). However, this article did not separate the children by age group. Given the somewhat different presentation expected for children under 2, as well as the possibility of more false positives for the youngest children, the current article examines screening in younger versus older toddlers from high- and low-risk samples.

The purpose of the current study was to explore two questions in these groups. First, is the false-positive rate significantly different for groups stratified by age and risk: older/high-risk versus younger/high-risk samples and older/low-risk versus younger/low-risk samples? Second, of the children who screen positive and then meet criteria for ASD on evaluation, does the clinical picture differ for the younger versus older and high-risk versus low-risk groups of children?

Methods

Sample

The study sample was drawn from a larger sample of families participating in a federally funded study of the M-CHAT screening tool. Procedures were approved by the University of Connecticut and Yale School of Medicine Institutional Review Boards. Participating toddlers were split into two age groups: younger (n screened = 4592, 16–23 months, mean 18.73 months) and older (n screened = 2184, 24–30 months, mean 25.12 months). Toddlers screened through the M-CHAT study are typically screened through either their early intervention or the pediatrician's office. In the former case, developmental concerns had been expressed about the child by parent or pediatrician, but no diagnosis had yet been made and no more than minimal services (1–2 hours per week for several weeks) had been provided. These children were designated as high risk (screened by early intervention providers), whereas the unselected well-child visit screening sample was designated as low risk. This resulted in four groups: younger/high-risk (n = 327, mean age = 20.72 months), younger/low-risk (n = 4265, mean age = 18.57 months), older/high-risk (n = 399, mean age = 26.84 months), older/low-risk (n = 1785, mean age = 24.74 months). Gender for the screening sample is shown in Table 1.

All children were recruited through a screening program run by the University of Connecticut. The current study includes the entire sample of children reported in the recent article by Kleinman et al. (2007) along with 2983 more recently screened children.

Table 1 Demographics of screening sample

	Total younger	Total older	Younger/ high-risk	Older/ high-risk	Younger/ low-risk	Older/ low-risk
Total screened	4592	2184	327	399	4265	1785
Boys screened (%)	2368 (51%)	1216 (56%)	227 (69%)	301 (75%)	2141 (50%)	915 (51%)
Girls screened (%)	2150 (47%)	949 (43%)	94 (29%)	94 (24%)	2056 (48%)	855 (48%)
Sex unreported (%)	74 (2%)	19 (1%)	6 (2%)	4 (1%)	68 (2%)	15 (1%)
Age at time of M-CHAT (SD)	18.73 (1.23)	25.12 (1.59)	20.72 (1.87)	26.84 (1.79)	18.57 (1.02)	24.74 (1.25)
Total M-CHAT score (SD) (0-23)	1.10 (2.21)	1.46 (2.62)	5.02 (5.60)	3.65 (4.36)	0.80 (1.26)	0.97 (1.69)
Critical M-CHAT score (SD) (0-6)	0.21 (0.80)	0.32 (0.99)	1.70 (2.10)	1.19 (1.73)	0.09 (0.40)	0.13 (0.57)

Of the 327 younger/high-risk toddlers screened, 164 screened negative; 29 screened positive, were given the follow-up interview and passed (screened negative); five families refused the follow-up interview and nine could not be contacted (i.e. either moved away or did not return multiple calls); and 95 failed the follow-up interview and were offered evaluations. Thus, of the 138 who needed a follow-up interview, 14 (10%) could not be contacted or refused. In addition, 25 children were offered evaluations without the follow-up interview, either because this was before the follow-up interview was instituted or because of heightened concern (i.e. high number of items failed or early interventionists requesting rapid evaluation). Of these 120 children who qualified for evaluations (95 + 25), 107 (89%) came for evaluation (mean age at evaluation 24 months) and 13 could not be reached or refused (see Figure 1 for a flow chart of sample flow/loss).

Of the 399 older/high-risk toddlers screened, 233 screened negative; 48 were given the follow-up interview and passed; three families refused the follow-up interview and nine could not be contacted. Eighty-two children failed the follow-up interview and were offered evaluations, and 24 children were offered evaluations without the follow-up interview. Thus, of the 142 who needed a follow-up interview, 12 (8%) could not be contacted or refused. Of these 106 children who qualified for evaluations, 96 (91%) came for evaluation (mean age at evaluation 30 months) and 10 could not be reached or refused (see Figure 2).

Of the 4265 younger/low-risk toddlers screened, 3971 screened negative; 206 were given the follow-up interview and passed; 11 families refused the follow-up interview and 20 could not be contacted; and 42 failed the follow-up interview and were offered evaluations. Thus, of the 279 who needed a follow-up interview, 31 (11%) could not be contacted or refused.

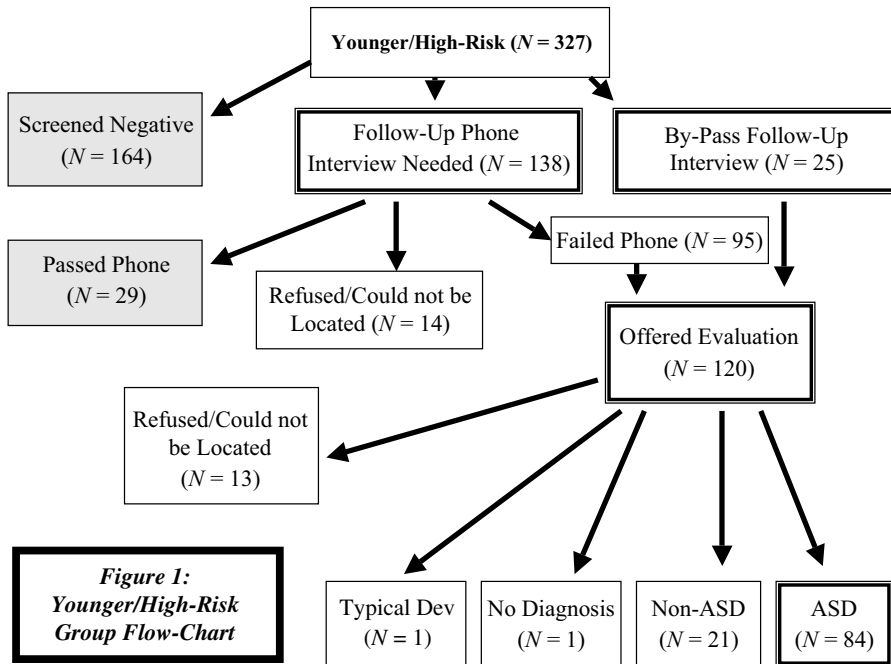


Figure 1

In addition, 15 children were offered evaluations without the follow-up interview, for reasons noted above. Of these 57 children who qualified for evaluations, 36 (63%) came for evaluation (mean age at evaluation 24 months) and 21 could not be reached or refused (see Figure 3).

