



HHS Public Access

Author manuscript

Autism. Author manuscript; available in PMC 2016 January 29.

Published in final edited form as:

Autism. 2008 September ; 12(5): 537–556. doi:10.1177/1362361308094502.

Screening for autism spectrum disorders in primary care settings

Diana L. Robins, Ph.D.

Georgia State University

Abstract

The need for autism-specific screening during pediatric well-child visits has been established. However, additional support for specific screening instruments is needed. The current study used the Modified Checklist for Autism in Toddlers (M-CHAT) and the M-CHAT Follow-Up Interview to screen 4797 children during toddler checkups. Of the 4797 cases, 466 screened positive on the M-CHAT; of the 362 who completed the follow-up interview, 61 continued to show risk for autism spectrum disorders (ASDs). A total of 41 children have been evaluated; 21 children have been diagnosed with ASD, 17 were classified with non-ASD delays, and three were typically developing. The PPV of M-CHAT plus interview was .57. It is notable that only four of the 21 cases of ASD were flagged by their pediatrician. These findings suggest that the M-CHAT is effective in identifying ASD in primary care settings. Future research will follow this sample longitudinally.

Keywords

Autism; screening; toddlers; M-CHAT

Screening offers the unique opportunity to alert primary care physicians and other healthcare providers to cases in the population that require further clinical attention. Effective screening is low cost in terms of time, money, and healthcare resources, and efficient in terms of maximizing sensitivity (the ability to detect the disorder in the sample) and specificity (the ability to detect wellness, or lack of the disorder). It is not possible to screen for all disorders that may affect young children; however, priority is generally given to those disabilities that have one or more of the following traits: high frequency of occurrence (e.g. hearing impairment), improved outcome if detected early (e.g. phenylketonuria), and efficient, low-cost screening methods available (e.g. heel stick to test newborns for multiple metabolic and genetic disorders). The American Speech–Language–Hearing Association has identified the principles of screening clearly and succinctly in the context of screening for hearing impairment (Gravel et al., 1993); the same principles can be applied to other disorders. Screening is warranted when (1) the cost of not detecting the disease is high, for example in terms of prevalence, severity of disease, cost of treatment, (2) diagnostic criteria are identified, (3) treatment is available, (4) early treatment is more effective than later treatment, and (5) an appropriate screening instrument is available. The new guidelines on developmental screening, issued by the American Academy of Pediatrics (AAP) in July 2006 (Duby et al., 2006) and emphasized again in November 2007 (Johnson and Myers,

2007) highlight the need for screening using standardized instruments to be incorporated into well-child visits for infants and toddlers.

Given the above criteria, it is imperative that screening for autism spectrum disorders (ASD) is conducted at the population level, also referred to as Level 1, low-risk, or first-stage screening. Evidence suggests that ASD is no longer rare; recent prevalence rates of disorders on the autism spectrum have been reported to be as high as 60–116 per 10,000 (Baird et al., 2006; Fombonne, 2006; Harrison et al., 2006; Moldin and Rubenstein, 2006; Williams et al., 2006). Diagnostic criteria are identified for autistic disorder and pervasive developmental disorder not otherwise specified (PDD-NOS) (American Psychiatric Association, 2000). Several models of intervention have been shown to improve communication, increase social relatedness, or reduce autistic symptomatology (Howlin, 2005; Jensen and Sinclair, 2002; Kasari et al., 2006; McConnell, 2002; Yoder and Stone, 2006). Increasing evidence suggests that early intervention results in increased developmental gains in domains such as communication, social interaction, and cognitive ability (Bryson et al., 2003; Dawson and Osterling, 1997; Dawson et al., 2000; Harris and Handleman, 2000; Lord, 1995; 1997; McGee et al., 1999; McGovern and Sigman, 2005; Rogers, 1996; Woods and Wetherby, 2003), greater independence (Gabovitch and Wiseman, 2005), and improved quality of life (Gabovitch and Wiseman, 2005).

Early identification also is critical for reducing the delay in the primary healthcare provider's referral to a specialist who can diagnose ASD (Koegel et al., 2005), which will reduce the burden of ASD on individuals and society at large. Although the validity of early diagnosis was initially questioned, several longitudinal studies have shown that the majority of diagnoses made around the second birthday are stable when children are re-evaluated at age 4 or older (Charman et al., 2005; Cox et al., 1999; Freeman and Cronin, 2002; Lord, 1995; Lord et al., 2006; Moore and Goodson, 2003; Stone et al., 1999), heightening the need for effective screening procedures. The weakest area of screening for ASD at the present time is data supporting the appropriateness of specific screening instruments. However, several promising instruments have been developed in the last decade (Baron-Cohen et al., 1992; 1996; Robins et al., 1999a; 2001; Siegel, 2004). Therefore, although one recent article argued against universal screening for ASD in the United Kingdom (Williams and Brayne, 2006), the consensus in the United States favors screening since the benefits are substantial. Much of the ASD screening literature has looked at two types of screening settings. Level 1 screening measures are appropriate for widespread use among a general pediatric population, whereas Level 2 screening instruments are designed for use in a subsample of the population identified as at risk for the disorder. Level 1 screening instruments must be brief and low cost, since many of the children screened are not at risk. In contrast, Level 2 screening measures can require more time or expertise to administer, since children in the Level 2 sample have a greater likelihood of having the disorder. In order to identify as many children as possible in the population, Level 1 screening through primary care providers is critical, given that until children are identified as being at risk, they are unlikely to see other professionals or specialists. For additional detail regarding the key issues in ASD screening, please refer to recent reviews (Carr and LeBlanc, 2007; Chakrabarti et al., 2005; Coonrod and Stone, 2005; Dumont-Mathieu and Fein, 2005; Mawle and Griffiths, 2006; Nadel and Poss, 2007; Robins and Dumont-Mathieu, 2006).

As a result of overall agreement that screening is necessary, a large inter-disciplinary group outlined practice parameters for improved early detection of ASD (Filipek et al., 1999; 2000). However, one recent survey of pediatricians (Sand et al., 2005) indicated that fewer than 25 percent are regularly incorporating screening with standardized instruments into well-child visits, and a second study found that fewer than 10 percent of pediatricians are using ASD-specific screening instruments (Dosreis and Weiner, 2006).

Furthermore, implementation and maintenance of screening pose significant challenges. One study found that although for the duration of the study, the age of identification dropped, the following year the age of diagnosis rose back to its original level (Holzer et al., 2006). The current AAP guidelines (Duby et al., 2006), released in July 2006, provide an update to these practice parameters, and recommend that primary care physicians incorporate standardized developmental screenings into three well-child visits (9, 18, and 24–30 months), with additional ASD-specific screening at 18 months. A reply to this policy statement, written by the AAP Autism Expert Panel (Gupta et al., 2007), suggested ASD-specific screening twice, at 18- and 24-month checkups, to identify children whose symptoms emerge later than 18 months. A specific instrument is not recommended by the 2006 policy statement, due to variability in standardization, psychometric properties, and the normative samples used in instrument development. However, the statement includes an extensive list of potential measures for healthcare providers to consider, and notes that additional research into the utility of specific screening instruments for Level 1 screening is critical to provide empirical support for screening best practices.

