A Screening Instrument for Autism at 18 Months of Age: A 6-Year Follow-up Study

GILLIAN BAIRD, F.R.C.PAED., TONY CHARMAN, PH.D., SIMON BARON-COHEN, PH.D., ANTONY COX, F.R.C.PSYCH., JOHN SWETTENHAM, PH.D., SALLY WHEELWRIGHT, B.A., AND AURIOL DREW, M.A.

ABSTRACT

Objectives: A population of 16,235 children aged 18 months was screened using the Checklist for Autism in Toddlers (CHAT) to identify childhood autism (CA). Two further screening procedures were conducted at age 3 and 5 years. The population was followed up at age 7 years in order to establish the sensitivity, specificity, and positive predictive value of the instrument. **Method:** A brief checklist assessing joint attention and pretend play behaviors was administered by primary health care practitioners when the children were 18 months old. Follow-up methods included screening through parents and health practitioners and checking medical and educational records. **Results:** Nineteen cases of CA were successfully identified by the CHAT at 18 months. At follow-up a total of 50 cases of CA were identified via all surveillance methods. Thus, the CHAT has a sensitivity of 38% and a specificity of 98% for identifying CA. The positive predictive value of the instrument was maximized by concentration on the highest-risk group. Repeated screening 1 month later increased the positive predictive value to 75% for identification of CA but reduced the sensitivity to 20%, although the specificity was close to 100%. The screen also identified cases of pervasive developmental disorder as well as children with language and other developmental disorders. **Conclusions:** The CHAT can be used to identify cases of autism and related pervasive developmental disorders at 18 months of age. It is emphasized that the CHAT is not a diagnostic instrument but can identify potential cases of autism spectrum disorders for a full diagnostic assessment. *J. Am. Acad. Child Adolesc. Psychiatry,* 2000, 39(6):694–702. **Key Words:** autism, pervasive developmental disorder, Asperger's syndrome, screening.

It is rare for autism to be diagnosed before 3 years of age, despite the fact that in most cases autism has an onset in infancy (Howlin and Moore, 1997; Kanner, 1943) and is the result of genetic factors affecting brain development very early in life (Szatmari et al., 1998). While a number of rating scales that measure severity of autistic symptoms exist, these are primarily used to assess clinically referred samples (e.g., Autism Behavior Checklist, Krug et al., 1980; Childhood Autism Rating Scale, Schopler et al., 1980; Pervasive Developmental Disorders Screening Test, Siegel, 1999). The present study reports the properties of a general population screen for autism that attempts to identify autism at an earlier age than it is commonly detected via usual surveillance methods. There is some, although not uncontroversial, evidence of the benefits of early intervention programs (McEachin et al., 1993; Rogers, 1998); therefore, the possibility of early identification of autism merits investigation.

The Checklist for Autism in Toddlers (CHAT) was designed to prospectively identify autism at 18 months of age (Table 1). The CHAT assesses pretend play, protodeclarative pointing, and gaze monitoring, by parental report and health practitioner observation through direct testing. In a first study (Baron-Cohen et al.,

Accepted December 29, 1999.

Dr. Baird, Prof. Cox, and Ms. Drew are with the Newcomen Centre, Guy's Hospital, London. Dr. Charman is with the Institute of Child Health, University College London. Dr. Baron-Cohen and Ms. Wheelwright are with the Departments of Experimental Psychology and Psychiatry, University of Cambridge, England. Dr. Swettenham is with the Department of Human Communication and Science, University College London.

This research was supported by 3 MRC project grants to S.B.C., A.C., and G.B. (1992–95, 1995–96, and 1996–99). The authors thank the health practitioners and families who took part in the study. They especially thank the consultant community pediatricians and speech and language therapists in the districts involved; Ly Kemal, Deborah Jones, Alex Cutting, and Emma Hernshaw for assistance with data collection; and Natasha Nightingale and Mary Marden for major administrative and secretarial support.

Reprint requests to Dr. Charman, Behavioural Sciences Unit, Institute of Child Health, 30 Guilford Street, London WC1N 1EH, England; e-mail: t.charman@ich.ucl.ac.uk.

TABLE 1
Checklist for Autism in Toddlers (CHAT)

Section A: Ask parent		
1. Does your child enjoy being swung, bounced on your knee, etc?	YES	NO
2. Does your child take an interest in other children?	YES	NO
3. Does your child like climbing on things, such as up stairs?	YES	NO
4. Does your child enjoy playing peek-a-boo/hide-and-seek?	YES	NO
5. Does your child ever PRETEND, for example, to make a cup of tea using a toy cup and teapot, or pretend other things?[a]	YES	NO
6. Does your child ever use his/her index finger to point, to ASK for something?	YES	NO
7. Does your child ever use his/her index finger to point, to indicate INTEREST in something?	YES	NO
8. Can your child play properly with small toys (eg. cars or bricks) without just mouthing, fiddling or dropping them?	YES	NO
9. Does your child ever bring objects over to you (parent) to SHOW you something?	YES	NO
Section B: GP or HV observation		
i. During the appointment, has the child made eye contact with you?	YES	NO
ii. Get child's attention, then point across the room at an interesting object and say "Oh look! There's a [name of toy]!" Watch child's face. Does the child look across to see what you are pointing at?	YES	NO*
iii. Get the child's attention, then give child a miniature toy cup and teapot and say, "Can you make a cup of tea?" Does the child pretend to pour out tea, drink it, etc.?[b]	YES	NO**
iv. Say to the child "Where's the light?" or "Show me the light." Does the child POINT with his/her index finger at the light?	YES	NO***
v. Can the child build a tower of bricks? (If so, how many?) (Number of bricks: ____)	YES	NO

Note: Reproduced with permission from Baron-Cohen et al., (1992), *British Journal of Psychiatry* Vol 161, p. 842.