Of the 1785 older/low-risk toddlers screened, 1646 screened negative; 81 were given the follow-up interview and passed; six families refused the follow-up interview and 13 could not be contacted. Twenty-nine children failed the follow-up interview and were offered evaluations, and 10 children were offered evaluations without the follow-up interview. Thus, of the 129 who needed a follow-up interview, 19 (15%) could not be contacted or refused. Of these 39 children who qualified for evaluations, 31 (80%) came for evaluation (mean age at evaluation 30 months) and eight could not be reached or refused (see Figure 4).

Chi-square analyses revealed no significant differences in rate of refusal for either the follow-up interview or the evaluation between the two age groups for either the high-risk or the low-risk samples, although families of the younger/low-risk children were more likely to refuse evaluation than families of either high-risk group ($\chi^2 = 16.84$, d.f. = 1, $p < 0.01$ relative to younger/high-risk group and $\chi^2 = 18.08$, $p < 0.01$ relative to older/high-risk group, both with Bonferroni correction for six pairwise comparisons).

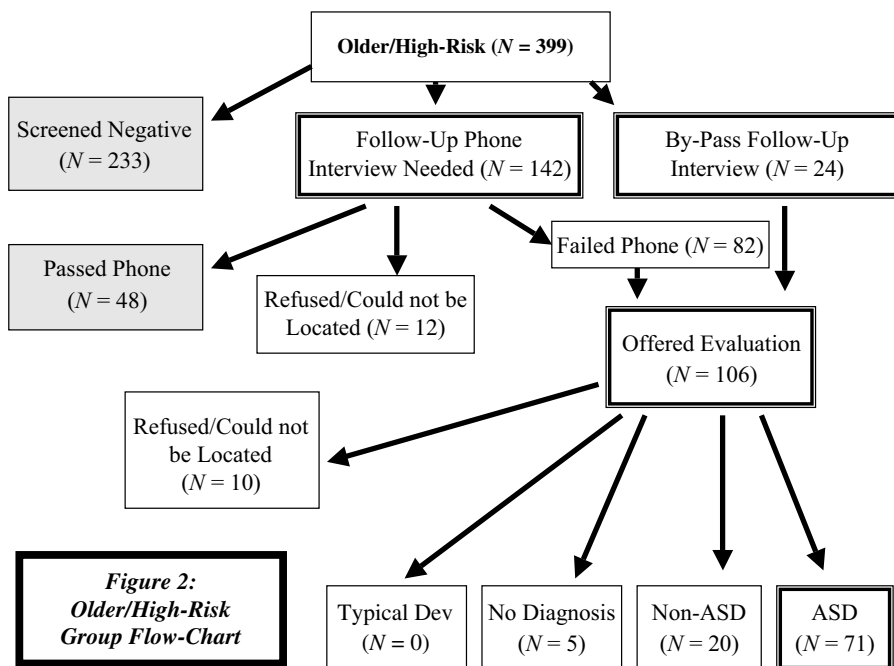


Figure 2

Children who screened positive were offered a free developmental/ diagnostic evaluation. The sample of evaluated children was then further divided by children who were diagnosed with an autism spectrum disorder (ASD) and those who were given another developmental diagnosis or no diagnosis. ASD is used to refer to children meeting DSM-IV-TR criteria for Autistic Disorder or pervasive developmental disorder not otherwise specified (American Psychiatric Association, 2000). Asperger’s disorder was not considered because of the young age of the sample. All children were evaluated using the same diagnostic measures: Autism Diagnostic Interview–Revised (ADI–R), Autism Diagnostic Observation Schedule (ADOS), clinical judgment based on DSM-IV-TR criteria for Autistic Disorder, and Childhood Autism Rating Scale (CARS). See procedures below for instrument descriptions and evaluation procedure.

Measures

The M–CHAT (Robins et al., 2001) This is a 23-item yes–no parent report screener for ASD. Initial failure on the screener is defined as any three items failed, or any two critical items failed. The critical items were identified by discriminant function analysis of children with and without

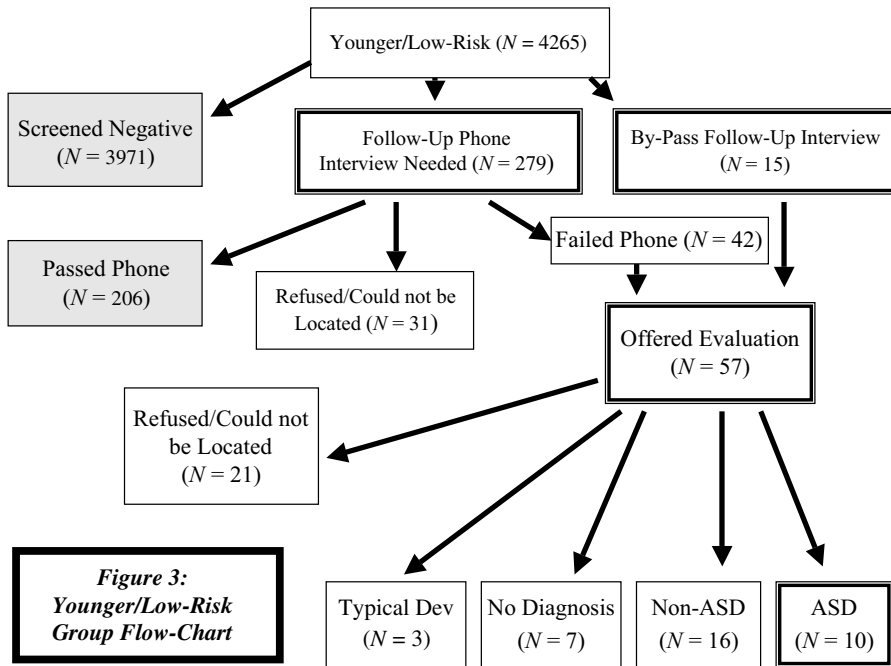


Figure 3

a disorder on the autism spectrum (Robins et al., 2001) and included items concerning joint attention (proto-declarative pointing, bringing to show, following a point), interest in other children, responding to name, and imitation. The scale, plus the follow-up interview, can be found at: <http://www2.gsu.edu/~wwwpsy/faculty/robins.htm>.

