Published in final edited form as:

J Autism Dev Disord. 2016 November ; 46(11): 3392–3407. doi:10.1007/s10803-016-2875-5.

Use of the Modified Checklist for Autism, Revised with Follow up-Albanian to Screen for ASD in Albania

Laura Brennan, MA,

Dept. of Psychology, University of Connecticut, Storrs, CT

Deborah Fein, Ph.D.² [Professor],

Depts. of Psychology and Pediatrics, University of Connecticut, Storrs, CT, tel: 413-519-1324, fax: 860-486-2760

Ariel Como, MD [Medical Director],

Tirana Medical University, Medical Director, “Domenick Scaglione” Regional Center for Autism, Tirana, Albania

Iris Carcani Rathwell, M.D. [Consultant Child and Adolescent Psychiatrist], and

South London and Maudsley Foundation NHS Trust, London

Chi-Ming Chen, Ph.D. [Assistant Professor]

Department of Psychology, University of Connecticut, Storrs, CT

Chi-Ming Chen: chi-ming.chen@uconn.edu

Abstract

The Modified Checklist for Autism in Toddlers Revised-Albanian screener (M-CHAT-R/-A) was used to screen 2,594 toddlers, aged 16-30 months, at well-child visits. Two hundred fifty three (9.75%) screened positive; follow up on failed items were conducted by phone with 127 (50%); the remainder were lost to follow-up. Twenty-six toddlers (21%) continued to screen positive; 19 received full evaluations, which assessed for ASD with the Autism Diagnostic Observation Schedule and developmental delays with the Parents Assessment of Developmental Status – Developmental Milestones. All evaluated children had significant delays; 17 of the 19 met criteria for Autism/ASD. Removal of three items improved performance. Although Albania and the US are quite different in culture and language, key features of autism appeared very similar.

²Correspondence concerning this article should be addressed to Deborah Fein, Dept. of Psychology, 406 Babbidge Rd., Storrs, CT 06268, 413-5191-324, deborah.fein@uconn.edu.

No other conflict of interest exists.

LB oversaw data collection, analyzed data, and wrote the dissertation on which this paper is based.

DF conceived of the study, wrote the grant which funded it, oversaw data collection and analysis, and contributed to writing of the paper.

AC conceived of the study, directed the data collection in Albania, trained the pediatricians on screening, and contributed to writing of the paper.

I C-R trained clinical staff on the ADOS, oversaw the translation process, and contributed to writing of the paper.

C-MC contributed to data analyses

This paper was based on the doctoral dissertation of Laura Brennan.

Keywords

autism; screening; Albania; M-CHAT-R/F; M-CHAT-R

Introduction

Significant disparities exist in the accessibility and quality of mental health services across the world (Saxena, Thornicroft, Knapp, & Whiteford, 2007). The availability of mental health resources depends largely on a country's level of economic development; individuals from lower- and middle-income countries (LMICs) have fewer mental health services than those from high-income countries (HICs; Patel, Kieling, Maulik, & Divan, 2013).

Autism Spectrum Disorder (ASD) is one of many developmental disorders for which there is limited availability of diagnostic and treatment services. ASD is a neurodevelopmental disorder that significantly impacts child functioning across both HICs and LMICs (Whiteford, Degenhardt, Rehm, Baxter, Ferrari, Erskine, ... & Burstein, 2013). ASD is characterized by a set of behavioral characteristics including difficulties in social communication, as well as restricted, repetitive behaviors or interests (American Psychiatric Association, 2013). The ASD global burden of disease is estimated to account for 4.2% (3.2-5.3%) of all disability-adjusted life years (DALY), a measure of overall disease burden (Whiteford et al., 2013). While prevalence rates are well-established in North America and Europe, studies investigating the prevalence of ASD in LMICs are fewer and vary dramatically in approach (Elsabbagh, Divan, Koh, Kim, Kauchali, Marcín et al., 2012; Patel et al., 2008; Wallace, Fein, Rosanoff, Dawson, Hossain, Brennan et al., 2012). Current U.S. estimates suggest that 1 out of 68 children carry an ASD diagnosis (Center for Disease Control, 2014). Recent global ASD prevalence estimates indicate 62 out of 10,000 children meet criteria for ASD; however, large regions such as South East Asia, Africa, and eastern areas of the Mediterranean were not represented in this estimate (Elsabbagh et al., 2012). Limited diagnostic resources in LMICs, and limited access to reliable, valid, and culturally appropriate screening and assessment tools, often restrict epidemiological studies such as those included in this estimate of global ASD prevalence (Elsabbagh et al., 2012; Soto, Linas, Jacobstein, Biel, Migdal, & Anthony, 2014; Wallace et al., 2012).

Recently, the number of ASD-specific organizations and the number of studies published in LMICs have increased (Daley, Singhal, & Krishnamurthy, 2013). The current study is the result of a program funded in Albania by Autism Speaks' Global Autism Public Health (GAPH) Initiative. This program aimed at increasing awareness of ASD in the region, improving early screening and diagnosis of ASD, training therapists in applied behavior analysis and developing video materials for a parent-training program. When this project was initiated, Albania was considered by the World Bank to be a lower middle-income country (The World Bank, 2012). However, in 2013, Albania's GNI per capita crossed the threshold into classification as upper middle-income country (The World Bank, 2013; see below). Despite this increase in GNI, Albania remains one of the poorest European countries (USAID, 2011), and the services available to parents of children with ASD outside of the

Albanian Children's Foundation and the Regional Center for Autism, where this study was conducted, remained very limited.

There is a clear need for low-cost, effective measures for detecting and diagnosing ASD in children living in countries with any income level, and especially those of low and middle incomes. It is now well documented that ASD can be reliably diagnosed in children aged 18-24 months (Kleinman, Ventola, Pandey, Verbalis, Barton, Hodgson ... & Fein, 2008) and even earlier (Zweigenbaum, Bryson, Brian, Smith, Roberts, Szatmari, Roncadin, Garon, & Vaillancourt, 2015), and that early identification and intervention are associated with more positive outcomes (Anderson, Liang, & Lord, 2014; Johnson & Myers, 2007; Orinstein et al, 2014; Robins and Dumont-Mathieu 2006; Rogers, 1996). The use of parent-report screening instruments is widely regarded as a first step in detecting young children at risk for ASD (Soto et al., 2014; National Research Council (US) Committee on Educational Interventions for Children with Autism, 2001; Szatmari, Bryson, Boyle, et al., 2003).

The American Academy of Pediatrics (AAP) recommends using an ASD-specific screener at 18- and 24-month well-child care (WCC) visits to screen for ASD symptoms (Johnson & Myers, 2007; Robins, Casagrande, Barton, Chen, Dumont-Mathieu, & Fein, 2014). WCC screening for ASD provides an opportunity for children who regularly attend WCC visits to receive services as early as possible. Using a WCC screening approach followed by timely diagnostic evaluations can minimize disparities in age at diagnosis (Herlihy et al., 2014) that are otherwise found across different ethnicities or socio-economic levels (Mandell, Listerud, Levy, & Pinto-Martin, 2002; Robins et al., 2014). In 2015, the US Preventive Services Task Force questioned the benefits of universal screening, indicating insufficient evidence supports this practice (McPheeters, Weitlauf, Vehorn, Taylor, Sathe, Krishnaswami, ... & Warren, 2016). However, the AAP and researchers in the field of ASD screening, diagnosis, and treatment have responded to this conclusion, asserting that evidence is sufficient to recommend universal autism screening, to reduce the age of detection and promote more equitable access to intervention (AAP, 2016; Robins et al, 2016; Zwaigenbaum et al., 2015; Coury, D.L., 2015).