[a] An alternative appropriate for North America would be ". . . for example, to feed himself/herself or a doll with milk and cookies from an empty cup and plate?"

[b] An alternative appropriate for North America would be, "Can you feed the doll/give the doll a drink?"

* To record YES on this item, ensure the child has not simply looked at your hand, but has actually looked at the object you are pointing at.

** If you can elicit an example of pretending in some other game, score a YES on this item.

*** Repeat this with "Where's the teddy?" or some other unreachable object, if child does not understand the word "light." To record YES on this item, the child must have looked up at your face around the time of pointing.

1992), the CHAT correctly predicted the 4 undetected cases of autism at 18 months of age from among 41 siblings of children with autism and therefore at higher risk of developing autism. A later study demonstrated that this procedure could be applied to community screening (Baron-Cohen et al., 1996). Ten children with childhood autism were prospectively identified from a population of 16,235 children screened at 18 months of age. A series of follow-up screening and surveillance procedures designed to identify all children from the population with childhood autism were conducted throughout the following 6 years. The intention was to test the CHAT screen's *sensitivity* (the proportion of children with a disorder identified by the screen), *specificity* (the proportion of children without the disorder whom the screen identified as normal), and *positive predictive value* (PPV) (the proportion of children with a positive screen result who have the disorder) (Aylward, 1997; Hall, 1984).

METHOD

Subjects

The eligible population was a 12-month birth cohort of children from 10 health districts in the South East Thames Health Region, England. Of the total population of 40,818 children aged 18 months who were eligible for screening, 16,235 (39.8%) were screened using the CHAT (age mean = 18.7 months, SD = 1.1 months). This proportion reflected the fact that only children aged 18 ± 2 months were included in the study. In most cases (n = 13,694; 84.3%) the CHAT screen was administered by primary health care practitioners during a routine 18-month developmental check. However, in a proportion of cases (n = 2,541; 15.7%), section A and items Bi, Bii, and Bv of the CHAT were mailed directly to parents. Forms were returned to the research team. Some children with profound developmental delay were excluded. Exclusions were not systematic or part of the design of the study but reflected the health practitioner's decision not to impose additional assessment on parents of children with preidentified developmental problems. A 2-stage screening procedure was adopted. Children who were initially screen-positive received a second administration of the screen 1 month later, conducted by the research team and not the primary health care practitioner. The intention was to minimize false positives.

On the basis of the earlier high-risk study, the risk thresholds for autism were set at 2 levels after the second CHAT administration. The

high risk for autism group included those children who failed all 5 key items on the CHAT: protodeclarative pointing (A7 + Biv), gaze monitoring (Bii), and pretend play (A5 + Biii) both initially and on repeated administration. The *medium risk for autism* group included those children who failed the 2 key items relating to protodeclarative pointing (A7 + Biv) but who did not meet criteria for group 1 (i.e., they passed at least one of the other items: A5, Bii, Biii) both initially and on repeated administration. Note that for parent-completed forms the high-risk threshold was failing all 3 key items administered (A5, A7, Bii), and children who met this profile were rescreened. No parent-completed medium-risk cases (which would involve failing A7 only) were rescreened.

Screening and Surveillance Methods

Five methods were used to determine the total number of children in the screened population with an autism spectrum disorder at age 7 years.

1. *Checklist for Autism in Toddlers.* The CHAT screen was administered at 18 months and repeated 1 month later on a selected group, as described above.

2. *Checklist for Referral.* All 16,235 children were rescreened with the Checklist for Referral (CR) at 3½ years (mean = 41.8 months, SD = 6.0 months); 12,770 checklists were returned (78.7% of the original sample). The remaining cases were untraceable or did not return the postal questionnaire. Children were screened by a health practitioner (8,508; 66.6%) or by a postal version of the questionnaire sent to the parents (4,262; 33.4%). The CR asks whether the child has been referred for assessment of autism and pervasive developmental disorder (PDD). According to CR returns, there were 22 (0.17%) referrals for autism spectrum disorders (excluding already identified children). Notes of all of these were examined. Where it was considered that the child might have an autism spectrum disorder, the child was assessed by the research team.

3. *Pervasive Developmental Disorders Questionnaire.* At 5½ years the 16,235 children were rescreened with the Pervasive Developmental Disorders Questionnaire (PDD-Q); 7,766 questionnaires were returned (47.8% of the original sample). The PDD-Q was mailed directly to parents. The PDD-Q asks parents about aspects of their child's social and pragmatic competence, as well as the presence of atypical behaviors characteristic of autism spectrum disorders. The key items relevant to social, pragmatic, and repetitive impairments were spontaneous chat, greeting, offering comfort, sharing interests, preoccupying interests, successful play with others, upset at changes in routine, insistence on sameness, and idiosyncratic interests. In a pilot study 37 of 40 children in whom Asperger's syndrome was already diagnosed failed more than 5 of the key items, and this threshold was used in the present study (Baron-Cohen et al., unpublished, 1999). Of 7,766 PDD-Q returns, 63 children (0.81%) scored above the cutoff. Eight children had been identified earlier in the project as having an autism spectrum disorder. In addition, 14 had received a non-PDD diagnosis locally (attention-deficit/hyperactivity disorder, language disorder, learning problems). Of the remaining 41 children, 9 were seen for a full clinical assessment by the research team: 2 had childhood autism; 1 had PDD; 1 had deficits in attention, motor control, and perception (DAMP); 3 had language disorder; 1 had general developmental delay; and only 1 was clinically normal. Details of the remaining 34 children were obtained from the survey of local services at 7 years, described below.