The Autism Diagnostic Observation Schedule–Generic (ADOS–G: Lord et al., 1999) This consists of semi-structured assessments (modules 1 and 2) of communication, social interactions and relatedness, play, and imagination. The assessments consist of 10 planned social interactions to encourage social initiations, responses, and opportunities for communication. The child is also given opportunities to engage in make-believe and imaginative play. On these measures, the child receives a score in the social domain, in the communication domain, and in the combined social and communication domains. Diagnostic classification is made by exceeding cutoff scores in these three areas (social, communication, and combined). A child can be classified as having Autistic Disorder, as having pervasive developmental disorder not otherwise specified, or as being non-autistic.

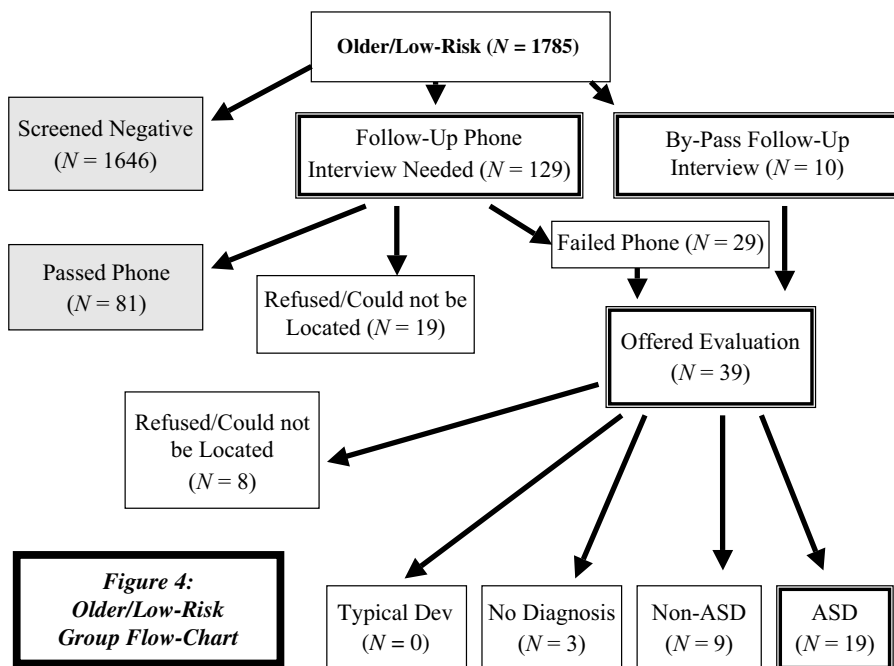


Figure 4

The Autism Diagnostic Interview–Revised (Rutter et al., 2005) This is a semi-structured clinician-based interview for parents or caregivers that evaluates the child’s communication, social development, play, and restricted, repetitive, and stereotyped behaviors. The interviewer scores each of the measure’s 123 questions with a 0–3 based on the severity of the behavior (0 = no behavior of this type and 3 = very severe behavior). For this study, we used the Toddler Edition, obtained from the authors. The ADI–R Toddler Edition has a scoring algorithm that is based on the DSM–IV and ICD–10 criteria for autism. The interview yields separate scores for each of three diagnostic domains (social interactions, communication, and repetitive and stereotyped behaviors). In order to meet diagnostic criteria for autism, a child has to meet the scoring criteria in each of the three domains separately. The ADI–R algorithm yields a classification of either Autistic Disorder or non-autistic; it does not consider PDD–NOS a possible diagnosis.

The Childhood Autism Rating Scale (CARS: Schopler et al., 1988) This consists of 15 items intended to measure the presence and severity of symptoms of pervasive developmental disorders. The child is rated on each

item based on the clinician's observation of the child's behavior throughout the testing as well as on the parent's report. The CARS includes items on socialization, communication, emotional responses, and sensory sensitivities. The clinician scores each of the 15 items from 1 to 4, with 1 indicating no impairment and 4 indicating severe impairment. Based on the child's combined score from the 15 items, he or she can be classified with mild, moderate, or severe autism, or no autism, with a cutoff of 30 for autism.

Clinical judgment Clinical judgment by experienced clinicians is considered to be the 'gold standard' for autism diagnosis (Klin et al., 2000; Spitzer and Siegel, 1990). In this study, the clinicians used the DSM-IV-TR criteria for Autistic Disorder (American Psychiatric Association, 2000) on which to base their clinical judgments. A diagnosis of Autistic Disorder or PDD-NOS was given if the child met the necessary DSM-IV-TR criteria.

The Vineland Adaptive Behavior Scales (Sparrow et al., 1984) This is a widely used parent interview that assesses adaptive functioning in the areas of socialization, communication, daily living, and motor skills. The clinician codes each item from 0 to 2, with 0 meaning that the child did not perform the particular behavior, 1 meaning that the child sometimes performed the behavior or that the skill was emerging, and 2 indicating that the child performed the behavior on a regular basis.

The Mullen Scales of Early Learning (Mullen, 1995) This is a standardized, individually administered test that assesses a child's level of cognitive development. The test measures ability in five domains: gross and fine motor, receptive and expressive language, and visual problem solving. All domains except gross motor were used to calculate an overall cognitive standard score (IQ), and only these four domains were administered.

Procedure

Younger and older toddlers from high- and low-risk sources were recruited from various pediatricians and early intervention offices throughout the states of Connecticut, Rhode Island, and Massachusetts. All children in the current study were screened by the Modified Checklist for Autism in Toddlers screener (M-CHAT: Robins et al., 2001).

Once a child screened positive, the family was contacted for a telephone follow-up. This conversation followed a script with specific examples in which all failed items were reviewed with a parent. When reviewing the failed responses with the parents, trained interviewers did not use the term 'fail'. Specific questions were repeated and parents were asked whether

their child's behavior was still present or had resolved since completing the screener. Parents were provided concrete examples of the behaviors targeted in each of the M-CHAT questions in order to help clarify responses. If the child continued to fail the M-CHAT after the telephone follow-up based on parent responses to increasingly specific probes, the family was told that their child was not doing some things that most other children of the same age were doing, and that an evaluation was recommended. If the family agreed to an evaluation, the child was given a free developmental and diagnostic evaluation. Since all of the children presented for evaluation because of failing the M-CHAT, some degree of risk for developmental disorder was present. This risk was known to the examiners, excluding the possibility of a completely blind assessment.