The most recent article from the Council on Children with Disabilities (Johnson and Myers, 2007) builds on the 2006 policy statement. It recommends routine ASD surveillance at every well-child visit. In addition, there are two indications for ASD-specific screening: (1) if the child is attending an 18- or 24-month visit, screening should occur regardless of the surveillance results; and (2) if at any other visit surveillance indicates risk for ASD, screening should occur.

The recent articles (Duby et al., 2006; Johnson and Myers, 2007) include mention of four ASD-specific instruments that have been studied in Level 1 screening samples.

The Checklist for Autism in Toddlers (CHAT) (Baron-Cohen et al., 1996) was the pioneer autism screen; however, it is currently under revision (Allison et al., 2006) to improve sensitivity (see Baird and colleagues, 2000, for detail). Two modifications of the CHAT also are mentioned. First, the Denver modification relaxes the scoring criteria, which will improve sensitivity. However, it has not yet been studied in a Level 1 sample. Second, the CHAT 23 was a combination of the M-CHAT (see below) and the CHAT's observation items, administered in Chinese; this also has not yet been studied in a Level 1 sample.

The Modified Checklist for Autism in Toddlers (M-CHAT) (Robins et al., 2001) was initially studied in a mixed sample of children from primary care settings ($n = 1122$) and early intervention sites ($n = 171$); most of the children diagnosed with ASD were from the latter sample. A more recent article (Kleinman et al., 2000) provides further evidence for the utility of the M-CHAT in both Level 1 and Level 2 samples. A new sample of 3309 cases

from Level 1 and 484 cases from Level 2 screening were screened between 16 and 30 months with the M-CHAT. The positive predictive value (PPV) of the entire sample was .36, identical to the initial M-CHAT study (Robins et al., 2001); when the follow-up interview was included in the calculation of PPV, it rose to .74. The PPV in the Level 1 sample was much lower than in the Level 2 sample (.11 versus .60), although when the follow-up interview was included the PPV rose to .65 and .76, respectively. A subset of cases was rescreened and re-evaluated around age 4; of 1416 children from combined Level 1 and Level 2 samples, including 120 children who received confirmatory clinical evaluations, seven missed cases were identified. Although sensitivity and specificity cannot be determined with certainty (given that only a subset of children have been followed at the second time point), the upper bound of sensitivity is .91.

The Pervasive Developmental Disorders Screening Test-II (PDDST-II) (Siegel, 2004) includes a Level 1 screen, but the validation sample reported in the manual is not from a Level 1 sample, and peer-reviewed studies of the PDDST-II have not yet been published.

Finally, the Social Communication Questionnaire (SCQ) (Rutter et al., 2003a) is designed for children over 4 years old, and is therefore not appropriate for screening at 18 months.

The current study presents a Level 1 sample of children screened for ASD with the M-CHAT during 18- and 24-month well-child checkups with primary care providers. It is the first article to screen an exclusively Level 1 sample using the M-CHAT, and also provides cross-validation for the Level 1 sample reported in Kleinman et al. (2008). Using the procedures and scoring described in Robins and colleagues (2001), children were referred for follow-up if initial screening demonstrated risk for ASD. The goal of the current study is to provide evidence for the utility of Level 1 screening for ASD in the primary care setting, which is critical given the AAP guidelines for universal ASD-specific screening.

Method

Recruitment and Study Procedures

All participants attended a well-child visit at the office of a participating healthcare provider in metro-Atlanta, GA between March 2005 and October 2007. Healthcare providers in the metro-Atlanta area were invited to participate by direct mailing and recruitment advertisements on the Georgia chapter of the American Academy of Pediatrics website and at their conferences. Currently, there are 42 sites participating in data collection: 40 are private practices, with a total of 136 participating physicians and nurse practitioners. The other two sites are public primary care clinics; one is staffed by four nurse practitioners, and the other is staffed by 40 pediatric residents supervised by eight attending pediatricians.

Parents of toddlers at 18- and 24-month well-child visits were invited to participate in the study. In order to account for families attending well-child visits off schedule, children 16–26.9 months were eligible for participation. One primary care clinic does not see children at 18 months. Families attending this site were invited to participate at 15- and 24-month visits (again, allowing for off-schedule visits, the minimum age was 14 months). Parents completed the packet, which included informed consent; demographic information,

including parent's name, child's name, contact information, child's date of birth, and child's sex; and the M-CHAT. Healthcare providers were asked to check a box labeled 'Office Use Only' to indicate concerns about ASD, regardless of the parent's M-CHAT responses. All physicians who flagged M-CHATs received follow-up calls to determine the nature of their concerns. Completed forms were collected by healthcare personnel and mailed to the researchers. Participating healthcare providers were not required to score the M-CHAT; however, they were offered the scoring instructions, and permitted to keep a copy of the child's M-CHAT in the patient file if they wished. Providers were encouraged to make early intervention referrals as usual, regardless of participation in the screening study.

Parents of children whose M-CHAT scores indicated risk for ASD (i.e. at-risk scores on any three M-CHAT items, or two of the six critical items) were called for the structured follow-up interview. The interview was tailored to include only those items for which the child demonstrated risk for ASD. The interview is designed to clarify and elicit specific examples of the child's typical behavior relevant to each M-CHAT item. If the child continued to demonstrate risk following the interview, using the same scoring criteria as the M-CHAT (any three items or two critical items), the family was invited for a complete clinical evaluation. The evaluation was conducted by a team consisting of one licensed psychologist (DLR), one doctoral student clinician, and one undergraduate research assistant to videotape the child. Evaluations took place in a large room in the Georgia State University Psychology Clinic, which contained an area for the child to work at a small table and play on the floor, and an adult seating area at the other end of the room for the parent interviews. The doctoral student administered the clinical measures of cognitive, language, motor, and adaptive functioning, and the licensed psychologist administered the diagnostic instruments. Following the end of testing, the examiners provided brief verbal feedback regarding performance across domains of functioning, including language, motor, non-verbal cognitive, adaptive, social, and play skills.