ASD Screening Tools

A number of validated measures are available to screen for ASD risk in children 18 months of age and older, including the Modified Checklist for Autism in Toddlers (M-CHAT; Chlebowski, Robins, Barton, & Fein, 2013; Robins, Fein, Barton, & Green, 2001) and the Infant Toddler Checklist (ITC; Johnson and Meyers, 2007; Zweigenbaum, Bryson, Lord, Rogers, Carter, Carver, ... & Yirmiya, 2009). A recently validated, revised version of the M-CHAT, the M-CHAT-Revised with Follow up (M-CHAT-R/F) has demonstrated strong psychometric properties, with estimates of sensitivity and specificity of .854 and 0.993 in a low-risk, population-based sample (Robins et al., 2014). For the purposes of this paper, the term M-CHAT-R-A will be used when referring to the screening and M-CHAT-R/F-A when referring to screening plus follow up.

Use of Translated and Culturally Adapted Versions of the M-CHAT and M-CHAT-R/F

No published studies to date have assessed the performance of the M-CHAT-R/F outside of the U.S., though it has been translated for use in at least 21 languages. Its predecessor, the M-CHAT, however, has been translated into at least 69 languages and has been found to have clinical utility in a number of international validation studies (Canal-Bedia, García-Primo, Martín-Cilleros, Santos-Borbujo, Guisuraga-Fernández, Herráez-García et al., 2010; Inada, Tomonori Koyama, Inokuchi, Kuroda, & Kamio, 2010; Kara, Mukaddes, Altinkaya, Guntepe, Gokcay, & Ozmen, 2014; Koh, Lim, Chan, Lin, Lim, Choo, & Magiati, 2014; Perera, Wijewardena, & Aluthwelage, 2009; Seif Eldin, Habib, Noufal, Farrag, Bazaid, Al-Sharbat, et al., 2008). Items involving imitation, pretend play, pointing, showing, and social referencing have been found to best differentiate diagnostic groups in studies conducted in Japan, China, and Singapore, suggesting a high degree of universality in the early symptom presentation of ASD (Inada et al., 2010; Koh et al., 2013; Wong, 2004). Although most studies of the original M-CHAT report good sensitivity, one Norwegian population study reported low sensitivity of 18 month M-CHAT assessed against diagnoses made later in childhood. There are several possible explanations for this: milder cases of ASD (PDD-NOS and Asperger's in DSM-IV), or ASD without intellectual or language impairment or requiring only some support (in DSM-V) might not be detectable at 18 months, because these children might meet early social and language milestones on time. In addition, in this study, diagnostic confirmation occurred years after the screener was completed. Given that symptom presentation may change over time, social and communication demands increase with age, and parent awareness of symptoms also likely change over time, the sensitivity of the parent-report screener at 18-months can be expected to be higher when diagnostic confirmation occurs concurrently with screening. In addition, the M-CHAT screener was added well into the Norwegian study, before which time, it contained several M-CHAT items but not the entirety of the screener. Finally, the follow-up phone interview was not completed in this study.

Limitations of translated screening and diagnosis methods

As noted above, researchers have begun to use translated versions of screening and diagnostic tools validated in HICs in an attempt to screen for ASD in middle- and low-income countries. This preserves resources within LMICs by providing a measure already shown to be useful, rather than developing new measures (Daley, Singhal, & Krishnamurthy, 2013; Soto et al., 2015). However, there is limited evidence on the utility of these measures in LMICs. The performance of such tools when used outside of the cultural and/or economic context in which they were validated depend on the assumption that ASD symptoms are “universal,”; there can be a limited appreciation of within-culture factors such as socioeconomic status (SES), literacy rates, and the impact of stigma associated with diagnosis (Daley et al., 2013; Matson, Worley, Fodstad, Chung, Suh, Jhin, ... & Furniss, 2011; Soto et al., 2015). A review of translated and culturally adapted ASD screeners by Soto and colleagues (2015) found that study authors who reported more rigorous adaptation procedures for a screener also more frequently suggested modifications including wording alterations, inclusion of more culturally appropriate information or examples, procedural revisions, or alternative scoring criteria (Soto et al., 2015).

The current study is the first international study to utilize the original two-stage screening procedures outlined by the M-CHAT-R/F. We used a translated version of the M-CHAT-R/F, the M-CHAT-R/F-Albanian (M-CHAT-R/F-A), administered to parents in Albania as part of routine WCC visits.

Background

The Republic of Albania lies in the southeastern region of Europe on the Balkan Peninsula (Institute of Statistics, Institute of Public Health [Albania] and ICF Macro, 2010). According to the WHO, over 98% of the Albanian population identified as ethnic Albanian, with the remaining 2% of the population identifying as Greek, Macedonian, Vlach, Roma, Bulgarian, or Serbian (WHO, 2015; Institute of Statistics, Institute of Public Health [Albania] and ICF Macro, 2010). The primary language is Albanian (Simmons and Slocum, 2014). According to UNICEF data, Albania has an adult literacy rate of 96.8% (UNICEF, 2013). Albania is currently a middle-income country, with a GNI per capita of \$4,460 (World Bank, 2014).

Albania's health care system is primarily public, with the state providing a majority of the health services, including prevention, diagnosis, treatment, and public health initiatives (Institute of Statistics, Institute of Public Health [Albania] and ICF Macro, 2010). In order to determine whether children have access to health services, public health research often looks to vaccination rates as a proxy for evaluating potential disparities in the health system (Gavi, n.d.). In Albania, the 2008-2009 Demographic and Health Survey revealed that 95% of children between 18-29 months received all of the necessary immunizations, with less than 1% receiving no immunizations at all (Institute of Statistics, Institute of Public Health [Albania] and ICF Macro, 2010). These findings suggest that most young children in Albania have access to primary care services.

Study Aims

The current study examined the internal consistency, reliability, and validity, of the M-CHAT-R/F-A in urban centers of Albania. Item-level analyses of the M-CHAT-R/F-A were also conducted to evaluate the ability of individual items to predict ASD symptoms in a low-risk sample of Albanian toddlers. Finally, the study examined potential modifications with the goal of a screening tool that best shows elevated risk for ASD while minimizing false positive rates, to reduce the time and cost of further evaluation.

Methods

Participants

A total of 3,133 children were screened using the M-CHAT-R-A during their well-child visits by one of 42 participating pediatricians in urban areas of Albania, primarily in Tirana. Children ($n=539$) were excluded for the following reasons: chronological age (CA) less than 16 months ($n=211$), CA greater than 36 months ($n=56$), no CA obtained due to missing birthdate or date of screening ($n=257$), and cases where protocol procedures were not followed correctly ($n=15$). Toddlers were only included if their parents spoke Albanian and provided informed consent at time of screening. The final sample included 2,594 toddlers

between the ages of 16 and 36 months, with a mean age at screening of 24 months ($SD = 2.83$ months).

Of the 2,594 toddlers included in the current sample, 50.1% were male ($n=1,300$), 49.7% were female ($n=1288$), and 0.2% did not have a sex indicated on their screener ($n=6$; see Table 1). Ethnicity data was not requested on the screener because Albanian collaborators raised concerns regarding the sensitivity of asking for such information and because the Albanian population is quite homogeneous. Ethnicity data for all children evaluated was collected ($n=20$), which indicated that 100% of children evaluated were identified as ethnic Albanian. A majority of M-CHAT-R-As (88%) were completed by mothers. Maternal education for toddlers screened was collected at time of screening (see Table 1). According to a 2008-2009 survey conducted in Albania, a majority of women in Albania (49%) have attained 8 years of education, with 1% reporting no education or incomplete primary school, 36% completing secondary school, and 13% with a university education or higher (Institute of Statistics, Institute of Public Health [Albania] and ICF Macro, 2010; see Table 1). Breaking the survey sample and the screened sample into individuals who completed secondary school vs. those who did not, the screened sample represents a higher level of maternal education than the population at large, ($X^2 = 13.36$, $p = 0.003$).

Procedures

The Institutional Review Board at the University of Connecticut provided oversight and approval of this study. All procedures were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. No ethical review boards existed in Albania at the time of this study, but Albanian personnel participating in data collection or study supervision completed University of Connecticut ethics training (in translation) and passed the ethics exam. Investigators provided pediatrician sites with a translated informed consents, a demographic form, the M-CHAT-R-A screener, and instructions to screen children between the ages of 16 and 36 months. An Albanian investigator (AC), a child psychiatrist, also provided training for participating pediatricians on characteristics of ASD in young children, as well as on how to administer the M-CHAT-R-A in their offices. Parents could decline participation by leaving the screener and demographic forms blank. Pediatrician sites did not track rates of refusal.