The evaluations took place at the Psychological Services Clinic at the University of Connecticut or at the Yale University School of Medicine Child Study Center. Evaluations were completed by a team of clinicians consisting of one licensed psychologist or developmental pediatrician who specialized in autism and one graduate or postdoctoral student. In almost all cases, parents and children stayed in the same room for the evaluation, allowing both evaluators to observe the child's behavior. Following the evaluation, both clinicians completed the CARS. The CARS completed by the psychologist or developmental pediatrician was used in the data analysis.

For a diagnosis of an ASD, the child's scores on the various measures were considered, but the final diagnosis was determined by the judgment of the clinician (Klin et al., 2000; Spitzer and Siegel, 1990). The 'non-ASD diagnosis' group represents those children who were diagnosed with a global developmental delay or developmental language disorder. Children were provided a diagnosis of global developmental delay if they demonstrated delays of at least 1.5 standard deviations on one of the 'non-language' domains of standardized testing (Mullen visual reception and fine motor and Vineland motor skills) and a delay of at least 1.5 standard deviations on at least one of the 'language' domains of standardized testing (Mullen expressive language and receptive language and Vineland communication). At least one of the delays in the 'non-language' or 'language' domains had to be demonstrated on the Mullen, i.e. not just by parent report. Children were given a diagnosis of developmental language disorder if they demonstrated delays of at least 1.5 standard deviations on at least two of the Mullen expressive language, Mullen receptive language, and Vineland communication, or a delay of at least 2 standard deviations on any one of these three language measures. Additionally, children in this diagnostic group could not have greater than 1.5 standard deviation delays on any other subscale. Children were labeled with 'no diagnosis' if they did not reach diagnostic threshold on the DSM-IV-TR or on formal assessment tools

(the Mullen or Vineland), but were noted to have mild difficulties with language, motor, social, or cognitive skills that warranted some intervention or monitoring, or that raised clinicians' concerns. Other children were determined to be functioning typically (Table 2).

For analyses of positive predictive power, the groups were compared by age within risk group, yielding two comparisons: older/high-risk versus younger/high-risk, and older/low-risk versus younger/low-risk. For the evaluation variables exploring developmental and diagnostic differences, the four groups were compared to each other in a 2×2 design (younger versus older, low versus high risk) and only those children diagnosed with ASD were included.

Table 2 Diagnostic classification of evaluated children

Diagnostic classification	Total younger (<i>n</i> = 143)	Total older (<i>n</i> = 127)	Younger/ high-risk (<i>n</i> = 107)	Older/ high-risk (<i>n</i> = 96)	Younger/ low-risk (<i>n</i> = 36)	Older/ low-risk (<i>n</i> = 31)
Autistic disorder	49	44	44	40	5	4
PDD-NOS	45	46	40	31	5	15
Non-ASD diagnosis	37	29	21	20	16	9
No diagnosis	8	8	1	5	7	3
Typical development	4	0	1	0	3	0
PPP for ASD	0.66	0.79	0.79	0.74	0.28	0.61
PPP for diagnosable dev. dis.	0.92	0.94	0.98	0.95	0.72	0.90

Results

Mean M-CHAT total and critical item scores, age, and results of evaluations are shown in Tables 1 and 2. Positive predictive power (PPP) is calculated as number of screen-positive cases diagnosed with ASD divided by total number of screen-positive cases, and is an estimate that a child screening positive will in fact be diagnosed with an ASD (see Table 2).

Of the 107 evaluated children in the younger/high-risk sample, 84 were diagnosed with ASD, for a PPP of 0.79 (95% CI = 0.713 to 0.867). Of the 96 evaluated children in the older/high-risk sample, 71 were diagnosed with an ASD, for a PPP of 0.74 (95% CI = 0.652 to 0.828). These PPPs did not differ (χ^2 (d.f. = 1) = 0.58, n.s.).

Of the 36 evaluated children in the younger/low-risk sample, 10 were diagnosed with an ASD, for a PPP of 0.28 (95% CI = 0.133 to 0.427). Of the 31 evaluated children in the older/low-risk group, 19 were diagnosed with an ASD, for a PPP of 0.61 (95% CI = 0.438 to 0.782). The PPP for the younger group was significantly lower than the PPP for the older group (χ^2 (d.f. = 1) = 7.62, $p < 0.01$).

PPP for any diagnosable DSM-IV developmental disorder, as defined above and not including those children diagnosed with ‘no diagnosis’ or determined to be typically developing, was: younger/high-risk 0.98, older/high-risk 0.95 (not significantly different), and younger/low-risk 0.72, older/low-risk 0.90 (χ^2 (d.f. = 1) = 3.49, $p = 0.06$) (see Table 2).

Analysis of variance (ANOVA), with risk level (high versus low) and age group (younger versus older) as between-subject factors, was used to compare the four groups diagnosed with ASD on the following autism measures: M-CHAT total and critical scores (see Table 3), CARS total score, DSM-IV-TR total score, ADOS sum of communication and social interaction scores, and ADI total score (see Table 4), and chi-square was used to compare the frequency of diagnoses of autistic disorder versus PDD-NOS (see Table 2). The four groups were also compared on the following developmental measures: Mullen Scales of Early Learning (visual reception, fine motor, receptive language, and expressive language), and Vineland Adaptive Behavior Scales (communication, daily living, socialization, and motor skills). Partial eta-squared (η_p^2) was used to measure effect size (Olejnik and Algina, 2003) (see Table 5). As a measure of effect size in factorial

Table 3 Demographics of evaluation sample diagnosed with ASD

	Younger/ high-risk (ASD)	Older/ high-risk (ASD)	Younger/ low-risk (ASD)	Older/ low-risk (ASD)	Interaction effect F-value	η_p^2
Total evaluated	84	71	10	19		
Boys evaluated	65 (77%)	60 (85%)	9 (90%)	16 (84%)		
Girls evaluated	19 (23%)	11 (15%)	1 (10%)	3 (16%)		
Age at M-CHAT	20.83 (2.11)	27.17 (1.66)	18.97 (1.96)	26.59 (1.93)	9.060 ^{a*}	0.048 ^a
M-CHAT total	11.94 (4.04)	9.87 (3.63)	8.10 (4.04)	9.95 (5.28)	5.281 ^{**}	0.029
M-CHAT critical items	4.42 (1.30)	3.55 (1.50)	2.80 (1.32)	3.58 (1.95)	7.146 [*]	0.038

* $p < 0.01$; ** $p < 0.05$.