The following diagnoses were made, based on clinician judgment, using DSM-IV criteria as a guide: autistic disorder and pervasive developmental disorder not otherwise specified (PDD-NOS). Additional classifications which are not included in DSM-IV were made based on clinician judgment: developmental language delay (DLD), which most closely mapped onto DSM-IV diagnoses of expressive language disorder, mixed receptive-expressive language disorder, and communication disorder not otherwise specified; global developmental delay (GDD), which most closely mapped onto mental retardation, although this term was not used with any families, given the young age at evaluation; and broader autism phenotype (BAP), which does not have a corresponding DSM-IV diagnosis. All non-ASD classification required that autism and PDD-NOS were ruled out. In addition, the following guidelines were used: classification with DLD was made when a child scored 2 SD below the mean on Mullen expressive, Mullen receptive, or Vineland communication scores, or 1.5 SD below the mean on two of these language scores; in addition, there were not commensurate delays noted in non-language scores (i.e. Mullen motor and visual reception, Vineland motor). Classification with GDD was made when a child scored 1.5 SD below the mean on at least one language score and at least one non-language score. A child was classified with BAP when the child presented with at least one social deficit found in the DSM-IV autism criteria, and at least one additional symptom from the communication or

behavioral autism criteria, but coupled with other strengths that ruled out a diagnosis of PDD-NOS. Diagnoses were made and explained to parents, and appropriate recommendations were provided. Families received a complete written report approximately 3 weeks following the assessment. This research was approved by the Institutional Review Board at the author's university.

Measures

The Modified Checklist for Autism in Toddlers (M-CHAT; Robins et al., 1999a) is a 23-item parent-report measure. The first nine items are from the Checklist for Autism in Toddlers (Baron-Cohen et al., 1992; 1996), and the remaining items were developed by the authors based on the literature and clinical judgment. The M-CHAT was validated on a sample of 1293 children (Robins et al., 2001) and found to demonstrate adequate internal consistency ($C = .85$). Preliminary psychometric values were reported, with the understanding that true sensitivity and specificity cannot be determined until follow-up is completed and diagnoses are confirmed. Estimates of sensitivity and specificity in this initial M-CHAT sample, based on the accuracy of classification by discriminant function analysis, were reported to be .87 and .99 respectively.

The M-CHAT was scored as reported in Robins and colleagues (2001), with a positive screen indicated by at-risk scores on any three items or two of the six critical items (items 2, 7, 9, 13, 14, 15). Items for which parents circled both 'yes' and 'no' were scored as at risk; similarly, items for which parents wrote in 'sometimes', 'not usually', or 'occasionally' were scored as at risk.

The M-CHAT Follow-Up Interview (Robins et al., 1999b) is a 5–20 minute interview containing specific probes for each M-CHAT item for which the child scored at risk. The interview is designed to elicit details about the child's behavior, including frequency and severity, and extracts specific examples of target behavior (e.g. pretend play). Inter-rater reliability analyses indicate that all items have acceptable kappas (.60–1.0) except item 21 ('Does your child understand what people say?', $K = .43$). In the current study, this interview was conducted on the telephone. The follow-up interview was scored in the same way as the M-CHAT. At-risk responses on any three items or two of the six critical items qualified the child for a diagnostic evaluation. The follow-up interview is designed to be administered by healthcare paraprofessionals with minimal experience with ASD; no formal training is required to use the interview, given that the flowchart format is explicit about decision points, and examples of target behaviors are included in the interview.¹

The clinical evaluation included the Mullen Scales of Early Learning (MSEL; Mullen, 1995), the Vineland Adaptive Behavior Scales–II (Sparrow et al., 2004), and three ASD diagnostic instruments. Clinical judgment was used to determine diagnosis, incorporating results from the Autism Diagnostic Observation Schedule, Module 1 (ADOS; Lord et al., 1999), the Autism Diagnostic Interview–Revised (ADI–R; Rutter et al., 2003b), and the Childhood Autism Rating Scale (CARS; Schopler et al., 1988). Some participants received the ADI–R Toddler version, an experimental version of the ADI–R obtained from the instrument's author. The toddler version eliminated items inappropriate for toddlers and

added items specifically relevant to very young children; however, the algorithm items remain identical to the standard ADI-R.

Statistical Analysis

M-CHAT and all other data were scored and double-entered into a File-Maker Pro database. Any inconsistencies were identified by the double-entry program and corrected. Children who were too young (<14 months) or too old (>26.9 months) were excluded from the study, although families of excluded children with at-risk M-CHAT scores were notified and it was recommended that they contact their physician regarding any developmental concerns. Analyses included descriptives, t-tests, and analysis of variance to investigate the performance of the M-CHAT in identifying children at risk for ASD.

Results

The sample included 4797 children who were screened during 15-, 18-, or 24-month well-child pediatric visits (mean age = 20.92 months, SD = 3.10 months, range = 14.03–26.97 months). The sample consisted of 2384 males (49.7%), 2280 females (47.5%), and 133 children for whom sex was not identified (2.8%). Ethnicity data were ascertained for 1177 participants: 779 were Caucasian (66.2%), 246 were African-American (20.9%), 32 were Hispanic/Latino (2.7%), 24 were Asian (2.1%), two were Hawaiian/Pacific Islander (.1%), and 94 were other ethnicities (8.0%). Refusal rates of families who chose not to participate have been ascertained from a subset of participating physicians (seven sites, 25 physicians).

Refusal rates range from less than 1 percent to 20 percent, and appear to vary based on the office staff's approach to introducing the M-CHAT study; this will be evaluated more systematically in future research.

Of the 4797 children screened, 466 (9.7%) required follow-up based on the scoring outlined in Robins et al. (2001) (see Figure 1). Of these 466 cases, 362 follow-up interviews were completed (77.7%) and 104 families were not reachable or declined to participate further (22.3%). The mean age at follow-up interview was 23.56 months (SD = 4.25, range 14.57–43.27); time between completion of the M-CHAT and the follow-up interview averaged 3.00 months (SD = 2.84, range 0–18.47 months). Mean M-CHAT score was not different for those families who did not complete the structured follow-up interview, relative to those families who completed the interview (4.30 and 4.00, respectively, $t(465) = 1.31$, $p = .19$). Although every attempt was made to complete the follow-up interview within 1 month of the initial M-CHAT screen, participants were invited to continue participating at any time. Overall, 62.5 percent of interviews were conducted within 3 months of the initial M-CHAT screen ($n = 222$), 26.5 percent were 3–6 months after the screen ($n = 94$), 9.3 percent were 6–12 months after ($n = 33$), and 1.7 percent were more than a year after ($n = 6$); seven children were missing a correct date and were not included in these calculations. Change in scores between initial M-CHAT and interview for children whose follow-up interview was completed within 1 month of the M-CHAT screen did not differ from children whose interview was delayed more than 1 month ($t(353) = -.929$, $p = .35$; mean score change for short delay = 2.90, mean score change for long delay = 3.10). It also is of note that the

average delay between M-CHAT and follow-up interview for those families who completed the evaluation was only 1.73 months (SD = 1.15, range 0–4.93 months).