Screening sites then sent completed demographic forms and M-CHAT-R-As to the supervising site at the Regional Center for Autism in Tirana, Albania, affiliated with the Albanian Children's Foundation, where they were processed and scored. Study staff telephoned parents whose children screened positive on Stage 1 of screening to conduct the Stage 2 follow up interview, during which time both items failed and items left blank were queried. For 126 families, staff was unable to contact parents over the phone or parents refused to participate in Stage 2 of the screening process. If a child continued to screen positive on the M-CHAT-R/F-A after the interview, meaning the child continued to fail two or more M-CHAT-R/F-A items, the parents were offered a diagnostic/developmental evaluation free of charge.

Diagnostic evaluations were conducted by a team consisting of a child psychiatrist, a pediatrician, and an experienced therapist from the Regional Center. The children were assessed using the following measures: Autism Diagnostic Observation Schedule (ADOS) (Lord et al., 2008); the Parent Evaluation of Developmental Status (PEDS; Glascoe 1997; Glascoe 2004), PEDS: Developmental Milestones (PEDS: DM) (Glascoe and Robertshaw, 2008) and a semi-structured parent-report interview assessing presence of ASD symptoms. Clinicians administering the ADOS completed reliability training to research standards and received distance supervision from a child psychiatrist (ICR) with extensive experience in ADOS administration and expertise in autism. She is a native Albanian speaker who received psychiatric training in the United Kingdom. In addition to double-scoring videotaped ADOS administrations for the purposes of achieving research reliability, ICR also scored tapes on children for which evaluators felt the coding was unclear. For the current study, ADOS diagnostic outcome was used.

Parents were then provided verbal feedback and a brief written summary from the research staff about their child's test results.

Measures

In the current study, all measures were translated to Albanian with authors' permission, back-translated to English, and reconciled by a bilingual Albanian-English child psychiatrist with expertise in ASD and practicing neurodevelopmental psychiatry in the United Kingdom (ICR). All authors then approved the back-translation for each translated measure.

Modified Checklist for Autism in Toddlers, Revised Version with Follow up in Albanian—(M-CHA-T-R/F-A) is an Albanian translation of the M-CHAT-R/F (Robins et al., 2014), a modified version of the M-CHAT. It has 20 yes-no questions, with 3 reverse scored, and 3 foil items (e.g., Does your child walk?) and was filled out by caregivers at well child pediatric visits. Stage 1 of screening consists of the M-CHAT-R, a paper and pencil measure involving 20 yes/no questions that takes approximately five minutes to complete (Robins et al., 2014). Stage 2 of screening, the M-CHAT-R/F (follow up), involves a follow-up interview over the telephone or in person during which parents of screened positive children are asked questions about items failed and examples of concerning behaviors are solicited (Robins et al., 2014). This two-stage screening process yielded a positive predictive value (PPV) for receiving an ASD diagnosis of 0.475 and a PPV of 0.946 for any developmental delay or concern (Robins et al., 2014).

The 20 items on the M-CHAT-R screener (Stage 1) were organized to decrease agreement bias; three of these items were reverse scored, meaning these items were intended to be answered “no” as the “passing” response, for example, “have you ever wondered if your child is deaf?” (Robins et al., 2014). Three “foil” items pertaining to motor skills were also included to allow parents of delayed toddlers to answer “yes” to at least some items on the screener. A child is considered to screen positive on the M-CHAT-R/F if he or she fails three or more items on the screener during the first stage of screening. The child continues to screen positive if he or she fails two or more items on the follow-up interview during Stage 2.

Parent Evaluation of Developmental Status (PEDS; Glascoe, 1997)—The PEDS is a 10 item parent-report screening measure that assesses parental concern about receptive and expressive language abilities, social-emotional functioning, behavior, self-help skills, global/cognitive functioning, current health, and/or fine- and gross-motor skills. A child's level of risk for developmental difficulties is determined based on the number of reported parent concerns for items shown to be predictive of developmental issues, or “predictive concerns”. The PEDS is highly correlated with intellectual functioning, language skills, academic performance, and adaptive behavior abilities (Glascoe, 2003). It has also demonstrated sensitivity estimates between 74-78% and specificity estimates between 70-80% in detecting developmental issues in children between the ages of 0 to 8 years (Glascoe, 2003).

PEDS: Developmental Milestones (PEDS: DM)—(Glascoe & Robertshaw, 2008). Parents also filled out the PEDS: DM Assessment Level during diagnostic evaluations in the current study. This measure was designed to provide age-equivalents and percentages of delay for each developmental domain. Percentage of delay greater than 25% indicates clinically significant concerns. The PEDS: DM demonstrates sensitivity and specificity ranging from 80-93% in children 16-30 months, the age range of the current study. Interrater reliability, test-retest reliability and readability are excellent (Brothers et al., 2008). Using the PEDS together with the PEDS: DM provides a more complete assessment of developmental level, as it combines parent concerns with specific developmental milestones (Glascoe & Robertshaw, 2008).

Autism Diagnostic Observation Schedule (ADOS)—The ADOS is a standardized semi-structured interview using play-based methods to assess for the presence of ASD symptoms (Lord, Risi, Lambrecht, Cook, Leventhal, DiLavore, et al., 2000). The ADOS was found to have strong interrater-reliability for both Modules 1 and 2 (Lord et al., 2000), with inter-rater agreement for assigning ASD vs. non-spectrum diagnoses at 100% for Modules 1 and 3 and 91% for Module 2 (Lord et al., 2000). The ADOS results in one of three outcomes: Autism, Autism Spectrum (milder or incomplete symptom cases), and Non-Spectrum. We will use “ASD” to refer to the Autism and Autism Spectrum categories combined. The Albanian translation of the ADOS was approved by its publisher, Western Psychological Services.

Data Analysis

Validation of the M-CHAT-R/F-A was assessed through internal consistency analysis, using Chronbach's alpha, and positive predictive value (PPV) for ASD specifically and for developmental delays in general for both initial and follow-up screening. Item level analyses were conducted using discriminant function analyses to determine whether particular items on the M-CHAT-R/F-A were predictive of diagnostic outcome (e.g., “ASD/Autism” or “Other developmental concerns”). Potential modifications to improve performance of the M-CHAT-R/F-A were then examined.

Results

Screening Results (M-CHAT-R/F-A)

Of the final sample of 2,594 children screened, 253 (9.75%) screened positive on the M-CHAT-R-A (see Figure 1). Follow up interviews were conducted using the M-CHAT-R/F-A with 127 (50.2%) of the screened positive toddlers. The additional 126 were lost to follow-up due to parent refusal or being unable to contact. A total of 26 of the 127 toddlers who received the follow-up interview (21%) continued to screen positive on the follow-up (M-CHAT-R/F-A), qualifying them for a free developmental/diagnostic evaluation. Of those 26 eligible for an evaluation, 20 (77%) attended, and 6 refused (see Fig 1 for a flow chart; see Table 2 for a summary of total scores by screening outcome).

Of the 20 children evaluated, 19 received the ADOS; the remaining family moved out of the country in the middle of the evaluation. Of the 19 toddlers with an ADOS, 17 met criteria for Autism Spectrum ($n=3$) or Autism ($n=14$) and 2 were classified as Non-Spectrum (see Figure 1).

Characteristics of Evaluated Sample—Sex, mean age at screening, initial and follow-up scores and maternal education are shown in Table 3. All of the M-CHAT-R-A screeners ($n=20$) were completed by mothers at their child's pediatrician offices.

The average total score on the M-CHAT-R-A screening tool for toddlers who met ADOS criteria for ASD was 11.71 (out of a possible 20) ($SD=4.07$; see Table 2) while the average total score for the two Non-ASD children was 9.50 ($SD=2.12$). On the follow up interview, toddlers who met ASD criteria had a mean score of 9.18 ($SD=3.59$) and Non-Spectrum toddlers had a mean score of 6.00 ($SD=2.83$). Statistical analyses were not run because of the small number of children in the Non-ASD group ($n=2$).