^a This value is a main effect for risk status, as the groups are selected to be significantly different on age and there was not a significant interaction effect.

Table 4 Age by risk evaluation comparisons: diagnostic measures

Diagnostic evaluation measures	Younger/ high-risk (n = 84)	Older/ high-risk (n = 71)	Younger/ low-risk (n = 10)	Older/ low-risk (n = 19)	Main effect: risk status F-value	η_p^2
ADI Toddler total	27.68 (6.68)	26.45 (6.95)	27.60 (4.50)	25.00 (11.67)	n.s.	
ADOS (a)(b) total	16.01 (4.53)	15.48 (4.67)	16.30 (2.50)	13.88 (4.60)	n.s.	
CARS total	33.78 (5.30)	32.99 (5.07)	31.38 (5.32)	31.47 (6.32)	n.s.	
DSM-IV-TR total	6.11 (1.86)	6.11 (1.84)	5.70 (1.64)	4.94 (1.84)	3.889 ^{**}	0.023

* $p < 0.01$; ** $p < 0.05$.

Table 5 Age by risk evaluation comparisons: cognitive and adaptive measures

<i>Evaluation measures</i>	<i>Younger/ high-risk (n = 84)</i>	<i>Older/ high-risk (n = 71)</i>	<i>Younger/ low-risk (n = 10)</i>	<i>Older/ low-risk (n = 19)</i>	<i>Main effect: η_p^2 age F-value</i>	
<i>Mullen T-scores</i>						
Visual reception	29.16 (9.26)	26.93 (8.60)	30.70 (13.41)	26.31 (10.57)	ns	
Fine motor	29.97 (11.59)	25.89 (8.25)	29.60 (8.58)	26.19 (10.24)	ns	
Receptive language	21.21 (5.06)	23.61 (8.15)	21.00 (3.16)	21.31 (3.30)	ns	
Expressive language	24.13 (6.11)	24.07 (8.48)	23.10 (3.63)	22.44 (5.02)	ns	
<i>Vineland standard scores</i>						
Communication	65.17 (5.45)	64.37 (8.40)	66.90 (3.90)	62.88 (7.71)	ns	
Daily living	70.64 (7.27)	65.04 (5.01)	73.40 (4.81)	65.47 (5.35)	25.842*	0.128
Socialization	69.02 (7.73)	65.17 (7.74)	71.60 (6.95)	66.88 (7.27)	6.765*	0.037
Motor skills	82.95 (10.83)	78.91 (11.64)	85.60 (12.48)	77.88 (15.25)	5.452**	0.030

* $p < 0.01$; ** $p < 0.05$.

designs, Olejnik and Algina (2003) have suggested replacing partial eta-squared with a new measure they call generalized eta-squared. They argue that the new measure is less influenced by the particulars of the research design, but they also point out that in a between-subjects design with factors that are measured (i.e. not manipulated), their new measure is equal to the more commonly used partial eta-squared. In order to avoid confusion, we have chosen to label this measure partial eta-squared, but we urge readers to consult Olejnik and Algina (2003) for a review of effect size measures in common experimental designs.

The younger and older groups differed significantly by age, as they were constituted to do; however, the time between screening and evaluation (approximately 3 months) was consistent between the groups.

The overall chi-square among the four groups for the proportion of children with ASD receiving a diagnosis of AD versus PDD-NOS was not significantly different (χ^2 (d.f. = 3) = 7.68, $p = < 0.1$).

Results of the ANOVA comparison of the four groups indicated significant risk and age \times risk effects in critical and total number of M-CHAT items failed (see Table 3). Examination of means shows that the effect of age was different in the risk groups, with the number of items failed going up with age in the low-risk group and down with age in the high-risk groups, and the younger/low-risk children having the lowest means. Effect sizes, however, were small.

CARS, ADI, and ADOS scores showed no significant main effects or interaction, and all effect sizes were very small. Number of DSM-IV-TR symptoms did show a risk effect, with the high-risk groups displaying more symptoms; the effect size was small ($\eta_p^2 = 0.023$) (see Table 4). On

developmental measures (see Table 5), there were no significant main effects for risk or age group or their interaction on any Mullen domain score, and all effect sizes were very small. There were also no significant effects for Vineland communication ($\eta_p^2 = 0.015$). The Vineland domains of daily living, socialization, and fine motor skills showed significant age effects, with the younger age group showing higher scores; daily living showed a medium effect size ($\eta_p^2 = 0.128$), and socialization and motor domains showed small effect sizes (0.037 and 0.03, respectively).

Discussion

The main purpose of the current study was to examine the outcome of screening younger versus older groups of toddlers, drawn from high- and low-risk samples. Diagnostic and developmental differences among children diagnosed with ASD were also assessed for the following four groups of children: younger/low-risk, older/low-risk, younger/high-risk, and older/high-risk. With regard to family participation, follow-up interview refusal rates did not differ by group, but the parents of the younger/low-risk children were more likely to refuse an offered evaluation than parents of children in the two high-risk groups.

Comparisons of PPP for ASD between age groups show that it is lowest for the younger/low-risk toddlers (0.28), those below the age of 24 months screened routinely at a well-child visit; their PPP was significantly lower than the older/low-risk toddlers (0.61). For children already identified to be at some developmental risk, PPP (0.79 for younger children and 0.74 for older children) did not differ by age. It is to be expected that PPP would be higher for the children where developmental concerns, possibly indicative of potential ASD, are already known to exist, and it appears that PPP for these children, even for the younger group, is relatively high. These findings are consistent with those of Kleinman and colleagues (2007), who found lower PPP for low-risk children (across both age groups); the current study included the Kleinman sample but added 2983 children. For the lower-risk groups, it is possible that some of the younger toddlers have mild or transient developmental delays that resolve, resulting in the lower predictive power. The PPP for the older/low-risk group may be higher than that of the younger children because delays in social communication in this age group (24 months and older) are less likely to be benign and transient. It should be noted that these PPP values assume that the follow-up interview was administered in order to reduce the false-positive rates, which was demonstrated in the Kleinman and colleagues (2007) article to be very important. The interview is available free of charge on the Internet (see above). In addition, the PPP for the younger/low-risk group may be the

lowest because the overall M-CHAT critical and total scores are significantly lower for this group, and therefore their degree of risk for ASD, as measured by M-CHAT scores, may be lower.