4. *Outside Referral to Clinical Center.* During the course of the project a number of children from the screened population were referred for diagnostic assessment at the regional center where the research team was based or at another regional center.

5. *Medical, Educational, Social Service, or Other Records.* At age 7 to 8 years, all children not already known to the research team who had a diagnosis of an autism spectrum disorder from local professionals were discussed with local teams. Special educational needs registers were perused and staff of special schools consulted, to ensure that all possible children were identified. IQ and language assessments were obtained from records. Information from the notes on each child's presentation at the age of 4 to 5 years was recorded and cases were classified by clinical consensus by G.B. and A.C. according to *ICD-10* criteria. Children identified by this check on case notes were not directly assessed by the research team.

Assessment Procedures for Children Seen by the Research Team

Children identified by methods 1, 2, 3, and 4 were directly assessed by the research team. Children were assigned *ICD-10* diagnoses, achieved as a result of the compilation of all available clinical, historical, and psychometric evidence, including administration of the Autism Diagnostic Interview-Revised (ADI-R) (Lord et al., 1994); assessment of IQ using the Leiter International Performance Scale (Leiter, 1952) or the Kaufman Assessment Battery for Children (Kaufman and Kaufman, 1983); language assessment using the Reynell Developmental Language Scales (Reynell, 1985), the Clinical Evaluation of Language Fundamentals-Preschool (Wiig et al., 1992), or the Clinical Evaluation of Language Fundamentals-Revised (Semel et al., 1987); and a structured interactive assessment measuring verbal and nonverbal social communicative competence. A consensus clinical judgment was reached by 3 clinicians (G.B., A.C., and A.D.), all highly experienced in the diagnosis of autism and related PPDs (Cox et al., 1999).

ICD-10 (World Health Organization, 1993) criteria for childhood autism (F84.0) were used. In addition, cases who met *ICD-10* criteria for Asperger's syndrome (F84.5), in that they had IQs greater than 70, had no significant delay in language milestones or autistic-type abnormality in early language development, and had current speech that was fluent but pragmatically impaired, were noted. However, there is disagreement about the interpretation of the hierarchical classification rules that both the *DSM* (American Psychiatric Association, 1994) and *ICD* systems follow with regard to Asperger's syndrome (Miller and Ozonoff, 1997). Here we note cases that clearly fit the criteria for Asperger's syndrome even though the severity of their autistic presentation is sufficient for them to meet full criteria for autism. Children who met criteria for atypical autism (F84.11) (in all cases this was due to *atypicality* in symptoms, not onset) were not distinguished from those who met criteria for the category *other pervasive developmental disorder* (F84.8), reserved for cases in which the *degree or severity* of abnormality was less than that required to meet criteria for childhood autism in one or more symptom areas. Hereafter we label these cases PDD.

RESULTS

Using all screening and surveillance methods, we identified 50 cases (47 boys, 3 girls) of childhood autism from the screened population, 5 of whom (all boys) also met criteria for Asperger's syndrome (26 cases diagnosed by the research team, 24 cases diagnosed by clinicians in the region). This gives a prevalence rate of 30.8 cases per 10,000 of childhood autism (95% confidence interval = 22.9 to 40.6). Forty-four cases (36 boys, 8 girls) of PDD were identified (17 diagnosed by the research team, 27 diagnosed by clinicians in the region), a prevalence rate of 27.1 cases per 10,000 (95% confidence interval = 19.7 to 36.4).

TABLE 2
Sensitivity, Specificity, and Positive Predictive Values for the High- and Medium-Risk Threshold at CHAT-1

	Sensitivity		Specificity		PPV	
	No.	(%)	No.	(%)	No.	(%)
CA						
High-risk group	10/50	(20.0)	16,157/16,185	(99.8)	10/38	(26.3)
Medium- + high-risk group	19/50	(38.0)	15,797/16,185	(97.6)	19/407	(4.7)
PDD						
High-risk group	1/44	(2.3)	16,154/16,191	(99.8)	1/38	(2.6)
Medium- + high-risk group	14/44	(31.8)	15,797/16,191	(97.6)	14/407	(3.4)
All PDDs						
High-risk group	11/94	(11.7)	16,114/16,141	(99.8)	11/38	(28.9)
Medium- + high-risk group	33/94	(35.1)	15,767/16,141	(97.7)	33/407	(8.1)

Note: CHAT = Checklist for Autism in Toddlers; PPV = positive predictive value; CA = childhood autism (including 5 cases of Asperger's syndrome); PDD = pervasive developmental disorder.