Evaluation Data: ASD Group: A total of 17 children met criteria for ASD, with 3 children meeting criteria for Autism Spectrum and 14 meeting criteria for Autism. As expected, scores for children who met Autism criteria were higher (more symptomatic) than children in the Autism Spectrum group on all ADOS domains (see Table 2).

The results from the PEDS indicated that the toddlers who met ADOS criteria for Autism had an average of 2.79 ($SD=0.58$) predictive and 2.50 ($SD=1.61$) non-predictive concerns. The children who met Autism Spectrum criteria had an average of 2.33 ($SD=0.58$) predictive concerns and 2.67 ($SD=0.58$) non-predictive concerns. Each toddler in the ASD groups received a high-risk classification on the PEDS screening tool based on the number of predictive concerns.

On the PEDS: DM, a percentage of delay score greater than 25% indicates a clinically significant level of delay in a particular area. The Autism Spectrum and Autism groups demonstrated significant delays across most developmental domains, with the exception of fine and gross motor skills. As a group, the average percentage of delay score for toddlers who met ADOS Autism criteria exceeded 25% in all but one domain, with their average Gross Motor score falling just below 25% at 24.25% ($SD=20.87\%$). Toddlers in the Autism

Spectrum group had an average percentage of delay score greater than 25% in the Self-Help, Receptive Language, Expressive Language, and Social Emotional domains; however, their average percentage of delay values for the Fine and Gross Motor domains were both less than 25% (M=15.52%, SD=17.68 and M=4.81%, SD=32.16%, respectively).

Evaluation Data: Non-ASD Group: Two evaluated children were classified as Non-ASD on the ADOS. The mean score for these children on the Communication domain of the ADOS was 3.50 (SD = 2.12), which is above the cut-off score of 2 for Autism Spectrum. In the Reciprocal Social Interaction domain, the mean score was 3.00 (SD = 1.41), with a combined Communication and Reciprocal Social Interaction average score of 6.50 (SD = 0.71); these fell below the Autism Spectrum cut-off score of 7. The Non-ASD classified children had an average score of 3.00 (SD = 1.41) in the Play domain and 1.00 (SD = 1.41) in the Restricted Interests and Repetitive Behavior (RRB) domain. Non-ASD toddlers had lower scores for each ADOS domain than both ASD groups, with greater differences between the ASD and Non-ASD toddlers in the Reciprocal Social Interaction and Repetitive Behavior domains.

On the PEDS, the average number of predictive concerns for the Non-ASD classified children was 3 (SD = 0) with an average of 3.5 (SD = 0.71) non-predictive concerns. The Non-ASD toddlers demonstrated significant delays in each domain assessed, with the exception of Gross Motor (M = 5.02%, SD = 12.54%). Both children in the Non-ASD group also met PEDS criteria for high risk of developmental delay. On the PEDS: DM, the Non-ASD group had lower overall percentage of delay scores than both ASD groups for the Self Help (M = 41.35%, SD = 23.12%), Receptive Language (M = 47.12%, SD = 39.43%), and Social Emotional (M = 35.04%, SD = 10.27%) domains, indicating less severe impairments in these areas. However, they showed large delays in Expressive Language (M = 54.49%, SD = 17.22%). Gross Motor delays were mild (M = 5.02%, SD = 12.54%), but Fine Motor skills were quite delayed (M = 39.42%, SD = 20.47%). These results suggest that while these two children did not meet Autism criteria on the ADOS, they demonstrated significant developmental delays, including in communication and language.

Lost to Follow Up Toddlers: As mentioned above, 50% of children who screened positive on the M-CHAT-R-A screener did not receive a follow up interview. Chi square analyses comparing toddlers who received the M-CHAT-R/F-A with those lost to follow up indicated that the groups did not differ significantly in terms of sex ($X^2(1, N = 253) = 1.434, p = 0.231$) or maternal education ($X^2(5, N = 253) = 10.287, p = 0.067$). Age at screening for both groups was not normally distributed, with skewness of 0.531 (SE = 0.215) and kurtosis of 0.898 (SE = 0.427). Mean age at screening for toddlers lost to follow up (M = 24.09 months, SD = 3.65) and toddlers who received the M-CHAT-R/F-A (M = 24.36 months, SD = 3.98) were not significantly different, by Mann-Whitney U test. M-CHAT-R-A total scores for the two groups were also non-normally distributed, with skewness of 1.52 (SE = 0.215) and kurtosis of 1.45 (SE = 0.427). A Mann-Whitney U test indicated total scores on the M-CHAT-R-A for children who did not receive follow up via phone interview (M = 5.13, SD = 3.82) did not differ significantly from those who did complete the follow-up (M =

5.50, SD = 4.09; $U=7641.00$, $Z = -0.672$, $p = 0.501$). Therefore, attrition was not attributable to differences on the demographic factors assessed or the M-CHAT-R-A original score.

Refused Evaluation: Six children did not receive evaluations due to parent refusal. To determine whether these toddlers differed significantly from the children who received an evaluation, independent-sample t-tests and chi square analyses were used to compare the two groups. Toddlers who did not receive evaluations did not differ significantly from those who did on chronological age at screening, maternal education, sex, or total score on the M-CHAT-R-A. Thus, attrition at this stage of the study also seems nonsystematic.

Performance of M-CHAT-R/F-A

Reliability: Internal reliability for the 20 M-CHAT-R-A screener items was good ($\alpha = 0.737$). Reliability for the follow-up interview score was also high ($\alpha = 0.935$).

Positive Predictive Value (PPV): The current study used the original scoring criteria outlined in the Robins et al. (2014) validation of the M-CHAT-R/F in the U.S., with screen positive cut off scores of 3 on the screening tool and 2 on follow up. The PPV of the M-CHAT-R-A before follow up, for any developmental delay, including autism, was 0.16. The PPV of the screening tool for ASD based on ADOS results was 0.14. These figures do not include children who were lost to follow up for the phone interview or refused evaluation. After follow up interview, the PPV of the M-CHAT-R/F-A for ASD was 0.895 (95% CI: 0.65-0.98) and 0.10 (95% CI: 0.79-1.00) for non-autistic developmental delay (no typically developing children were found on evaluation).

Other Performance Factors: Screen negative rate of the original screener was 90%. Follow-up of these screen negative children was not possible, due to limited resources for evaluating children who screened negative. Of the children who screened positive initially, 79.5% went on to screen negative on the follow-up interview, reinforcing the necessity of reducing false positives with the follow-up interview. The combination of limited data in this area and the small sample size of evaluated children precluded examiners from using ROC curve analyses to evaluate sensitivity and specificity more precisely.

Prevalence of ASD Detected

Although the current study was not intended to be a population-based study of prevalence, we can estimate the percent of the population that screening and diagnostic procedures detected. Seventeen children met ADOS criteria for ASD out of a total 2,594. However, approximately 50% of the screened positive toddlers ($n = 126$) were lost to follow up, with no significant differences between children who received the follow-up interview and those who were lost to follow up. We therefore assumed these children would have met criteria for ASD at a similar rate if evaluated. Thus, prevalence was estimated by assuming that 34 children would have met criteria for ASD out of the total sample ($n = 2594$), resulting in an estimate of 1.31% of the population diagnosed with an ASD.

Item Response Stability

The frequency with which item responses changed between initial screening and follow up on the M-CHAT-R/F-A was also examined. Items that changed the least between the 2 stages of screening included 8, 1, and 6 (see Table 4). On these items, screen positive (failed) responses changed to screen negative (passed) at rates of 42%, 43%, and 45%, respectively. The largest change in initial failure rates on the screener and passing rates after follow up was found for Items 2, 5, and 12, with rates of failure decreasing by 98%, 89%, and 98%, respectively (see Table 4). These three items differed in several respects: they were the only reverse scored items; they touched on sensory issues, and they referred to positive symptoms rather than developmental delays (e.g., “Does your child make unusual finger movements near his/her eyes?”; see Table 5 for content of these items).