The PPP rates of our study suggest that clinical decisions about age of screening should balance the earliest possible identification against the possibility of unnecessary referrals for each clinical application. Whether the PPP of 0.61 for the older/low-risk children is acceptable is a clinical decision that could be made based on the degree of clinician concern about the child, the time that will elapse before another screening is planned, the likelihood of harm to the family from an inaccurate positive screen, and the availability of early intervention. Although the PPP of 0.28 for the younger/low-risk group is low, it is important to note that 72 percent of the screen-positive children in this group were diagnosed with either an ASD, a language delay, or a global developmental delay, and that only 8 percent of this group were judged to be fully typical. Therefore, unnecessary alarm is not likely to be a serious risk for the large majority of even the younger/low-risk children, especially if parents are told that a positive screen is indicative of the need for further assessment, not of a specific diagnosis.

Although the stability of these diagnoses for this age group are not known, we used specific criteria, similar to those used by many early intervention providers, for these diagnoses, and it is likely that the children needed and could benefit from intervention services in the areas on which they showed delays. Thus, the M-CHAT seems to have limited specificity for the youngest children for whom no prior developmental concerns were noted, detecting ASD but also other developmental conditions. When one considers the other three groups (both older groups and the younger/high-risk group), the overwhelming majority of the screen-positive children did qualify for a developmental diagnosis; therefore, for these children, it seems clear that the risks of a false positive in a typical child, and consequent unnecessary alarm, are low, and this risk is relatively low even for the youngest, low-risk group.

The second major purpose of the study was to compare the children diagnosed with ASD from the four age \times risk groups to see if they differed in diagnostic severity or degree of developmental delay. With regard to scores on the autism-specific measures, no significant differences in severity of symptomatology were found across the autism measures, with the exception of the DSM-IV-TR symptoms, in which the high-risk groups demonstrated slightly but significantly more symptoms than the low-risk groups. In addition, M-CHAT total and critical scores were slightly lower in the younger/low-risk toddlers than in the older/low-risk groups, while the opposite trend held for the high-risk children. Both of these findings, however, showed small effect sizes. For many of the remaining diagnostic

variables, however, the four toddler groups with ASD received very similar scores with regard to classification and severity. This evidence suggests that the four groups are similar in severity of clinical presentation, although specific items might differ between groups.

Within the domains of cognitive functioning, there were no group differences. On adaptive functioning, the younger children tended to receive higher standard scores, but there was no difference by risk. We suggest that the most likely explanation for these differences, which were small, rests in the psychometric properties of the tests, where older children have a greater range of items administered, resulting in a greater range of possible scores. Another possibility is that the nature of the earliest items in social, daily living, and motor scores are within the capacity of children with ASD, but that later items become more cognitively complex or socially demanding. For example, early social items include social orientation to parents, which is shown by most children with ASD, whereas later ones focus more on social orientation to peers, which is more problematic. Therefore, the higher social skills expected of the older children would result in lower standard scores. It is also possible that some of the pairwise comparisons between the groups may have reached significance by chance and do not reflect differences in developmental functioning.

Our results suggest that older and younger toddlers detected by autism screening have a similar degree of symptomatology on the CARS, ADI, and ADOS. It is possible that diagnostic measures designed specifically for use with these young children might be more effective in detecting age differences related to autism symptomatology (although we used the ADI Toddler Form, the algorithm items remain the same as on the standard version). It is also possible that a detailed investigation of specific autism diagnostic items, rather than the use of a total score, may reveal specific differences between symptoms in younger versus older toddlers. To our knowledge, the literature has not investigated specific symptom differences in toddler populations based on referral risk, which may also be useful.

The current sample was limited by a relatively small sample size, particularly in the number of children who were evaluated from the low-risk samples (57 younger, 45 older). While the small sample sizes of these groups are reflective of the screen-positive rates in the general population, it is possible that larger samples would change the PPP values. Another limitation of the study is the fact that the stability of specific developmental delays in this young age group are not known, although prior data on our sample (Kleinman et al., 2008; Suter et al., 2007) suggest that these diagnoses are relatively stable, and in particular, that they do not tend to evolve into a diagnosis of ASD at a later age. Our finding of higher refusal rates for the evaluation in the younger/low-risk children is consistent with that of Dietz

and colleagues (2006), who found high evaluation refusal rates with an even younger unselected sample from the Netherlands. It is possible that 24-month-old children with delays are more likely to raise parental concerns than 18-month-old children, and therefore that evaluation might be more acceptable at that age and that parents may be more likely to comply with the screening and evaluation process. These findings may indicate that parents of very young, low-risk children might need more of an introduction to, or physician support through, the screening process. If such facilitation improves participation in autism screening programs, it may also be effective for improving participation in community screening programs for other medical conditions. As developmental and autism-specific screening becomes more standard practice at well-child visits, as recommended by the AAP (Johnson et al., 2007), pediatricians may be more able to prepare parents for routine screening and follow-up evaluations.

Another limitation is that the assessments were not done blind with respect to screening status, which would be highly desirable given the research context. We did rely heavily on results of standardized testing and structured diagnostic tests (e.g. ADOS), lessening the probability of bias in diagnosis, but replication of results with a blind assessment would be a strong confirmation. It is important to note that in many of the settings in which the M-CHAT will be used, clinicians are not blind to the presenting symptoms of the child. A final limitation is that we did not have the resources to follow the large number of children who screened negative, and therefore data about differential sensitivity for the two age groups are lacking. Kleinman and colleagues (2007) did attempt to ascertain cases of ASD that were missed by the M-CHAT, and found relatively few, but these missed cases were not examined for age at initial screening.

Overall, our results support the efficacy of ASD screening in young children, as recommended by the AAP, although with less specificity for ASD in the younger/low-risk children. The majority of children in all groups identified by the screening process were found to have a diagnosable developmental delay of some type, suggesting the need for intervention or at least heightened surveillance. Continued study of the emergence of ASD in children at risk may help to hone the screening items that are appropriate for each age group, and increase the PPP of autism screening.

References

- AMERICAN ACADEMY OF PEDIATRICS (2006) 'Identifying Infants and Young Children with Developmental Disorders in the Medical Home: An Algorithm for Developmental Surveillance and Screening', *Pediatrics* 118: 404–19.
- AMERICAN PSYCHIATRIC ASSOCIATION (2000) *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn text rev. (DSM-IV-TR). Washington, DC: APA.