One-Stage Administration

The sensitivity, specificity, and PPVs for childhood autism and PDD from the 1-stage administration of the CHAT (CHAT-1) are shown in Table 2. Thirty-eight children (0.23%) met the threshold for high risk for autism and 369 met the threshold for medium risk for autism (2.27%), giving a cumulative number of 407 (2.51%) identified as being at risk for autism. Ten of the 38 children in the high-risk category had a diagnosis of childhood autism, a PPV of 26.3%. One additional child in the high-risk group had a diagnosis of PDD, giving an overall PPV for all PDDs of 28.9%. Seventeen of the 407 children who met the threshold for medium risk (*including* those in the high-risk group) had a diagnosis of autism. An additional 2 children in this group met criteria for Asperger's syndrome/autism (PPV 4.7%), and a further 14 met criteria for PDD (PPV 3.4%). The overall PPV for all PDDs using the medium-risk threshold was 8.1%. Thus, from the single administration and using the medium- and high-risk thresholds, the sensitivity of the CHAT screen was 38.0% (19/50 cases) for detection of childhood autism (including cases meeting criteria for Asperger's syndrome) and 31.8% (14/44 cases) for detection of PDD. Taking an autism spectrum disorder approach, the overall sensitivity of the screen was 35.1% (33/94 cases: 17 autism, 2 Asperger's syndrome/autism, 14 PDD).

Of the 27 children from the high risk for autism group who did not receive autism spectrum diagnoses, 4 children were identified at follow-up as having language disorders and 3 as having other developmental disabilities at follow-up, giving a PPV for all developmental disorders of 47.4%. From the medium risk for autism group, of the 347 children who did not receive autism spectrum diagnoses, 25 children were identified as having language disorders, 5 as having moderate or severe developmental delay, and 6 as having other developmental disabilities, giving a PPV for all developmental disorders of 16.7%. The results of the 1-stage administration are summarized in Figure 1.

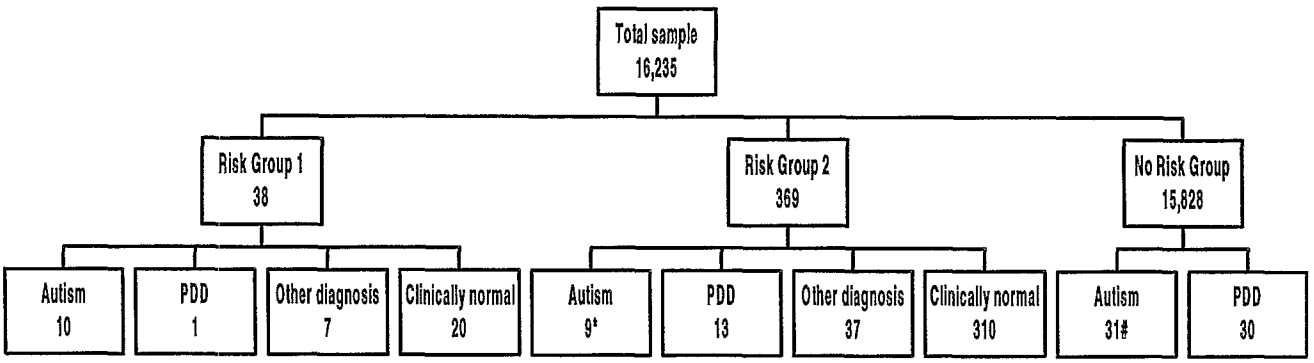

Fig. 1 Identification of childhood autism and pervasive developmental disorder (PDD) from the 1-stage administration of the Checklist for Autism in Toddlers. *Includes 2 cases meeting criteria for Asperger's syndrome. #Includes 3 cases meeting criteria for Asperger's syndrome.

Two-Stage Administration

All 38 children who met high risk for autism threshold on the first administration of the CHAT had a second CHAT administered (CHAT-2) by the research team 1 month later. Owing to resource constraints, it was only possible to directly assess a limited number of cases. Approximately half of the 369 cases with a CHAT-1 medium-risk profile were rescreened to identify the first 22 cases who continued to reach this threshold on the CHAT-2. The remainder of cases did not receive a second administration of the CHAT. Specificity and PPV will be calculated only on the basis of these cases.

After this second screen, only 12 children (0.07%) met the threshold for high risk for autism, and 22 children continued to meet the threshold for medium risk for autism (0.14%), giving a cumulative number of 34 children (0.21%) identified as at risk for autism (high- and medium-risk thresholds combined). Nine of the 12 children in the high-risk category had a diagnosis of childhood autism, a PPV of 75.0% (Table 3). One additional child in the high-risk group had a diagnosis of PDD, giving an overall PPV for all PDDs of 83.3%. Ten of the 34 children who met the threshold for medium risk (*including* those in the high-risk group) had a diagnosis of autism (PPV 29.4%). An additional 10 children in this group met criteria for PDD. The overall PPV for all PDDs using the medium-risk threshold was 58.8% (20 of 34 children received an autism spectrum diagnosis). From the 2-stage use of the CHAT and using the combined medium- and high-risk threshold, the sensitivity of the CHAT screen was 20.0% (10/50 cases) for detection of childhood autism children (including Asperger's syndrome) and 22.7% (10/44 cases) for detection of PDD. Taking an autism spectrum disorder approach, the overall sensitivity of the screen was 21.3% (20 of 94 cases: 10 autism, 10 PDD).