Discriminant Function Analysis

A discriminant function analysis (DFA) was used to evaluate the ability of the 20 items on the M-CHAT-R-A to predict ADOS outcome of an ASD. The sample was divided into two categories for comparison, with the first group comprised of toddlers who met ASD criteria on the ADOS ($n=17$) and the second consisting of toddlers who screened negative on the M-CHAT-R-A or M-CHAT-R/F-A ($n=2442$). DFA revealed one function, consisting of 12 items with strong predictability for an ASD (see Table 5 for items and standardized canonical discriminant function coefficients). These items consisted primarily of questions related to social interaction, including nonverbal communication, response to name, social smiling, shared interest in items or activities, receptive language abilities, eye contact, imitation, and interest in other children (see Table 5). Item 1, “If you point at something across the room, does your child look at it?” was most predictive of ASD, with Item 10, “Does your child respond when you call his or her name?” the second most predictive of ASD. Notably, Item 1 was one of the items that changed least frequently between initial screening and follow up. The least predictive items consisted primarily of the three foil items (Items 4, 13, and 20) assessing symptoms unrelated to ASD, and the three items that were reverse scored and referred to sensory symptoms (Items 2, 5, and 12, see Table 5). Of these 6 items, 5 of the least predictive items identified by the DFA were also items that changed most frequently during Stage 2 of screening, which further suggests that these items may be particularly hard for parents to understand. Though not a foil or reverse scored item, Item 6, “Does your child point with one finger to ask for something or to get help?” was also one of the least predictive items (0.086), perhaps because of shared variance with Item 1, involving pointing for joint attention. In addition, when administered during Stage 2 of screening, 97% of toddlers passed Item 6, suggesting that this item may not been understood by parents. Item 12, “Does your child get upset by everyday noises?” was revealed to be the least predictive of outcome (-0.035). While 47% of all parents endorsed this symptom on the screener, 98% of the toddlers who received Stage 2 of the M-CHAT-R/F-A went on to pass the item when administered over the phone (see Table 4). Based on its instability over the two stages of screening and minimal predictive value, Item 12 was the poorest performing item on the M-CHAT-R/F-A.

Removal of Three Problematic Items—The performance of the M-CHAT-R/F-A was then reevaluated after omitting the three items most frequently changed over the two stages

of screening and found to be the least predictive by DFA (Items 2, 5, 12; leaving the 3 foil items). Reliability for Stage 1 of screening increased from 0.737 on the 20-item version to 0.886 without the three omitted items, and remained stable for Stage 2 at 0.931 (compared to 0.935). In addition, screen negative rate on initial screening increased to 96% compared to 90% of the original M-CHAT-R-A. On the original screener, 20% of children who received the follow-up interview continued to screen positive; with the 3 items removed, this increased to 40% (see Table 6).

If these items were omitted, only two cases determined to require an evaluation by the original M-CHAT-R-A no longer met criteria for an evaluation. Since the evaluation was refused by these two families, their diagnosis is unknown. No child who was evaluated would have been missed.

Results also indicated a possible alternative cut-off score for Stage 1 of M-CHAT-R/F-A without the three problematic items. Examination of the frequency of total scores by screening outcome revealed that no child who received an evaluation and subsequent diagnosis scored lower than 5 items failed on Stage 1 of the currently suggested version of the screener. *This suggests a possible alternative scoring criterion of 5 total items failed on Stage 1 of the screener when Items 2, 5, and 12 are omitted, with the stage 2 cutoff of 2 failed items retained.*

With higher reliability, lower screen positive rate, and fewer children reverting to screen negative on the follow-up interview, the revised, 17 item version may be a more effective screening tool for ASD in Albanian. Of course, it is not possible to state this with confidence until the revised screener is tested with these changes already made.

Discussion

M-CHAT-R/F-A Performance

This is the first study conducted outside of the U.S. to examine the utility of a translated version of the recently validated M-CHAT-R/F. Both stages of the M-CHAT-R/F-A were translated and adapted for use in Albanian-speaking communities to detect ASD in toddlers aged 16-36 months at their WCC visits. Results indicated excellent reliability for the 2-Stage M-CHAT-R/F-A in the current sample ($\alpha = 0.935$), a figure higher than that of the original M-CHAT-R/F ($\alpha = 0.79$; Robins et al, 2014).

Of the 2,594 children initially screened with the M-CHAT-R-A, 9.75% screened positive; of those reached for follow-up, 20% continued to screen positive. The comparable figures in the US were 7.2% screen positive, and 37% continuing to screen positive on follow-up (Robins et al, 2014). In Albania, PPV for ASD at Stage 1 of screening was 0.14 (in the US the comparable figure was also 0.14). This figure increased dramatically after Stage 2 follow-up was conducted, yielding PPV of 0.89 for ASD (in the US, the comparable figure was .475; Robins et al, 2014). PPV for any developmental problem was 1.00 (but it is highly likely that this would be lower in a larger sample). The comparable PPV for any developmental problem in the US was 0.946 (Robins et al, 2014).

Prevalence of ASD in Albania was estimated at 1.31%. Current estimates in the U.S. propose a similar rate of 1.47%, while global estimates are much lower, suggesting that 0.62% of the world's population has ASD (Center For Disease Control, 2014; Elsabbagh et al., 2012). Caution is required when interpreting the estimated prevalence in Albania, since this was not intended as an epidemiological study.

Of those with follow-up, about 20% of toddlers in Albania, and 37% in the US (Robins et al, 2014), continued to screen positive, stressing the need for the follow-up to eliminate many false positive cases, which were higher in Albania. A drawback of the two-stage screening process is the resources required to conduct follow up interviews with caregivers, especially in low resource settings (e.g., money needed to call caregivers, time required by staff to administer items via phone and solicit examples, making multiple calls to families before contact made, etc.).

Attrition rates such as those found in the current study (50% of screened positive toddlers lost to follow up and 23% refused evaluation after continuing to screen positive on follow up interview) further contribute to the difficulties posed by the two stage process. Attrition due to caregiver inaccessibility (e.g., phone number changed, moved, etc.) can preclude at-risk children from receiving further services. Moreover, cultural differences in expectations regarding child development and in attitudes involving stigma of developmental disorders may also contribute to parent refusal to discuss concerns over the phone or to receive a developmental evaluation for their child (Koh et al., 2014; Samadi and McConkey, 2015). The problem of attrition across the two stage screening process is not confined to the current sample. In the U.S., 18% of children were lost to follow up after Stage 1 of screening and 37% either refused or did not complete evaluations following Stage 2 (Robins et al., 2014). Studies examining the two stage screening process with translated and/or culturally adapted versions of the original M-CHAT screener in Japan and South Korea also reported high rates of toddlers lost to follow up between initial screening and the follow up interview, with both studies reporting they were unable to follow up with 39% of their screened samples (Kamio, Inada, Koyama, Inokuchi, Tsuchiya, & Kuroda, 2014; Seung, et al., 2015). Kamio and colleagues (2014) proposed that the high false positive rate found in their study (17%) may have been a contributing factor to attrition in their sample. The authors suggested that Japanese caregivers may have experienced unnecessary distress as a result of their child being identified incorrectly as being at risk for ASD, which may have impacted participation rates in follow up (Kamio et al., 2015). The current study did not find any significant differences between toddlers who were lost to follow up or refused evaluations and those who completed the procedures.

Researchers continue to examine alternatives to established screening protocols that will best suit the needs of each population and context, including measures for decreasing attrition. One study of the M-CHAT's performance in Turkey addressed cultural differences found in caregiver response patterns there, by having the screener administered by a staff person verbally, and offering follow up interviews in-person at the time of screening (Kara et al., 2012). Currently, the original authors of the M-CHAT and M-CHAT-R/F in English are examining the utility of an electronic version of the M-CHAT-R screener, with follow up being administered immediately following the screener electronically. If found to be a valid

administration process of the M-CHAT-R/F in the U.S., dissemination and evaluation of this technology should be examined in other regions, particularly those where higher attrition rates are found. Additionally, modifications to the screening tool itself, as well as to cut off scores, have been shown to improve M-CHAT performance when adapted for use outside of the U.S. setting (Koh et al., 2014; Inada et al., 2011; Wong, Hui, Lee, Leung, Ho, Lau, ... & Chung, 2004; Kamio et al., 2015). Such modifications were also explored in the current study and will be discussed below.