- BAIRD, G., CHARMAN, T., BARON-COHEN, S., COX, A., SWETTENHAM, J., WHEELWRIGHT, S. & DREW, A. (2000) 'A Screening Instrument for Autism at 18 Months of Age: A Six-Year Follow-Up Study', *Journal of the American Academy of Child and Adolescent Psychiatry* 39: 694–702.
- BAIRD, G., CHARMAN, T., COX, A., BARON-COHEN, S., SWETTENHAM, J., WHEELWRIGHT, S. & DREW, A. (2001) 'Current Topic: Screening and Surveillance for Autism and Pervasive Developmental Disorders', *Archives of Disease in Childhood* 84: 468–75.
- BARANEK, G. (1999) 'Autism During Infancy: A Retrospective Video Analysis of Sensory-Motor and Social Behaviours at 9–12 Months of Age', *Journal of Autism and Developmental Disorders* 29: 213–24.
- BRYSON, S.E., ZWAIGENBAUM, L., BRIAN, J., ROBERTS, W., SZATMARI, P., ROMBOUGH, V., ET AL. (2007) 'A Prospective Case Series of High-Risk Infants Who Developed Autism', *Journal of Autism and Developmental Disorders* 37 (1): 12–24.
- CHARMAN, T. & BAIRD, G. (2002) 'Practitioner Review: Diagnosis of Autism Spectrum Disorder in 2- and 3-Year-Old Children', *Journal of Child Psychology and Psychiatry* 43: 289–305.
- CHAWARSKA, K., KLIN, A., PAUL, R., & VOLKMAR, F. (2007) 'Autism Spectrum Disorder in the Second Year: Stability and Change in Syndrome Expression', *Journal of Child Psychology and Psychiatry* 48: 128–38.
- COURCHESNE, E., CARPER, R., AKSHOOMOFF, N. (2003) 'Evidence of Brain Overgrowth in the First Year of Life in Autism', *Journal of American Medical Association* 290 (3): 337–44.
- DEMENTIEVA, Y., VANCE, D., DONNELLY, S., ELSTON, L., WOLPERT, C., RAVAN, S., DELONG, G., ABRAMSON, R., WRIGHT, H. & CUCCARO, M. (2005) 'Accelerated Head Growth in Early Development of Individuals with Autism', *Pediatric Neurology* 32 (2): 102–8.
- DIETZ, C., SWINKELS, S., VAN DAALLEN, E., VAN ENGELAND, H. & BUITELAAR, J.K. (2006) 'Screening for Autistic Spectrum Disorder in Children Aged 14–15 Months. II: Population Screening with the Early Screening of Autistic Traits Questionnaire (ESAT). Design and General Findings', *Journal of Autism and Developmental Disorders* 36: 713–22.
- EAVES, L.C., WINGERT, H. & HO, H.H. (2006) 'Screening for Autism: Agreement with Diagnosis', *Autism* 10 (3): 229–42.
- FILIPEK, P., ACCARDO, P., BARANEK, G., COOK, E. JR, DAWSON, G., GORDON, B., GRAVEL, J., JOHNSON, C., KALLEN, R., LEVY, S., MINSHEW, N., PRIZANT, B., RAPIN, I., ROGERS, S., STONE, W., TEPLIN, S., TUCHMAN, R. & VOLKMAR, F. (1999) 'The Screening and Diagnosis of Autism Spectrum Disorders', *Journal of Autism and Developmental Disorders* 29 (6): 439–75.
- FINE, S.E., WEISSMAN, A., GERDES, M., PINTO-MARTIN, J., ZACKAI, E.H., MCDONALD-MCGINN, D.M. & EMANUEL, B.S. (2005) 'Autism Spectrum Disorders and Symptoms in Children with Molecularly Confirmed 22q11.2 Deletion Syndrome', *Journal of Autism and Developmental Disorders* 35 (4): 461–70.
- GILLBERG, C., EHLERS, S., SCHAUMANN, H., JOACOBSEN, G., DAHLGREN, S., LINDBLOM, R., BAGENHOLM, A., TJUUS, T. & BLINDER, B. (1990) 'Autism under Age 3 Years: A Clinical Study of 28 Cases Referred for Autistic Symptoms in Infancy', *Journal of Child Psychology and Psychiatry* 6: 921–34.
- GOLDBERG, W.A., OSANN, K., FILIPEK, P.A., LAULHERE, T., JARVIS, K., MODAHL, C., FLODMAN, P. & SPENCE, M.A. (2003) 'Language and Other