Of the 2 children from the high risk for autism group who did not receive an autism spectrum diagnosis, 1 child had a severe expressive and receptive language disorder and only 1 child was developmentally normal, giving a PPV for all developmental disorders of 83.3%. From the 12 children in the medium risk for autism group at CHAT-2 who did not receive an autism spectrum diagnosis, 7 had language disorders, 2 had general developmental delay, 1 had cerebral palsy, and at follow up only 2 had no clinical diagnosis, giving a PPV for all developmental disorders of 29.4%. Thus, almost all children who met criteria for high or medium risk for autism at CHAT-2 had significant developmental disorders that warranted assessment. The results of the 2-stage administration are summarized in Figure 2.

Comparing Characteristics of Screen True Positives and False Negatives

Characteristics of the 24 CHAT-1 screen true positives (13 childhood autism [11 boys, 2 girls], 11 PDD [10 boys, 1 girl]) were compared with those of the 19 screen false negatives (13 childhood autism [12 boys, 1 girl], 6 PDD [5 boys, 1 girl]) who were directly assessed by the research team. The screen true positives were assessed at a younger age than the screen false negatives (Table 4) (analysis of variance; autism: $F_{1,24} = 6.94$, $p < .05$; PDD: $F_{1,17} = 13.3$, $p < .01$). Because of noncompliance, IQ data were available for only 9 of the true-positive children and 9 of the false-negative children with autism, but for all children

TABLE 3
Sensitivity, Specificity, and Positive Predictive Values for the High- and Medium-Risk Threshold at CHAT-2

	Sensitivity		Specificity		PPV	
	No.	(%)	No.	(%)	No.	(%)
CA						
High-risk group	9/50	(18.0)	16,182/16,185	(100)	9/12	(75.0)
Medium- + high-risk group	10/50	(20.0)	16,161/16,185	(99.9)	10/34	(29.4)
PDD						
High-risk group	1/44	(2.3)	16,180/16,191	(99.9)	1/12	(8.3)
Medium- + high-risk group	10/44	(22.7)	16,167/16,191	(99.9)	10/34	(29.4)
All PDDs						
High-risk group	10/94	(10.6)	16,139/16,141	(100)	10/12	(83.3)
Medium- + high-risk group	20/94	(21.3)	16,127/16,141	(99.9)	20/34	(58.8)

Note: CHAT = Checklist for Autism in Toddlers; PPV = positive predictive value; CA = childhood autism (including 5 cases of Asperger's syndrome); PDD = pervasive developmental disorder.

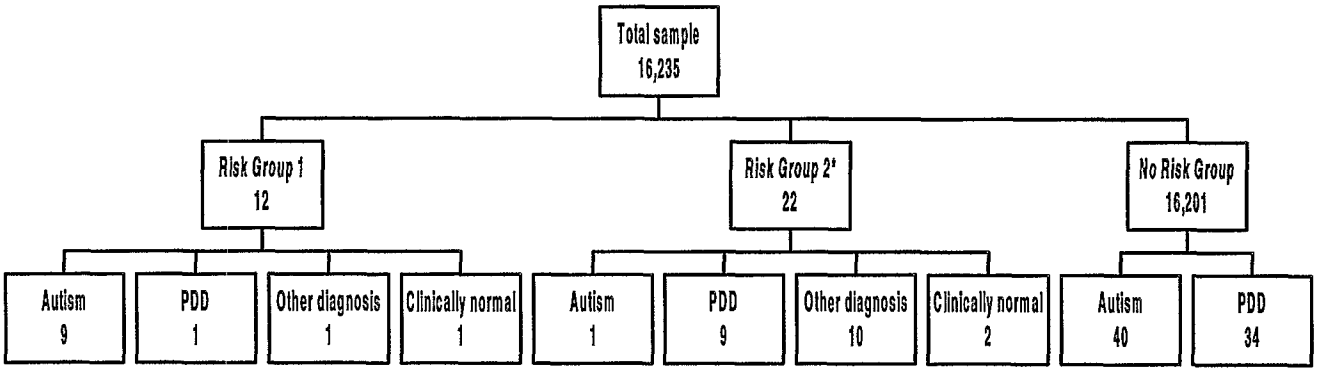

Fig. 2 Identification of childhood autism and pervasive developmental disorder (PDD) from the 2-stage administration of the Checklist for Autism in Toddlers (CHAT). *Only half of the 369 CHAT-1 medium-risk group were administered the CHAT-2.

with PDDs. Language data were available for 12 and 5 of the true-positive and false-negative children with autism, and for 8 and 6 of the true-positive and false-negative children with PDDs, respectively. There were no differences between either the childhood autism or PDD true-positive and false-negative children in terms of language ability (all $p > .10$). The false-negative children with autism had a higher IQ than the true-positive children with autism ($F_{1,16} = 6.20, p < .05$). Conversely, the true-positive children with PDDs had significantly higher IQs than the false-negative children with PDDs ($F_{1,15} = 5.47, p < .05$).

Using one-way analysis of variance with age and IQ entered as covariates, we found no differences between the ADI-R domain scores of true positives and false negatives, either for cases of autism or PDD. However, there were several significant covariate effects, with older children scoring higher than younger children. The ADI-R has an established algorithm cutoff for an autism diagnosis (Lord et al., 1994). Eleven of the 13 true-positive and 12 of the 13 false-negative children with autism reached criterion, as did 1 of 11 and 3 of 6 true-positive and false-negative children with PDD diagnoses, respectively. There were no differences between the true positives and false negatives on ADI-R items that record parental report of early concern: age of first concern, time of onset with hindsight, and loss of skills. Next, we compared the ADI-R dimension scores of the subjects with childhood autism and PDD who were screen true positive and false negative. For the true positives, the subjects with childhood autism scored significantly higher than those with PDD on all 3 dimensions of the ADI-R (analysis of covariance, all $p < .05$). For the false negatives, there were no differences between the subjects with childhood autism and those with PDD (analysis of covariance, all $p > .10$).