Of the 362 participants who completed the telephone interview, 301 (83.1%) required no additional follow-up. Sixty-one children (16.9%) continued to demonstrate risk for ASD and were invited to participate in the clinical evaluation, and four children passed the M-CHAT or interview but were flagged by the healthcare provider, totaling 65 children invited to participate in the clinical evaluation. Six children were excluded for severe neurological, physical, visual, or hearing deficits that precluded the child's ability to complete the standardized evaluation measures. Five families were not reachable to schedule the evaluation, 13 families declined to attend the evaluation, and 41 agreed to participate in the evaluation.

Mean age at time of evaluation was 24.34 months (SD = 3.92, range = 16.77–33.80). Twenty children were evaluated before the second birthday (11 in the non-ASD group and nine in the ASD group). M-CHAT scores of families who completed the follow-up interview but did not attend the evaluation were not significantly different from those who attended the evaluation ($t(60) = 1.049$, $p = .3$), nor were the follow-up interview total scores different among the groups ($t(55) = 1.360$, $p = .18$). Of the 41 children who have been evaluated, 21 were diagnosed with ASD (51.2% of evaluated cases: five autism, 16 PDD-NOS); of the remaining 20 evaluated cases, four were classified with broader autism phenotype (BAP), meaning that they demonstrated notable features of ASD but not above the clinical threshold; 11 with language delay; one with global cognitive delay; one with other non-ASD developmental delay; and three were found to be typically developing. For all analyses, children with BAP were classified as non-ASD. All children were seen within 10 months of completing the M-CHAT (mean time between M-CHAT and evaluation = 3.30 months, SD = 1.94 months, range 1.20–9.87 months).

Physicians flagged 19 M-CHATs indicating concern about the child's development. Upon follow-up with these physicians, nine reported concerns about ASD specifically, whereas the other nine were concerned with language and speech delays (and one did not respond to requests for more information); only the nine cases for whom the physician had ASD concerns were evaluated. It is of note that of the nine children for whom the physician had ASD concerns, six were identified by the M-CHAT, and five of those cases continued to meet eligibility for the evaluation after the follow-up interview. In only four cases did the physician flag a case that was not also determined to be at risk based on M-CHAT plus interview; however, due to the physician's concern, each of these cases was evaluated. Of these four cases for whom the M-CHAT score was not of concern but the pediatrician suspected ASD, all were classified with non-ASD delays: one child was classified with language delay, one with global developmental delay, and two with BAP. Of the five cases for whom M-CHAT plus interview score would have qualified for the evaluation regardless of the pediatrician's flag, one child received no diagnosis, two were diagnosed with PDD-NOS, and two with autism. It is notable that all four of the pediatrician flagged children diagnosed with ASD were identified by the M-CHAT and interview, although only one of the two classified with BAP was identified based on M-CHAT score, and he did not demonstrate risk on the follow-up interview and would not have been seen for evaluation

had it not been for the physician's concern. Therefore, if one considers the 'hit rate' of physician concerns to be four of the nine cases for whom there were ASD concerns, the PPV for physician concern is .44; however, it is notable that the upper bound of sensitivity is quite low, given that physicians only noted concerns in four of the 21 children diagnosed with ASD (.19).

Children who were evaluated were divided into two groups: children who were not diagnosed with ASD (non-ASD: $n = 20$), and children who were diagnosed with ASD ($n = 21$). No significant difference in age of screening, telephone follow-up, or evaluation was observed ($ps > .3$). Comparison of total and critical scores on the M-CHAT and follow-up interview indicated that children with ASD had higher scores than the non-ASD sample on the M-CHAT total and critical score, and the follow-up interview critical score, and a trend toward higher follow-up interview score (Table 1). Comparison of cognitive, language, motor, and adaptive functioning levels between the two groups indicated that the ASD group was significantly more impaired on all scales on the Mullen Scales of Early Learning, and on the communication domain and the adaptive behavior composite of the Vineland-II, than the non-ASD sample, and the ASD group showed a trend for more severe impairment on the socialization domain of the Vineland-II (Table 2), with small to moderate effect sizes. With the exceptions of Mullen expressive language and Vineland-II adaptive behavior composite, these differences remained significant when the three children who were typically developing were removed from the sample, suggesting that even among the subsample of children with delays, the ASD sample demonstrated greater impairment than children with non-ASD delays.

Positive predictive value (PPV) is the proportion of cases identified in the screening who are diagnosed with ASD. Given that the cases who did not complete participation cannot be classified as ASD or non-ASD, PPV for M-CHAT alone and M-CHAT plus follow-up interview were calculated based on completed cases. Of the 362 cases who initially failed the M-CHAT and completed the follow-up interview, 21 were diagnosed with ASD, indicating PPV of .058. However, most children who failed the M-CHAT passed the follow-up interview. Given that the follow-up interview is an integral part of the M-CHAT procedure, it is reasonable to calculate PPV based on the combined M-CHAT plus interview score. Therefore, the PPV for M-CHAT plus interview was calculated as 21 of 37 screen-positive completed cases (excluding the four cases who passed the M-CHAT or interview but were evaluated based on the pediatrician's concerns of ASD), which brings the PPV to .57.

Discussion

This study demonstrates the feasibility of Level 1 screening for ASD in the primary care setting. To date, 21 children have been identified with ASD from a sample of 4797 toddlers, based on the results of a brief screening instrument completed during the 15-, 18-, or 24-month well-child visit. It is of note that only four of the 21 children diagnosed with ASD were flagged by healthcare providers (in addition to five false-positive pediatrician flags), suggesting that standardized screening measures are critical for the early detection of ASD to supplement other pediatric practices. In these four cases, the M-CHAT and follow-up

interview also indicated risk for ASD. Pediatrician concern in the absence of M-CHAT plus interview risk led to four additional evaluations; however, none of these children received a diagnosis of ASD (please note that the clinician was blind to M-CHAT total score and whether the pediatrician flagged the M-CHAT until after the diagnostic evaluation was complete, although the clinician was aware that most children evaluated demonstrated risk for ASD on the M-CHAT). Therefore, these preliminary findings suggest that developmental surveillance alone is not sufficient to identify all children with ASD, and does not reduce the false negative rate of standardized screening with the M-CHAT. This leads to the conclusion that screening will improve the pediatrician's ability to refer toddlers at risk for ASD. However, it also highlights that at-risk scores on the M-CHAT plus interview, combined with the pediatrician's flag for ASD concerns, were highly accurate in identifying ASD cases (80% of cases for whom both M-CHAT plus interview and pediatrician concerns were indicated were diagnosed with ASD). Although the M-CHAT and follow-up interview identified a number of false-positive cases (16 of the non-ASD cases were identified based on M-CHAT plus interview), 17 of the 20 false-positive cases demonstrated significant language or global cognitive delays, which warranted intervention. The typically developing children had parents who were extremely vigilant to ASD concerns; for example, one child had an older sibling with ASD, and another had a cousin with ASD.