Parents of each evaluated child reported high numbers of predictive concerns on the PEDS, indicating that developmental delays were found in both ASD and non-ASD toddlers who screened positive on the M-CHAT-R-A and follow-up. Findings suggest that rather than being delayed across all developmental skill areas, the ASD groups, on average, demonstrated patterns of delays consistent with their diagnoses (e.g., social/emotional functioning, communication, adaptive skills), with relative sparing of motor skills, consistent with other literature (Ventola, Kleinman, Pandey, Wilson, Esser, Boorstein, ... & Green, 2007).

Results of the DFA demonstrated that joint attention items on the M-CHAT-R-A best predict ASD classification (see Table 5). The item that was most predictive of ASD classification was Item 1, "If you point at something across the room, does your child look at it?" Response to name, social smiling, showing items, and following instructions were also highly predictive of outcome. Each of the reverse scored and foil items were found to be least predictive of outcome. This was by design with the foil items; the reverse scored items described autism symptoms rather than developmental milestones, and may be hard to understand for families who have not observed these behaviors; they were found to change frequently on follow-up interview. Pictures or video clips might help families make accurate judgments about their children. Unexpectedly, Item 6, "Does your child point with one finger to ask for something or to get help?" also poorly predicted ASD classification in the current sample. However, this result may be due to the shared variance with Item 1, which also inquires about pointing, a result also demonstrated by the DFA conducted for the original M-CHAT study (Robins et al., 2001).

Several internationally conducted M-CHAT studies supported current findings regarding predictive value of items on the initial screener. Items 1, 7, 8, 9, 10, 14, 16, 17, and 18 were found to perform well by multiple investigators across different region of the world, including Japan, China, Turkey, South Korea, and several Middle Eastern countries (Inada et al., 2011; Wong et al., 2004; Kara et al., 2012; Seung et al., 2015; Seif Edin et al., 2008). Each of these items assesses a child's capacity for joint attention and responsiveness to social stimuli, which have also been shown to best differentiate between children with ASD and other developmental concerns in previous literature (Ventola et al., 2007).

International studies examining the performance of the M-CHAT also reported findings related to problematic or less predictive items that were consistent with the current study. Three such studies found Items 12 and 5 (sensitivity to noise and finger movements near the eyes), the poorest performing items in the current study, to be less predictive in their samples as well (Canal-Bedia et al., 2011; Kara et al., 2012; Seung et al., 2015). Authors of a

validation study for the M-CHAT conducted in Spain proposed that higher failure rates on Items 12 and 5 (Items 11 and 18 on the original M-CHAT) were possibly due to parents' misunderstanding the question (Canal-Bedia et al., 2011). In South Korea, Seung and colleagues (2015) reported that parents' responses on M-CHAT items equivalent to M-CHAT-R/F Items 12 and 5 changed more frequently after follow up than others. These authors reported that the higher initial failure rates were potentially a result of misunderstanding items due to "culturally-based" differences revealed during follow up interviews (Seung et al., 2015). In the case of Item 5 involving "unusual finger movements," the authors suggest that South Korean parents may have had difficulty understanding the intent of the question due a common form of dyadic play involving pointing (Seung et al., 2015). Further, the South Korean study authors proposed a shorter version of the Korean M-CHAT excluding these two items, as well as the other two reverse-coded items found in the original M-CHAT (the M-CHAT-R/F has three such items), due to the high failure rates in their sample (Seung et al., 2015). A study conducted in Japan similarly proposed creating a short form of the Japanese M-CHAT, also excluding the four reverse scored items on the original screener, suggesting that fewer false positives would result (Inada et al., 2011).

Differences were also found between results of the current study and those reported in the literature. Items found to be predictive in the current study, such as social referencing and imitation were found to be similarly predictive in studies conducted in Japan and China, but less predictive in studies conducted in Spain, Turkey, and several Arabic speaking countries (Kara et al., 2012; Inada et al., 2011; Wong et al., 2004; Seif Edin et al., 2008). The converse was also true, with Items 3 (pretend play) and 11 (pointing to request) found to perform poorly in the current Albanian sample, while studies conducted in Japan found both items to be predictive and others in China and Turkey found Item 11 to perform well in their samples (Kara et al., 2012; Inada et al., 2011; Wong et al., 2004).

It is difficult to determine the precise causes for such differences in item performance around the world. However, as the field of ASD research continues to expand to include other regions of the globe, researchers have begun to explore potential differences in cultural perceptions regarding child development, completion of written screening tools, and of an ASD diagnosis itself. Researchers in this field have found that the salience of items examining social and communication skills differ between countries and cultures (Daley et al., 2013; Perera et al., 2009; Samadi and McConkey, 2015; Kara et al., 2012; Koh et al., 2014; Albores-Gallo, Roldán-Ceballos, Villarreal-Valdes, Betanzos-Cruz, Santos-Sánchez, Martínez-Jaime, ... & Hilton, 2012). Cultural attitudes toward the propriety of frequent eye contact or pointing between children and adults are suggested as one example of an effect on screener performance, as such behaviors in young children may be considered disrespectful (Albores-Gallio et al., 2012). Other investigators suggest that mothers in LMICs may not notice or interpret their child's delays in communication or social skills as impairments (Daley et al., 2013; Ertem, Atay, Dogan, Bayhan, Bingoler, Gok, ... & Isikli, 2007). The optimal format for collecting information about a child's development may also be culturally dependent. For example, caregiver familiarity with completing paper and pencil questionnaires or with responding to yes/no questions has been shown to influence screening results (Kara et al., 2012; Wong et al., 2004; Inada et al., 2011). Once these factors are better

understood in a particular context, modifications to the screener should then be made to improve its utility in that population.

Proposed Modifications to M-CHAT-R/F-A

Data analysis also reevaluated the M-CHAT-R/F-A, omitting the three most problematic items, Items 12, 5, and 2. While further research must be conducted to evaluate this shortened version of the screener as a stand-alone measure, analyses of data from the current sample indicate that the omission of these items would improve performance by increasing the internal consistency of the screening tool and decreasing initial screen positive rates, thereby reducing the number of follow up interviews required. Also based on this data, a total cut off score of 5 or greater was suggested.

Limitations

Limitations of the current study include inability to determine the number of true negative cases, due to limited resources for evaluating screen negative children before or after follow up. As a result, sensitivity could not be established. Previous studies have suggested that higher functioning children, with fewer or milder developmental delays, who are later diagnosed with ASD are less likely to screen positive when screened at 18 months of age than more severely affected children (Eaves, Winger, & Ho, 2006; Stenberg et al., 2014). It may be that higher functioning children who screened negative at this time may later meet criteria for ASD. This could only be determined with a population study to find autism cases in middle childhood. Since there were no differences in the number of items failed or other demographic factors between the continued and the attrition groups, it is not likely that there was systematic bias in attrition of a disproportionate number of children with ASD, but milder cases could certainly have been missed in both groups.

Due to the current sociopolitical context in Albania, collecting data on race and ethnicity was discouraged. The current investigators therefore cannot be certain that results would hold for ethnic minorities in Albania or Albanian-speaking populations in other countries. However, given that more than 99% of children attend WCC in Albania, the current results are probably representative of the Albanian speaking community in Albania (Institute of Statistics, Institute of Public Health [Albania] and ICF Macro, 2010). Results also indicated that the current screened sample reported higher maternal education levels than the general population in Albania. This may be due to parents with higher education being more willing to complete screeners, or to the fact that mothers of toddlers are younger than the population at large and therefore more educated, or to the fact that screening was done in urban areas.

In addition, although a large number of children were screened, only 19 were fully evaluated; making the conclusions about screening efficacy tentative. Finally, the suggested modifications of the screener need to be tested as a stand-alone measure.