TABLE 4
Age, IQ, Language, and Autism Diagnostic Interview-Revised (ADI-R) Subdomain Scores for the True Positives Versus False Negatives

	No.	Age (Months)	IQ	Lang. Receptive *z* Score	Lang. Expressive *z* Score	RSI	VC[a]	NVC	RSB
Autism									
TP	13	47.5 (14.3)[b]	65.7 (21.8)[b]	−2.3 (0.7)	−2.4 (0.9)	16.8 (6.4)	17.7 (6.2)	11.4 (2.1)	4.2 (2.6)
FN	13	61.5 (12.2)	88.7 (16.9)	−2.8 (0.5)	−3.3 (0.1)	19.1 (6.6)	15.3 (3.2)	11.7 (3.2)	4.9 (2.0)
PDD									
TP	11	46.4 (7.5)[c]	99.5 (15.0)[d]	−1.1 (1.2)	−1.4 (1.2)	7.8 (3.7)	8.1 (3.1)	12.0 (—)	1.8 (1.6)
FN	6	65.0 (13.8)	83.2 (10.9)	−2.0 (1.0)	−2.0 (1.0)	13.3 (7.2)	14.7 (7.1)	— (—)	3.5 (2.3)

Note: Values represent mean scores (SD). TP = true positive; FN = false negative; PDD = pervasive developmental disorder. ADI-R domains: RSI = Reciprocal Social Interaction; VC = Verbal Communication; NVC = Nonverbal Communication; RSB = Repetitive and Stereotyped Behavior.

[a] Seven true positives and 3 false negatives with autism, and 1 true positive with PDD, had insufficient language to score verbal items.

[b] TP < FN, $p < .05$.

[c] TP < FN, $p < .01$.

[d] TP > FN, $p < .05$.

DISCUSSION

With a 1-stage administration, the sensitivity of the CHAT in identifying autism using the high-risk threshold was low but was increased by the use of the medium-risk threshold. The screen also identified cases of PDD. At both thresholds the specificity was high (>95%) for childhood autism and PDD. However, with the 1-stage administration the PPV was low, particularly at the medium-risk threshold. When a 2-stage screening procedure was used, the PPV in detecting autism was increased to 75% for the high-risk group and 83% when PDD was included. However, the second administration also led to a proportion of children with autism and PDD falling out of the risk groups, decreasing the screen's sensitivity. The specificity for childhood autism and PDD using the 2-stage administration was close to 100%.

Although the high-risk threshold was set on the basis of a previous study (Baron-Cohen et al., 1992), the lower risk threshold had not been previously tested. The choice of the lower risk threshold was guided by the fact that previous studies have indicated that impairments in joint attention are the most discriminating feature of young children with autism (Charman, 1997). In fact, only 1 child with autism and 3 with PDD failed *only* the 2 joint attention items measuring protodeclarative pointing, and 3 children with autism and 1 child with Asperger's syndrome failed *only* the 2 pretend play items, from the 5 key items on the CHAT. Thus, it appears that the *combination* of failing joint attention *and* pretend play items at 18 months of age may indicate risk for autism and related PDDs.

Possible Explanations of Screen False Negatives

Screen false negatives could have arisen for several reasons. First, parents were asked to report whether their child had "ever" pointed or produced examples of simple pretend play. If the screen had asked if children only "rarely" produced such behaviors, its sensitivity might have been higher, although at a cost to its PPV and specificity.

Alternatively, false negatives might have arisen because the discriminations that are required to identify the type of impairments in pretend play and joint attention behaviors that characterize young children with autism are not elicited by a simple parental questionnaire or single observation. It is noteworthy that on interview when their child was aged 3 to 7 years, many parents of screen false negatives reported that their child had never pointed for interest, although they had reported the presence of such behavior at the CHAT administration at 18 months. Subsequently, many parents indicated that their children's pointing at 18 months occurred in the context of requesting but not sharing interests. Furthermore, almost all cases of disagreements between parental report and practitioner observation consisted of parents reporting that their child pointed for interest, which was not verified by practitioner observation, suggesting that discrimination of protodeclarative acts may be particularly difficult for parents.

Another possible explanation is that a subgroup of children with autism may exist in whom pretend play and joint attention are intact at 18 months. This may be "late-onset" autism, whereby after a period of normal development a child loses skills or his/her development plateaus (Volkmar and Cohen, 1989). However, this phenomenon would account for a minority of cases only, since by contemporaneous clinic records only 9 of the 31 false negatives with childhood autism were reported to have shown a period of regression after the 18-month screen.

Differences Between the Screen True Positives and False Negatives

There were few systematic differences between the screen true positives and false negatives, and those that were found are not easily interpreted. True-positive children with autism had a lower IQ than the false-negative children with autism, but conversely the true-positive children with PDD had a higher IQ than the false-positive children with PDD. In terms of severity of autistic symptoms as measured by the ADI-R, there were no differences between either the autism or PDD true positives and false negatives. However, the true-positive cases with PDD had the lowest ADI-R scores, perhaps reflecting differences in the diagnosis of PDD between the research team and the clinicians in the region.