The group of children diagnosed with ASD had higher M-CHAT total and critical scores, and higher follow-up interview critical scores, than the children who were not diagnosed with ASD. In order to rule out the possibility that the three typically developing children impacted these analyses, the t-tests were run excluding these three participants; M-CHAT total, interview total, and interview critical scores remained statistically significant ($p = .017-.049$), and M-CHAT critical score showed a trend toward significance ($p = .062$). The effect sizes for these analyses are consistent with the findings from the sample presented in Robins and colleagues (2001). The ASD group demonstrated significantly more severe impairment compared to the non-ASD group on all Mullen scales (fine motor, visual reception, receptive language, expressive language, and early learning composite), and the communication domain and adaptive behavior composite on the Vineland-II, with a trend toward significance for the socialization domain. This suggests that even when a sample is selected for further evaluation based on M-CHAT ASD risk, the children with ASD are likely to show greater impairment across clinical domains of functioning. Evidence for greater clinical impairment in ASD relative to non-ASD delays is consistent with Ventola and colleagues (2006), who examined a much larger mixed-stage (i.e. Level 1 and Level 2) sample of children screened using the M-CHAT.

Although this was not a prevalence study, the topic merits consideration. In comparison to the range of 6–11 per 1000 described in the literature (Baird et al., 2006; Fombonne, 2006; Harrison et al., 2006; Moldin and Rubenstein, 2006; Williams et al., 2006), the current study identified four cases per 1000. However, not all children who screened positive continued participation throughout the entire study. Although it is impossible to know whether those participants who discontinued participation would show the same rate of ASD, it is of note that there were no significant differences between the M-CHAT and follow-up interview scores of those who completed the study and those who did not complete the study. If one

were to presume that the rate of ASD is similar for the 22 percent of cases who did not complete the follow-up interview, it is possible that one child with ASD may have been lost due to incomplete data. Furthermore, given that 24 families continued to show risk for ASD after the M-CHAT follow-up interview but did not attend the diagnostic evaluation, it is likely that several more children with ASD were lost due to incomplete data, based on the data that 21 of the 41 evaluated cases were diagnosed with ASD. Therefore, it is possible that were we to ascertain 100 percent participation from all toddlers who demonstrated ASD risk on the M-CHAT, the identification rate by the M-CHAT would fall within the range of published prevalence rates of ASD.

However, one also must consider that participation bias may have led to an even higher ASD prevalence among this sample than would be predicted from population-based prevalence estimates. It is important to note that this conjecture does not replace measurement of the M-CHAT's sensitivity, which will be accomplished by rescreening the entire sample at age 4.

Criticism about current recommendations for ASD screening (Williams and Brayne, 2006) highlights that ASD screening research is still in the early stages of development and cross-validation, that better case definition is required before successful screening programs can be developed, and that screening instruments need further study before they are recommended for use (Mawle and Griffiths, 2006). It can be argued that the diagnostic criteria in the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2000) provide sufficient case definitions for autism and PDD-NOS, the two disorders on the autism spectrum likely to be detected in toddlers. It is true that disorders that are behaviorally defined will not have the same level of agreement as biologically defined disorders, but this should not be an obstacle to screening. It is unlikely that instruments with perfect sensitivity and specificity can be developed for disorders like ASD that are behaviorally defined. There will always be a tradeoff between false positives (which reduce positive predictive value and specificity) and false negatives (which reduce sensitivity). It would be a tremendous clinical disservice to assume that screening should not be recommended until the research supporting specific screening instruments and procedures is unequivocal; it will take several more years before complete follow-up data are available for large samples of children screened at toddler pediatric visits. Furthermore, sensitivity and specificity may not be the 'gold standard' by which to evaluate the utility of a screening instrument (Camp, 2006). An instrument that is able to detect some cases of ASD earlier than would be identified by general clinical practice provides the advantage of earlier referral for diagnostic evaluation and intervention. The psychometric properties of the M-CHAT have not been fully studied in an exclusively Level 1 sample; however, preliminary findings indicate that the M-CHAT does identify a significant number of ASD cases during toddler pediatric check-ups (20 of 3309 in Kleinman et al. (2008), and 21 of 4797 in the current study). Although the PPV is not as high as was hoped, particularly from the M-CHAT alone without the follow-up interview, the PPV of the M-CHAT plus follow-up interview is .57 in the current study and .65 in Kleinman et al. (2008). It also is of note that nearly all cases identified by the M-CHAT plus interview are children with significant delays who warrant intervention. Therefore it is reasonable, as Bryson and colleagues (2003)

urge, to use caution in interpreting screening results until additional data are published, but to advocate Level 1 screening for ASD in all toddlers.

This conclusion is consistent with the recent AAP policy statement (Dubey et al., 2006) and autism-specific screening guidelines (Johnson and Myers, 2007), which call for ASD-specific screening at all 18- and 24-month well-child visits, alongside routine ASD surveillance and broadband screening for other developmental disorders. The M-CHAT is a promising standardized instrument to facilitate Level 1 ASD screening, and can be administered with minimal intrusion on the healthcare provider's office practice. Although not expected to detect non-ASD delays, the M-CHAT can be used in conjunction with a general screening instrument to maximize early detection of ASD. Furthermore, the structured follow-up interview can be administered by a primary healthcare provider or paraprofessional during the well-child visit, in order to determine risk for ASD and need for referral immediately.

This study demonstrates the utility of Level 1 screening for ASD in toddlers using the M-CHAT in the primary care setting. It also reframes the M-CHAT as a two-step screening instrument; the use of the paper-and-pencil M-CHAT screen without the follow-up interview is not advocated in the primary care setting at this time. Approximately 90 percent of children will not demonstrate risk on the M-CHAT; however, for the 10 percent who do show risk on the initial M-CHAT screen, use of the follow-up interview should be integrated into the well-child visit. The interview has been designed for use by professionals and paraprofessionals with minimal experience with ASD, and can usually be conducted in 5–15 minutes. Use of the interview brings the PPV to .57, which is a moderate level. Future research will further evaluate the psychometric properties of the instrument in a large-scale longitudinal Level 1 screening sample, and will directly compare the sensitivity of the M-CHAT with broadband developmental screening, in order to provide empirical support for the recommendation for Level 1 screening using an ASD-specific instrument as a supplement to general developmental screening.