- Regression: Assessment and Timing', *Journal of Autism and Developmental Disorders* 33 (6): 607–16.
- HORNER, R., CARR, E., STRAIN, P., TODD, A. & REED, H. (2002) 'Problem Behavior Interventions for Young Children with Autism: A Research Synthesis', *Journal of Autism and Developmental Disorders* 32 (5): 423–46.
- JOHNSON, C., MYERS, S. & THE COUNCIL ON CHILDREN WITH DISABILITIES OF THE AMERICAN ACADEMY OF PEDIATRICS (2007) 'Identification and Evaluation of Children with Autism Spectrum Disorders', *Pediatrics* 120 (5): 1183–215.
- KLEINMAN, J., ROBINS, D., VENTOLA, P., PANDEY, J., BOORSTEIN, H., ESSER, E., WILSON, L., ROSENTHAL, M., SUTERA, S., VERBALIS, A., HODGSON, S., BARTON, M., GREEN, J., DUMONT-MATHIEU, T., VOLKMAR, F., KLIN, A., CHAWARSKA, K. & FEIN, D. (2007) 'The Modified Checklist for Autism in Toddlers: A Follow-Up Study Investigating the Early Detection of Autism Spectrum Disorders', *Journal of Autism and Developmental Disorders*. Online first, 20 September.
- KLEINMAN, J., VENTOLA, P., PANDEY, J., VERBALIS, A., BARTON, M., HODGSON, S., GREEN, J., DUMONT-MATHIEU, T., ROBINS, D. & FEIN, D. (2008) 'Diagnostic Stability in Very Young Children with Autism Spectrum Disorders', *Journal of Autism and Developmental Disorders* 38 (4): 606–15.
- KLIN, A., LANG, J., CICCETTI, D. & VOLKMAR, F. (2000) 'Brief Report: Interrater Reliability of Clinical Diagnosis and DSM-IV Criteria for Autistic Disorder: Results of the DSM-IV Autism Field Trial', *Journal of Autism and Developmental Disorders* 30 (2): 163–7.
- LANDA, R. & GARRETT-MAYER, E. (2006) 'Development in Infants with Autism Spectrum Disorders: A Prospective Study', *Journal of Child Psychology and Psychiatry* 47 (6): 629–38.
- LANDA, R.J., HOLMAN, K.C. & GARRETT-MAYER, E. (2007) 'Social and Communication Development in Toddlers with Early and Late Diagnosis of Autism Spectrum Disorders', *Archives of General Psychiatry* 64: 853–64.
- LORD, C. (1995) 'Follow-Up of Two-Year-Olds Referred for Possibly Autism', *Journal of Child Psychology and Psychiatry* 36: 1–18.
- LORD, C., RUTTER, M., DILAVORE, P.C. & RISI, S. (1999) *Autism Diagnostic Observation Schedule—WPS Edition*. Los Angeles, CA: Western Psychological Services.
- LORD, C., SHULMAN, C. & DILAVORE, P. (2004) 'Regression and Word Loss in Autistic Spectrum Disorders', *Journal of Child Psychology and Psychiatry* 45 (5): 936–55.
- LYUSTER, R. (2005) 'Early Regression in Social Communication in Autism Spectrum Disorders: A CPEA Study', *Developmental Neuropsychology* 3: 311–36.
- MAWLE, E. & GRIFFITHS, P. (2006) 'Screening for Autism in Pre-School Children in Primary Care: Systematic Review of English Language Tools', *International Journal of Nursing Studies* 43 (5): 623–36.
- 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: 359–72.
- MULLEN, E. (1995) *The Mullen Scales of Early Learning*. Circle Pines, MN: American Guidance.
- OLEJNIK, S. & ALGINA, J. (2003) 'Generalized Eta and Omega Squared Statistics: Measures of Effect Size for Some Common Research Designs', *Psychological Methods* 8: 434–47.
- 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: 247–57.

- 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 (2): 131–44.
- ROGERS, S. (2001) 'Diagnosis of Autism before the Age of 3', *International Review of Research in Mental Retardation* 23: 1–31.
- RUTTER, M., LECOUTEUR, A. & LORD, C. (2005) *Autism Diagnostic Interview—Revised—WPS Edition*. Los Angeles, CA: Western Psychological Services.
- SALLOWS, G. & GRAUPNER, T. (2005) 'Intensive Behavioral Treatment for Children with Autism: Four-Year Outcome and Predictors', *American Journal of Mental Retardation* 110 (6): 417–38.
- SCHOPLER, E., REICHLER, R.J. & RENNER, B.R. (1988) *The Childhood Autism Rating Scale*. Los Angeles, CA: Western Psychological Services.
- SCHREIBMAN, L. (2000) 'Intensive Behavioral/Psychoeducational Treatments for Autism: Research Needs and Future Directions', *Journal of Autism and Developmental Disorders* 30 (5): 373–8.
- SPARROW, S.S., BALLA, D.A. & CICCHETTI, D.V. (1984) *Vineland Adaptive Behavior Scales*. Circle Pines, MN: American Guidance Service.
- SPITZER, R.L. & SIEGEL, B. (1990) 'The DSM-III-R Field Trial of Pervasive Developmental Disorders', *Journal of the American Academy of Child and Adolescent Psychiatry* 29 (6): 855–62.
- STONE, W., HOFFMAN, E., LEWIS, S. & OUSLEY, O. (1994) 'Early Recognition of Autism', *Archives of Pediatric and Adolescent Medicine* 148: 174–9.
- 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 and Allied Disciplines* 40 (2): 219–26.
- STONE, W.L., COONROD, E.E. & OUSLEY, O.Y. (2000) 'Brief Report. Screening Tool for Autism in Two-Year-Olds (STAT): Development and Preliminary Data', *Journal of Autism and Developmental Disorders* 30: 607–12.
- SUTERA, S., PANDEY, J., ESSER, E.L., ROSENTHAL, M.A., WILSON, L.B., BARTON, M., GREEN, J., HODGSON, S., ROBINS, D.L., DUMONT-MATHIEU, T. & FEIN, D. (2007) 'Predictors of Optimal Outcome in Toddlers Diagnosed with Autism Spectrum Disorders', *Journal of Autism and Developmental Disorders* 37 (1): 98–107.
- TUCHMAN, R.F. & RAPIN, I. (1997) 'Regression in Pervasive Developmental Disorders: Seizures and Epileptiform Electroencephalogram Correlates', *Pediatrics* 99 (4): 560–6.
- WERNER, E., DAWSON, G., OSTERLING, J. & DINNO, N. (2000) 'Brief Report: Recognition of Autism Spectrum Disorder before One Year of Age: A Retrospective Study Based on Home Videotapes', *Journal of Autism and Developmental Disorders* 30 (2): 157–162.
- WILLIAMS, J. & BRAYNE, C. (2006) 'Screening for Autism Spectrum Disorders; What Is the Evidence?', *Autism* 10 (1): 11–35.
- WONG, V., HUI, L., LEE, W., LEUNG, L., HO, P., LAU, W., FUNG, C. & CHUNG, B. (2004) 'A Modified Screening Tool for Autism (Checklist for Autism in Toddlers [CHAT–23]) for Chinese Children', *Pediatrics* 114: e166–e176.
- YIRMIYA, N. & OZONOFF, S. (2007) 'The Very Early Autism Phenotype', *Journal of Autism and Developmental Disorders* 37 (1): 1–11.
- ZWAIGENBAUM, Z., BRYSON, S., ROGERS, T., ROBERTS, W., BRIAN, J. & SZATMARI, P. (2005) 'Behavioral Manifestation of Autism in the First Year of Life', *International Journal of Developmental Neuroscience* 23: 143–52.

ZWAIGENBAUM, L., THURM, A., STONE, W., BARANEK, G., BRYSON, S.,
IVERSON, J., KAU, A., KLIN, A., LORD, C., LANDA, R., ROGERS, S. &
SIGMAN, M. (2007) 'Studying the Emergence of Autism Spectrum Disorders in
High-Risk Infants: Methodological and Practical Issues', *Journal of Autism and
Developmental Disorders* 37 (3): 466–80.