Conclusions

The current study revealed strong performance for the use of the M-CHAT-R, Albanian version to detect ASD in toddlers who received follow up after screening positive during a WCC visit in Albanian speaking communities. Although there are no hard data on this point,

our Albanian medical colleagues report that most children outside of the study are identified as possibly having autism much later than the current screening age (or not at all). Children in the current study received timely referrals for treatment at the Regional Center for Autism. Results of this study support the conclusion that while ASD screening tools may require cultural adaptations to best meet the needs of different contexts, the core symptoms of ASD (e.g., following a point, bringing to show, shared affect, response to name, gaze following) appear to be consistent across cultures.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank the children, families, and pediatricians in Albania for their time and participation in this project. We are especially grateful to Dr. Liri Berisha, founder and head of the Albanian Children's Foundation for her initiating the autism work in Albania, and her unending support. To the therapists and other staff of the Albanian Children's Foundation and the Regional Center for Autism, we are very grateful for your dedication to the parents and children and your participation in many roles on the screening project. We are particularly thankful to Erjona Capari, Ariana Caushi, Miranda Hadermataj and Dr. Aida Bushati of Tirana Medical University for your incredible attention to detail, personal dedication to the children, and commitment to this study and to Dr. Bushati and Ms. Hadermataj for conducting the ADOS. We wish to thank Dr. Lynn Brennan, BCBA-D, and Jamie Hughes, BCBA, for their incredible dedication, hours beyond counting, and invaluable leadership on all parts of the Albanian screening and parent training projects. Thank you also to Dr. Marianne Barton for your wisdom and support throughout this project. We are deeply grateful to Autism Speaks, for funding Screening, Diagnosis and Parent Training for Young Children with ASD in Albania, of which this study was a part, and to Andy Shih of Autism Speaks for his indefatigable efforts to improve life for children and families affected by autism throughout the world.

This project was funded by a grant from Autism Speaks (Screening, Diagnosis, and Parent Training for Young Children with Autism in Albania). AC and IR received consulting fees from this grant.

References

- Albores-Gallo L, Roldán-Ceballos O, Villarreal-Valdes G, Betanzos-Cruz BX, Santos-Sánchez C, Martínez-Jaime MM, et al. Hilton CL. M-CHAT Mexican version validity and reliability and some cultural considerations. *ISRN neurology*. 2012; 2012
- American Academy of Pediatrics. AAP Statement on U.S. Preventive Services Task Force Final Recommendations on Autism Screening. 2016. Retrieved from: <https://www.aap.org/en-us/about-the-aap/Committees-Councils-Sections/Council-on-Children-with-Disabilities/Pages/Recent-Information.aspx?nfstatus=401&nftoken=00000000-0000-0000-0000-000000000000&nfstatedescription=ERROR%3a+No+local+token#sthash.ipMIDMIO.dpuf>
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. Arlington: American Psychiatric Publishing; 2013.
- Anderson DK, Liang JW, Lord C. Predicting young adult outcome among more and less cognitively able individuals with autism spectrum disorders. *Journal of Child Psychology and Psychiatry*. 2014; 55:485–494. DOI: 10.1111/jcpp.12178 [PubMed: 24313878]
- Ashwood KL, Buitelaar J, Murphy D, Spooren W, Charman T. European clinical network: autism spectrum disorder assessments and patient characterisation. *European child & adolescent psychiatry*. 2015; 24(8):985–995. [PubMed: 25471824]
- Autism Speaks Global Autism Public Health Initiative. (n.d.). Retrieved 5/1/2015 from: https://www.autismspeaks.org/science/research-initiatives/global-autism-public-health#gaph_goals
- Brothers KB, Glascoe FP, Robertshaw NS. PEDS: Developmental Milestones--An Accurate Brief Tool for Surveillance and Screening. *Clinical Pediatrics*. 2008; 47(3):271–279. [PubMed: 18057141]

- Canal-Bedia R, García-Primo P, Martín-Cilleros MV, Santos-Borbujo J, Guisuraga-Fernández Z, Herráez-García L, et al. Modified Checklist for Autism in Toddlers: Cross-Cultural Adaptation and Validation in Spain. *Journal of Autism and Developmental Disorders*. 2010; 41(10):1342–1351.
- Center for Disease Control. Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years —Autism and Developmental Disabilities Monitoring Network, 11 Sites. United States: 2014. 2010
- Chlebowski C, Robins DL, Barton ML, Fein D. Large-scale use of the modified checklist for autism in low-risk toddlers. *Pediatrics*. 2013; 131(4):e1121–e1127. [PubMed: 23530174]
- Coury DL. Babies, Bathwater, and Screening for Autism Spectrum Disorder: Comments on the USPSTF Recommendations for Autism Spectrum Disorder Screening. *Journal of Developmental & Behavioral Pediatrics*. 2015; 36(9):661–663. [PubMed: 26421531]
- Daley TC, Singhal N, Krishnamurthy V. Ethical considerations in conducting research on autism spectrum disorders in low and middle income countries. *Journal of autism and developmental disorders*. 2013; 43(9):2002–2014. [PubMed: 23283629]
- Elsabbagh M, Divan G, Koh YJ, Kim YS, Kauchali S, Marcín C, et al. Global Prevalence of Autism and Other Pervasive Developmental Disorders. *Autism Research*. 2012; 5(3):160–179. [PubMed: 22495912]
- Ertem IO, Atay G, Dogan DG, Bayhan A, Bingoler BE, Gok CG, et al. Isikli S. Mothers' knowledge of young child development in a developing country. *Child: care, health and development*. 2007; 33(6):728–737.
- Eaves LC, Wingert H, Ho HH. Screening for autism Agreement with diagnosis. *Autism*. 2006; 10(3): 229–242. [PubMed: 16682396]
- Gavi. Health equity. n.dRetrieved from: <http://www.gavi.org/About/Value/Health-equity/#sthash.EaG1n7oc.dpuf>
- Glascoe, FP. Parents' evaluations of developmental status: A method for detecting and addressing developmental and behavioral problems in children. Nashville, TN: Ellsworth & Vandermeer Press Ltd; 1997.
- Glascoe FP. Parents' Evaluation of Developmental Status: How Well Do Parents' Concerns Identify Children With Behavioral and Emotional Problems? *Clinical Pediatrics*. 2003; 42(2):133–138. [PubMed: 12659386]
- Glascoe, FP.; Robertshaw, NS. Parents' Evaluation of Developmental Status: Developmental Milestones (PEDS:DM). Nashville, Tennessee: Ellsworth & Vandermeer Press, LLC; 2008.
- Herlihy LE. Racial/ethnic and Socioeconomic Differences in Screening Toddlers for Autism Spectrum Disorders Using the M-CHAT. 2014
- Institute of Statistics, Institute of Public Health [Albania] and ICF Macro. Albania Demographic and Health Survey 2008-09. Tirana, Albania: Institute of Statistics, Institute of Public Health and ICF Macro; 2010.
- Inada N, Tomonori Koyama P, Inokuchi E, Kuroda M, Kamio Y. Research in Autism Spectrum Disorders. *Research in Autism Spectrum Disorders*. 2010; 5(1):330–336.
- Johnson CP, Myers SM. American Academy of Pediatrics Council on Children With Disabilities. Identification and evaluation of children with autism spectrum disorders. *Pediatrics*. 2007; 120(5): 1183–1215. [PubMed: 17967920]
- Kamio Y, Haraguchi H, Stickley A, Ogino K, Ishitobi M, Takahashi H. Brief Report: Best Discriminators for Identifying Children with Autism Spectrum Disorder at an 18-Month Health Check-Up in Japan. *Journal of autism and developmental disorders*. 2015; 1–7. [PubMed: 25428291]
- Kamio Y, Haraguchi H, Stickley A, Ogino K, Ishitobi M, Takahashi H. Brief Report: Best Discriminators for Identifying Children with Autism Spectrum Disorder at an 18-Month Health Check-Up in Japan. *Journal of autism and developmental disorders*. 2015; 45(12):4147–4153. [PubMed: 26189180]
- Kara B, Mukaddes NM, Altinkaya I, Guntepe D, Gokcay G, Ozmen M. Using the Modified Checklist for Autism in Toddlers in a well-child clinic in Turkey: Adapting the screening method based on culture and setting. *Autism*. 2014; 18(3):331–338. [PubMed: 23175752]