Acknowledgements

This article would not have been possible without the support and participation of the healthcare providers and families in the metro-Atlanta, GA area. I would also like to thank my colleagues Deborah Fein, Lauren Adamson, and Ann Hazzard for support, and the graduate, post-baccalaureate, and under-graduate students in my lab for assistance with data collection and management, particularly Sharlet Anderson, Nicolle Angeli, Lama Farran, Kimberly Oliver, Lisa Wiggins, AnneMarie Newman, Lauren Stites, Jamie Zaj, Lyntovia Ashe, Jasmine Brigham, Esther Choi, Shelley Hinkle, Steve Hirt, Amy Jones, Margaret Jones, Puja Joshi, Steve Kohn, Cassie Lovett, Molly Locklear, Melissa Nikolic, Cristina Parfene, Catherine Shelton, Gina Vanegas, and Amber Wimsatt. The principal investigator (DLR) has had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Finally, I would like to acknowledge grant 2R56HD 035612–10A1, which provided partial support for this study.

References

- Allison, C.; Baron-Cohen, S.; Wheelwright, S.; Charman, T.; Brayne, C. Development of the Q-CHAT, a revised screening instrument for autism spectrum conditions in toddlers between 18–24 months. Paper presented at the International Meeting for Autism Research; Montreal, Canada. 2006 Jun. 2006
- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition, Text Revision ed.). Arlington, VA: American Psychiatric Association; 2000.

- Baird G, Charman T, Baron-Cohen S, Cox A, Swettenham J, Wheelwright S, et al. A screening instrument for autism at 18 months of age: A 6-year follow-up study. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2000; 39(6):694–702. [PubMed: 10846303]
- Baird G, Simonoff E, Pickles A, Chandler S, Loucas T, Meldrum D, et al. Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *Lancet*. 2006; 368(9531):210–215. [PubMed: 16844490]
- Baron-Cohen S, Allen J, Gillberg C. Can Autism Be Detected at 18 Months - the Needle, the Haystack, and the CHAT. *British Journal of Psychiatry*. 1992; 161:839–843. [PubMed: 1483172]
- Baron-Cohen S, Cox A, Baird G, Swettenham J, Nightengale N, Morgan K, et al. Psychological markers in the detection of autism in infancy in a large population. *British Journal of Psychiatry*. 1996; 168(2):158–163. [PubMed: 8837904]
- Bryson SE, Rogers SJ, Fombonne E. Autism spectrum disorders: Early detection, intervention, education, and psychopharmacological management. *Canadian Journal of Psychiatry-Revue Canadienne De Psychiatrie*. 2003; 48(8):506–516. [PubMed: 14574826]
- Camp BW. What the clinician really needs to know: Questioning the clinical usefulness of sensitivity and specificity in studies of screening tests. *Journal of Developmental and Behavioral Pediatrics*. 2006; 27(3):226–230. [PubMed: 16775521]
- Carr JE, LeBlanc LA. Autism spectrum disorders in early childhood: An overview for practicing physicians. *Primary Care: Clinics in Office Practice*. 2007; 34:343–359. [PubMed: 17666231]
- Chakrabarti S, Haubus C, Dugmore S, Orgill G, Devine F. A model of early detection and diagnosis of autism spectrum disorder in young children. *Infants and Young Children*. 2005; 18(3):200–211.
- Charman T, Taylor E, Drew A, Cockerill H, Brown JA, Baird G. Outcome at 7 years of children diagnosed with autism at age 2: predictive validity of assessments conducted at 2 and 3 years of age and pattern of symptom change over time. *Journal of Child Psychology and Psychiatry*. 2005; 46(5):500–513. [PubMed: 15845130]
- Coonrod, EE.; Stone, WL. Screening for autism in young children. In: Volkmar, FR.; Paul, R.; Klin, A.; Cohen, D., editors. *Handbook of autism and pervasive developmental disorders, Vol. 2: Assessment, interventions, and policy*. 3rd. Vol. 2. Hoboken, NJ: John Wiley & Sons, Inc.; 2005. p. 707-729.
- Cox A, Klein K, Charman T, Baird G, Baron-Cohen S, Swettenham J, et al. Autism spectrum disorders at 20 and 42 months of age: Stability of clinical and ADI-R diagnosis. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. 1999; 40(5):719–732.
- Dawson, G.; Osterling, J. Early intervention in autism: Effectiveness and common elements of current approaches. In: Guralnick, MJ., editor. *The Effectiveness of Early Intervention*. Baltimore, MD: Paul H. Brookes; 1997. p. 307-326.
- Dawson G, Ashman SB, Carver LJ. The role of early experience in shaping behavioral and brain development and its implications for social policy. *Development and Psychopathology*. 2000; 12(4):695–712. [PubMed: 11202040]
- Dosreis S, Weiner CL. Autism spectrum disorder screening and management practices among general pediatric providers. *Journal of Developmental and Behavioral Pediatrics*. 2006; 27(2):S88–S94. [PubMed: 16685190]
- Duby JC, Lipkin PH, Macias MM, Wegner LM, Duncan P, Hagan JF, et al. Identifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. *Pediatrics*. 2006; 118(1):405–420. [PubMed: 16818591]
- Dumont-Mathieu T, Fein D. Screening for autism in young children: The modified checklist for autism in toddlers (M-CHAT) and other measures. *Mental Retardation and Developmental Disabilities Research Reviews*. 2005; 11(3):253–262. [PubMed: 16161090]
- Filipek PA, Accardo PJ, Baranek GT, Cook EH, Dawson G, Gordon B, et al. The screening and diagnosis of autistic spectrum disorders. *Journal of Autism and Developmental Disorders*. 1999; 29(6):439–484. [PubMed: 10638459]
- Filipek PA, Accardo PJ, Ashwal S, Baranek GT, Cook EH, Dawson G, et al. Practice parameter: Screening and diagnosis of autism - Report of the Quality Standards Subcommittee of the