Prevalence

We consider the current prevalence rate of autism of 30.8 per 10,000 cases as a *provisional* figure only. It is higher than the established figure of 4 to 5 per 10,000 (Fombonne, 1999), although some other studies have found rates of 10 cases per 10,000 (Arvidsson et al., 1997) and 20 per 10,000 (Honda et al., 1996). In comparison with other epidemiological studies, this group has a higher male-female ratio (15.7:1) and a higher IQ (60% estimated to have IQs greater than 70), perhaps suggesting overlap with Asperger's syndrome. Although the research team has clinically assessed 26 of the 50 children, and all

but 3 children with autism met the well-established ADI-R cutoff, the remaining children have not been directly assessed by the research team. Their diagnosis reflects current practice in the districts in the study. Confirmation of the diagnosis of all cases identified at age 9 years in the screened *and* nonscreened children from the original birth cohort will form the basis of a future study.

The total prevalence rate of autism spectrum disorders (childhood autism plus PDD) identified in the cohort is 57.9 per 10,000 (95% confidence interval = 46.8–70.9). This prevalence figure is higher than in many previous studies. For the broader autism spectrum, prevalence rates have varied from 16 to 21 per 10,000 (Fombonne, 1999; Wing and Gould, 1979). Comparison of prevalence rates across different studies is impeded by the different base populations studied, differences in recruitment and identification methods, differences in diagnostic interpretation, and changes in diagnostic criteria over time. The current screened sample was drawn from a larger total eligible population, and it may be that a bias operated in which children were presented for the screening and which children health practitioners chose to screen. However, it is not expected that either factor will boost the proportion of children with autism in the sample screened because it is known that at least some children with profound developmental delay who are at high risk for autism were excluded.

One difference between the current study and previous studies is that subjects were identified via 5 different methods over the course of 6 years, and this may be a factor in the higher ascertainment. Another factor may be a relatively small population base that allows close contact between a network of secondary and tertiary level health services, resulting in more accurate ascertainment of cases than is possible in a large population. A recent prevalence study which relied on a similar network found a prevalence rate at age 5 years of 21.1 per 10,000 for childhood autism (Honda et al., 1996).

Limitations

There are several limitations to the present study. First, only 40% of the eligible total population were screened within the ±2-month time limit around 18 months of age. Furthermore, some children with profound handicap were excluded by the health practitioners. In addition, because of resource constraints, only a proportion of the medium-risk group at CHAT-1 received a second administration. Therefore, we do not know the efficiency of the screen in a total population or outside this time window. Second, because the second administration of the screen was conducted by the research team, we do not know how the screen would operate when conducted on both occasions by community health practitioners. Third, we have no data on the reliability of the instrument. Systematic data on the reliability of parental report and health practitioner observation across repeated administrations of screen, and direct comparison of the utility of parent-only versus parent and practitioner administration, would be valuable. Fourth, to test the efficiency of using different lower risk thresholds, we would ideally have assessed a proportion of children who failed every combination of the key items on the first administration. However, this would have entailed assessing many hundreds of children and again was not possible because of resource constraints. Fifth, the prevalence figures include children diagnosed by clinicians in the region as well as children who were assessed by the research team. Although case notes were inspected and cases were discussed with clinicians, confirmation of the prevalence figures requires systematic assessment of cases across both the screened and nonscreened samples from the total eligible birth cohort.

Clinical Implications

This study demonstrated that autism can be prospectively identified at 18 months of age. Using a single administration the CHAT has a moderate sensitivity and low PPV. This has implications for health service resources. Both the need to maintain existing surveillance techniques and the possibility of onward referral and assessment of clinically normal children need to be taken into account if the CHAT is to be used as a screening instrument. A 2-stage administration improves the PPV to a good level but reduces the sensitivity, with some cases being eliminated by the second screen. However, used as a 2-stage procedure, the CHAT has an extremely low false-positive rate (i.e., the specificity is close to 100%) and the PPV for autism spectrum disorders is greater than 50%. In light of the need for genetic advice so that parents can make informed decisions about increasing their family size (Simonoff, 1998), and with encouraging evidence of the benefits of early intervention programs (Rogers, 1998), early detection of autism may result in considerable benefit to children and their families.

REFERENCES

American Psychiatric Association (1994), *Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV)*. Washington, DC: American Psychiatric Association

Arvidsson T, Danielsson B, Forsberg P, Gillberg C, Johansson M, Kjellgren G (1997), Autism in 3–6-year-old children in a suburb of Goteborg, Sweden. *Autism Int J Res Pract* 1:163–173

Aylward GP (1997), Conceptual issues in developmental screening and assessment. *J Dev Behav Pediatr* 18:240–249

Baron-Cohen S, Allen J, Gillberg C (1992), Can autism be detected at 18 months? The needle, the haystack, and the CHAT. *Br J Psychiatry* 161:839–843

Baron-Cohen S, Cox A, Baird G et al. (1996), Psychological markers of autism at 18 months of age in a large population. *Br J Psychiatry* 168:158–163

Charman T (1997), The relationship between joint attention and pretend play in autism. *Dev Psychopathol* 9:1–16