- American Academy of Neurology and the Child Neurology Society. *Neurology*. 2000; 55(4):468–479. [PubMed: 10953176]
- Fombonne, E. Past and future perspectives on autism epidemiology. In: MSO; Rubenstein, JLR., editors. *Understanding Autism: From Basic Neuroscience to Treatment*. Boca Raton, FL: Taylor & Francis Group; 2006. p. 25-48.
- Freeman BJ, Cronin P. Diagnosing autism spectrum disorder in young children: An update. *Infants and Young Children*. 2002; 14(3):1–10.
- Gabovitch, EM.; Wiseman, ND. Early identification of autism spectrum disorders. In: Zager, D., editor. *Autism spectrum disorders: Identification, education, and treatment*. 3rd. Mahwah, NJ: Lawrence Erlbaum Associates; 2005. p. 145-172.
- Gravel JS, Marttila J, Nerbonne MA, Nozza RJ, Scott DM, Smedley T, et al. Report on audiologic screening. *American Journal of Audiology*. 1993; 4:24–40. [PubMed: 26650023]
- Gupta VB, Hyman SL, Johnson CP, et al. Identifying children with autism early. *Pediatrics*. 2007; 119:152–153. [PubMed: 17200280]
- Harris SL, Handleman JS. Age and IQ at intake as predictors of placement for young children with autism: A four- to six-year follow-up. *Journal of Autism and Developmental Disorders*. 2000; 30(2):137–142. [PubMed: 10832778]
- Harrison MJ, O'Hare AE, Campbell H, Adamson A, McNeillage J. Prevalence of autistic spectrum disorders in Lothian, Scotland: an estimate using the “capture-recapture” technique. *Archives of Disease in Childhood*. 2006; 91(1):16–19. [PubMed: 15886261]
- Holzer L, Mihailescu R, Rodrigues-Degauff C, Junier L, Muller-Nix C, Halfon O, et al. Community introduction of practice parameters for autistic spectrum disorders: Advancing early recognition. *Journal of Autism and Developmental Disorders*. 2006; 36(2):249–262. [PubMed: 16447011]
- Howlin P. The effectiveness of interventions for children with autism. *Journal of Neural Transmission-Supplement*. 2005; (69):101–119. [PubMed: 16355605]
- Jensen VK, Sinclair LV. Treatment of autism in young children: Behavioral intervention and applied behavior analysis. *Infants and Young Children*. 2002; 14(4):42–52.
- Johnson CP, Myers SM. Identification and evaluation of children with autism spectrum disorders. *Pediatrics*. 2007; 120(5):1183–1215. [PubMed: 17967920]
- Kasari C, Freeman S, Paparella T. Joint attention and symbolic play in young children with autism: a randomized controlled intervention study. *Journal of Child Psychology and Psychiatry*. 2006; 47(6):611–620. [PubMed: 16712638]
- Kleinman JM, Robins DL, Ventola PE, Pandey J, Boorstein H, Esser E, et al. 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*. 38(5):827–39. [PubMed: 17882539]
- Koegel LK, Koegel RL, Nefdt N, Fredeen R, Klein EF, Bruinsma YEM. First STEP - A Model for the early identification of children with autism spectrum disorders. *Journal of Positive Behavior Interventions*. 2005; 7(4):247–252.
- Lord C. Follow-up of two-year-olds referred for possible autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. 1995; 36(8):1365–1382.
- Lord C. Why is more serious scientific consideration not given to alternative and less conventional theories and treatments in the field of autism? *Journal of Autism and Developmental Disorders*. 1997; 27(3):349–349.
- Lord, C.; Rutter, M.; DiLavore, PC.; Risi, S. *Autism Diagnostic Observation Schedule - WPS Edition*. Los Angeles: Western Psychological Services; 1999.
- Lord C, Risi S, DiLavore PS, Shulman C, Thurm A, Pickles A. Autism from 2 to 9 years of age. *Archives of General Psychiatry*. 2006; 63(6):694–701. [PubMed: 16754843]
- Mawle E, Griffiths P. Screening for autism in pre-school children in primary care: Systematic review of English Language tools. *International Journal of Nursing Studies*. 2006; 43(5):623–636. [PubMed: 16387303]
- McConnell SR. Interventions to facilitate social interaction for young children with autism: Review of available research and recommendations for educational intervention and future research. *Journal of Autism and Developmental Disorders*. 2002; 32(5):351–372. [PubMed: 12463515]

- McGee GG, Morrier MJ, Daly T. An incidental teaching approach to early intervention for toddlers with autism. *Journal of the Association for Persons with Severe Handicaps*. 1999; 24(3):133–146.
- McGovern CW, Sigman M. Continuity and change from early childhood to adolescence in autism. *Journal of Child Psychology and Psychiatry*. 2005; 46(4):401–408. [PubMed: 15819649]
- Moldin, SO.; Rubenstein, JLR., editors. *Understanding Autism: From Basic Neuroscience to Treatment*. Boca Raton, FL: Taylor & Francis Group; 2006.
- Moore V, Goodson S. How well does early diagnosis of autism stand the test of time? Follow-up study of children assessed for autism at age 2 and development of an early diagnostic service. *Autism*. 2003; 7(1):47–63. [PubMed: 12638764]
- Mullen, E. *The Mullen Scales of Early Learning*. Circle Pine, MN: American Guidance; 1995.
- Nadel S, Poss JE. Early detection of autism spectrum disorders: Screening between 12 and 24 months of age. *Journal of the American Academy of Nurse Practitioners*. 2007; 19(8):408–417. [PubMed: 17655570]
- Robins DL, Dumont-Mathieu TM. Early screening for autism spectrum disorders: Update on the modified checklist for autism in toddlers and other measures. *Journal of Developmental and Behavioral Pediatrics*. 2006; 27(2):S111–S119. [PubMed: 16685177]
- Robins DL, Fein D, Barton ML. *Modified Checklist for Autism in Toddlers*. 1999 self-published.
- Robins DL, Fein D, Barton M. *Modified Checklist for Autism in Toddlers (M-CHAT) Follow-up Interview*. 1999 self-published.
- Robins DL, Fein D, Barton ML, Green JA. 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*. 2001; 31(2):131–144. [PubMed: 11450812]
- Rogers SJ. Brief report: Early intervention in autism. *Journal of Autism and Developmental Disorders*. 1996; 26(2):243–246. [PubMed: 8744493]
- Rutter, M.; Bailey, A.; Lord, C. *The Social Communication Questionnaire (SCQ)*. Los Angeles: Western Psychological Services; 2003.
- Rutter, M.; LeCouteur, A.; Lord, C. *Autism Diagnostic Interview, Revised*. Los Angeles: Western Psychological Services; 2003.
- Sand N, Silverstein M, Glascoe FP, Gupta VB, Tonniges TP, O'Connor KG. Pediatricians' reported practices regarding developmental screening: Do guidelines work? Do they help? *Pediatrics*. 2005; 116(1):174–179. [PubMed: 15995049]
- Schopler, E.; Reichler, RJ.; Renner, BR. *The Childhood Autism Rating Scale (CARS)*. Los Angeles: Western Psychological Services; 1988.
- Siegel, B. *Pervasive Developmental Disorders Screening Test-II (PDDST-II)*. San Antonio, TX: Harcourt Assessment, Inc.; 2004.
- Sparrow, SS.; Cicchetti, DV.; Balla, DA. *Vineland Adaptive Behavior Scales, Second Edition (Vineland-II)*. Bloomington, MN: Pearson Assessment; 2004.
- Stone WL, Lee EB, Ashford L, Brissie J, Hepburn SL, Coonrod EE, et al. Can autism be diagnosed accurately in children under 3 years? *Journal of Child Psychology and Psychiatry and Allied Disciplines*. 1999; 40(2):219–226.
- Ventola P, Kleinman J, Pandey J, Wilson L, Esser E, Boorstein H, et al. Differentiating between autism spectrum disorders and other developmental disabilities in children who failed a screening instrument for ASD. *Journal of Autism and Developmental Disorders*. 2006; 37(3):425–436. [PubMed: 16897377]
- Williams J, Brayne C. Screening for autism spectrum disorders - What is the evidence? *Autism*. 2006; 10(1):11–35. [PubMed: 16522708]
- Williams JG, Higgins JPT, Brayne CEG. Systematic review of prevalence studies of autism spectrum disorders. *Archives of Disease in Childhood*. 2006; 91(1):8–15. [PubMed: 15863467]
- Woods JJ, Wetherby AM. Early identification of and intervention for infants and toddlers who are at risk for autism spectrum disorder. *Language Speech and Hearing Services in Schools*. 2003; 34(3):180–193.