Cox A, Klein K, Charman T et al. (1999), Autism spectrum disorders at 20 and 42 months of age: stability of clinical and ADI-R diagnosis. *J Child Psychol Psychiatry* 40:719–732

Fombonne E (1999), The epidemiology of autism. *Psychol Med* 29:769–786

Hall DMB (1984), *The Child With a Handicap*. Oxford, England: Blackwell

Honda H, Shimizu Y, Misumi K, Niimi M, Ohashi Y (1996), Cumulative incidence and prevalence of childhood autism in children in Japan. *Br J Psychiatry* 169:228–235

Howlin P, Moore A (1997), Diagnosis in autism: a survey of over 1,200 patients in the UK. *Autism Int J Res Pract* 1:135–62

Kanner L (1943), Autistic disturbance of affective contact. *Nerv Child* 2:217–250

Kaufman AS, Kaufman NL (1983), *Kaufman Assessment Battery for Children*. Circle Pines, MN: American Guidance Service

Krug DA, Arick J, Almond P (1980), Behaviour checklist for identifying severely handicapped individuals with high levels of autistic behavior. *J Child Psychol Psychiatry* 21:221–229

Leiter RG (1952), *Leiter International Performance Scale*. Wood Dale, IL: Stoetling

Lord C, Rutter M, Le Couteur A (1994), Autism Diagnostic Interview-Revised. *J Autism Dev Disord* 24:659–686

McEachin JJ, Smith T, Lovaas OI (1993), Long-term outcome for children with autism who received early intensive behavioral treatment. *Am J Ment Retard* 97:359–372

Miller JN, Ozonoff S (1997), Did Asperger cases have Asperger disorder? *J Child Psychol Psychiatry* 38:247–251

Reynell JK (1985), *Reynell Developmental Language Scales,* 2nd ed. Windsor Berks, England: NFER Nelson

Rogers SJ (1998), Empirically supported comprehensive treatments for young children with autism. *J Clin Child Psychol* 27:167–178

Schopler E, Reichler RJ, DeVellis R, Daly K (1980), Towards objective classification of childhood autism: Childhood Autism Rating Scale (CARS). *J Autism Dev Disord* 10:91–103

Semel E, Wiig EH, Secord W (1987), *Clinical Evaluation of Language Fundamentals-Revised.* San Antonio, TX: Psychological Corporation

Siegel B (1999), Detection of autism in the 2nd and 3rd years: the Pervasive Developmental Disorders Screening Test (PDDST). Presentation at the biennial meeting of the Society for Research in Child Development, Albuquerque, April

Simonoff E (1998), Genetic counseling in autism and pervasive developmental disorders. *J Autism Dev Disord* 28:447–456

Szatmari P, Jones MB, Zwaigenbaum L, MacLean JE (1998), Genetics of autism: overview and new directions. *J Autism Dev Disord* 28:351–368

Volkmar FR, Cohen DJ (1989), Disintegrative disorder or "late onset" autism. *J Child Psychol Psychiatry* 30:717–724

Wiig EH, Secord W, Semel E (1992), *Clinical Evaluation of Language Fundamentals-Preschool.* San Antonio, TX: Psychological Corporation

Wing L, Gould J (1979), Severe impairments of social interaction and associated abnormalities in children: epidemiology and classification. *J Autism Dev Disord* 9:11–29

World Health Organization (1993), *Mental Disorders: A Glossary and Guide to Their Classification in Accordance With the 10th Revision of the International Classification of Diseases: Research Diagnostic Criteria (ICD-10).* Geneva: WHO

Gastrointestinal Abnormalities in Children With Autistic Disorder. Karoly Horvath, MD, PhD, John C. Papadimitriou, MD, PhD, Anna Rabsztyn, Cinthia Drachenberg, MD, J. Tyson Tildon, PhD

Objectives: Our aim was to evaluate the structure and function of the upper gastrointestinal tract in a group of patients with autism who had gastrointestinal symptoms. *Study Design:* Thirty-six children (age: 5.7 ± 2 years, mean ± SD) with autistic disorder underwent upper gastrointestinal endoscopy with biopsies, intestinal and pancreatic enzyme analyses, and bacterial and fungal cultures. The most frequent gastrointestinal complaints were chronic diarrhea, gaseousness, and abdominal discomfort and distension. *Results:* Histologic examination in these 36 children revealed grade I or II reflux esophagitis in 25 (69.4%), chronic gastritis in 15, and chronic duodenitis in 24. The number of Paneth's cells in the duodenal crypts was significantly elevated in autistic children compared with non-autistic control subjects. Low intestinal carbohydrate digestive enzyme activity was reported in 21 children (58.3%), although there was no abnormality found in pancreatic function. Seventy-five percent of the autistic children (27/36) had an increased pancreatico-billary fluid output after intravenous secretin administration. Nineteen of the 21 patients with diarrhea had significantly higher fluid output than those without diarrhea. *Conclusions:* Unrecognized gastrointestinal disorders, especially reflux esophagitis and disaccharide malabsorption, may contribute to the behavioral problems of the non-verbal autistic patients. The observed increase in pancreatico-billary secretion after secretin infusion suggests an upregulation of secretin receptors in the pancreas and liver. Further studies are required to determine the possible association between the brain and gastrointestinal dysfunctions in children with autistic disorder. **J Pediatr** 1999;135:559–563. Reproduced with permission from Mosby-Year Book, Inc.