Yoder P, Stone WL. Randomized comparison of two communication interventions for preschoolers with autism spectrum disorders. *Journal of Consulting and Clinical Psychology*. 2006; 74(3):426–435. [PubMed: 16822100]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

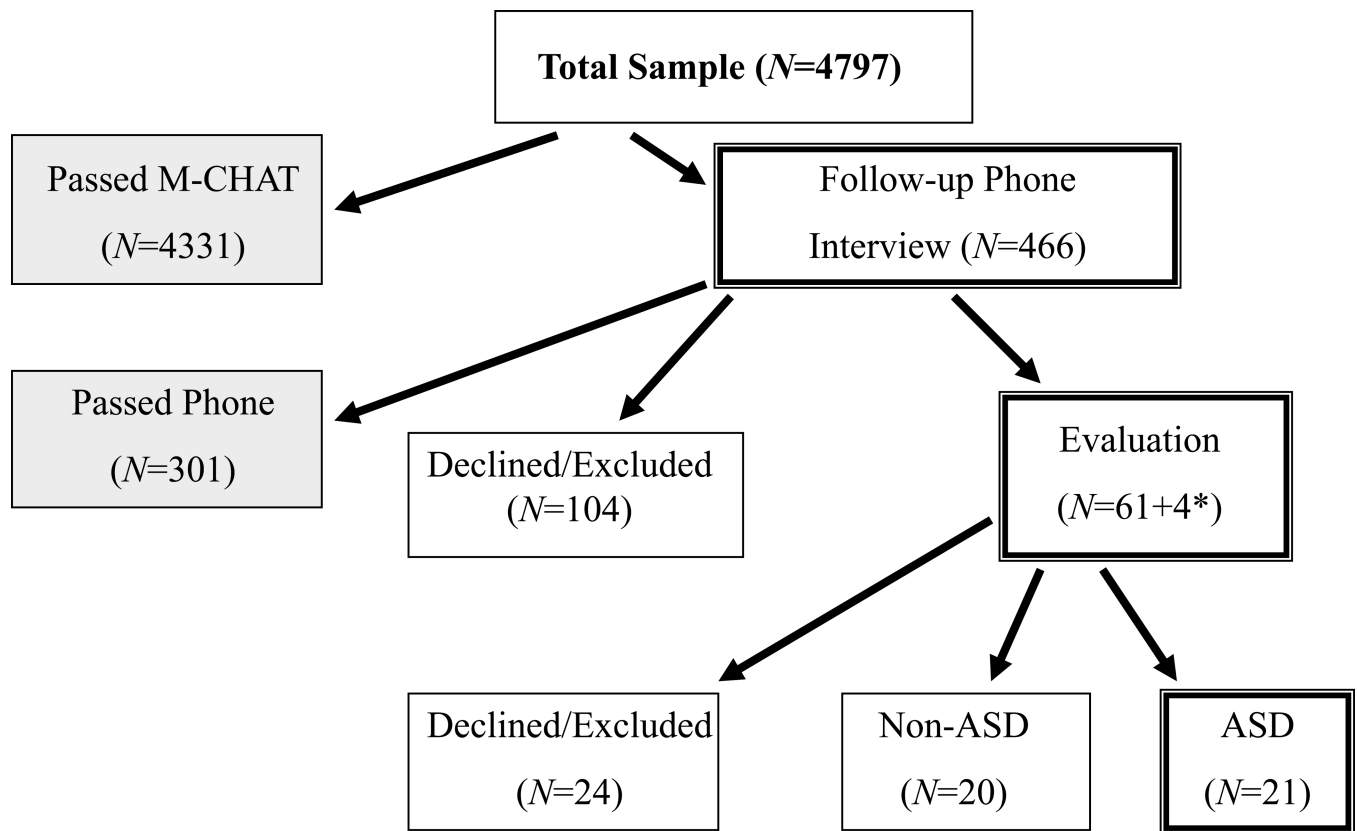


Figure 1. Flowchart of participation, based on risk for ASD

* Four additional cases were invited for evaluation based on the physician's concern, though the child passed the M-CHAT or Follow-Up Interview. None of these children was diagnosed with ASD.

Table 1

M-CHAT and follow-up interview scores for children evaluated based on M-CHAT risk

	NonASD (n=20)	ASD (n=21)	<i>t</i>	<i>Eta</i> ²
M-CHAT total (<i>SD</i>)	5.50 (3.33)	7.95 (3.15)	-2.421*	.131
M-CHAT critical (<i>SD</i>)	1.95 (1.40)	3.05 (1.53)	-2.395*	.128
Follow-up Interview total (<i>SD</i>)	4.83 (2.64)	6.52 (3.06)	-1.831 ^{<i>t</i>}	.083
Follow-up Interview critical (<i>SD</i>)	1.67 (1.03)	2.76 (1.55)	-2.557*	.150

*
p<.05^{*t*}
p<.1

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Clinical data for all evaluated children.

Scale	Subtest	NonASD (n=20)	ASD (n=21)	t	Etat ²
Mullen Scales of Early Learning					
	Fine Motor T (SD)	39.50 (11.95)	27.24 (10.56)	3.485**	.237
	Visual Reception T (SD)	43.80 (17.61)	26.81 (10.02)	3.821**	.272
	Receptive Language T (SD)	41.80 (17.42)	23.95 (8.79)	4.110**	.378
	Expressive Language T (SD)	34.55 (10.34)	27.48 (10.58)	2.164*	.107
	Early Learning Composite SS (SD)	90.95 (41.75)	66.29 (18.20)	2.573*	.145
Vineland Adaptive Behavior Scales - II					
	Communication SS (SD)	88.60 (12.30)	79.24 (13.22)	2.344*	.123
	Daily Living SS (SD)	95.65 (12.30)	91.05 (10.62)	1.265	.039
	Socialization SS (SD)	87.60 (12.82)	81.62 (7.92)	1.807 ^t	.077
	Motor SS (SD)	93.80 (10.60)	92.62 (9.21)	.381	.004
	Adaptive Behavior Composite SS (SD)	89.55 (10.23)	83.43 (8.06)	2.134*	.105

**
p<.01

*
p<.05

^t
p<.1

Note. Scales on the MSEL are measured by T-scores, the Early Learning Composite is a Standard Score. Scales on the Vineland-II are measured by Standard Scores.