- Kleinman JM, Ventola PE, Pandey J, Verbalis AD, Barton M, Hodgson S, et al. Fein D. Diagnostic stability in very young children with autism spectrum disorders. *Journal of autism and developmental disorders*. 2008; 38(4):606–615. [PubMed: 17924183]
- Koh HC, Lim SH, Chan GJ, Lin MB, Lim HH, Choo SHT, Magiati I. The Clinical Utility of the Modified Checklist for Autism in Toddlers with High Risk 18–48 Month Old Children in Singapore. *Journal of Autism and Developmental Disorders*. 2014; 44(2):405–416. [PubMed: 23812662]
- Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, DiLavore PC, et al. The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*. 2000; 30(3)
- Mandell DS, Listerud J, Levy SE, Pinto-Martin JA. Race differences in the age at diagnosis among Medicaid-eligible children with autism. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2002; 41(12):1447–1453. [PubMed: 12447031]
- Matson JL, Worley JA, Fodstad JC, Chung KM, Suh D, Jhin HK, et al. Furniss F. A multinational study examining the cross cultural differences in reported symptoms of autism spectrum disorders: Israel, South Korea, the United Kingdom, and the United States of America. *Research in Autism Spectrum Disorders*. 2011; 5(4):1598–1604.
- McPheeters ML, Weitlauf A, Vehorn A, Taylor C, Sathe NA, Krishnaswami S, et al. Warren ZE. Screening for Autism Spectrum Disorder in Young Children. 2016
- National Research Council (US) Committee on Educational Interventions for Children with Autism. *Educating Children with Autism*. Washington, DC: National Academies Press; 2001.
- Patel V, Kieling C, Maulik PK, Divan G. Improving access to care for children with mental disorders: a global perspective. *Archives of disease in childhood*. 2013; 98(5):323–327. [PubMed: 23476001]
- Patel V, Chowdhary N, Rahman A, Verdeli H. Improving access to psychological treatments: Lessons from developing countries. *Behaviour Research and Therapy*. 2011; 49(9):523–528. [PubMed: 21788012]
- Patel V, Flisher AJ, Nikapota A, Malhotra S. Promoting child and adolescent mental health in low and middle income countries. *Journal of Child Psychology and Psychiatry*. 2008; 49(3):313–334. [PubMed: 18093112]
- Perera H, Wijewardena K, Aluthwelage R. Screening of 18-24-Month-Old Children for Autism in a Semi-Urban Community in Sri Lanka. *Journal of Tropical Pediatrics*. 2009; 55(6):402–405. [PubMed: 19401407]
- Robins DL, Casagrande K, Barton M, Chen CMA, Dumont-Mathieu T, Fein D. Validation of the Modified Checklist for Autism in Toddlers, Revised With Follow-up (M-CHAT-R/F). *Pediatrics*. 2014; 133(1):37–45. [PubMed: 24366990]
- 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 & Behavioral Pediatrics*. 2006; 27(2):S111–S119. [PubMed: 16685177]
- 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):131–143. [PubMed: 11450812]
- Robins DL, Adamson LB, Barton M, Connell JE Jr, Dumont-Mathieu T, Dworkin PH, et al. Newschaffer C. Universal Autism Screening for Toddlers: Recommendations at Odds. *Journal of autism and developmental disorders*. 2016; 46(5):1880–1882. [PubMed: 26820633]
- Rogers SJ. Brief report: Early intervention in autism. *Journal of autism and developmental disorders*. 1996; 26(2):243–246. [PubMed: 8744493]
- Samadi SA, McConkey R. Screening for Autism in Iranian Preschoolers: Contrasting M-CHAT and a Scale Developed in Iran. *Journal of autism and developmental disorders*. 2015; 45(9):2908–2916. [PubMed: 25911978]
- Saxena S, Thornicroft G, Knapp M, Whiteford H. Resources for mental health: scarcity, inequity, and inefficiency. *The Lancet*. 2007; 370(9590):878–889.

- Seif Eldin A, Habib D, Noufal A, Farrag S, Bazaid K, Al-Sharbati M, et al. Use of M-CHAT for a multinational screening of young children with autism in the Arab countries. *International Review of Psychiatry*. 2008; 20(3):281–289. [PubMed: 18569180]
- Seung H, Ji J, Kim SJ, Sung I, Youn YA, Hong G, et al. Youm HK. Examination of the Korean Modified Checklist of Autism in Toddlers: Item Response Theory. *Journal of autism and developmental disorders*. 2015; 45(9):2744–2757. [PubMed: 25847755]
- Simmons and Slocum. Indo-European Languages, Balkan Group. Albania: 2014. Retrieved from: <http://www.utexas.edu/cola/centers/lrc/general/ie-lg/Albanian.html>
- Soto S, Linas K, Jacobstein D, Biel M, Migdal T, Anthony BJ. A review of cultural adaptations of screening tools for autism spectrum disorders. *Autism*. 2015; 19(6):646–661. [PubMed: 25008216]
- Stenberg N, Bresnahan M, Gunnes N, Hirtz D, Hornig M, Lie KK, et al. Schjølberg S. Identifying children with autism spectrum disorder at 18 months in a general population sample. *Paediatric and perinatal epidemiology*. 2014; 28(3):255–262. [PubMed: 24547686]
- Szatmari P, Bryson SE, Boyle MH, et al. Predictors of outcome among high functioning children with autism and Asperger syndrome. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. 2003; 44(4):520–528.
- UNICEF. Albania, statistics. 2013. Retrieved from: http://www.unicef.org/infobycountry/albania_statistics.html
- USAID. Albania: 2011-2015, Country Development Cooperation Strategy. 2011. Retrieved from: https://www.usaid.gov/sites/default/files/documents/1863/CDCS_USAID_Albania.pdf
- Ventola P, Kleinman J, Pandey J, Wilson L, Esser E, Boorstein H, et al. Green J. Differentiating between autism spectrum disorders and other developmental disabilities in children who failed a screening instrument for ASD. *Journal of autism and developmental disorders*. 2007; 37(3):425–436. [PubMed: 16897377]
- Wallace S, Fein D, Rosanoff M, Dawson G, Hossain S, Brennan L, et al. A Global Public Health Strategy for Autism Spectrum Disorders. *Autism Research*. 2012; 5(3):211–217. [PubMed: 22605577]
- Whiteford HA, Degenhardt L, Rehm J, Baxter AJ, Ferrari AJ, Erskine HE, et al. Burstein R. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *The Lancet*. 2013; 382(9904):1575–1586.
- The World Bank. Newest country classifications. 2012. Retrieved from: <http://data.worldbank.org/news/newest-country-classifications-released>
- The World Bank. Newest country classifications. 2013. Retrieved from: <http://data.worldbank.org/news/new-country-classifications>
- The World Bank. [accessed 15 May 2014] How We Classify Countries. 2015. Retrieved from: <http://data.worldbank.org/about/country-classifications>
- WHO. mhGAP intervention guide for mental, neurological and substance use disorders in non-specialized health settings: mental health Gap Action Programme (mhGAP). 2010
- WHO. Metrics: Disability-Adjusted Life Year (DALY). n.d Retrieved from: http://www.who.int/healthinfo/global_burden_disease/metrics_daly/en/
- WHO. Meeting report: autism spectrum disorders and other developmental disorders: from raising awareness to building capacity. 2013. Retrieved from: http://apps.who.int/iris/bitstream/10665/103312/1/9789241506618_eng.pdf
- WHO. Albania. 2015. Retrieved from: <http://www.who.int/countries/alb/en/>
- Wong V, Hui LHS, Lee WC, Leung LSJ, Ho PKP, Lau WLC, et al. Chung B. A modified screening tool for autism (Checklist for Autism in Toddlers [CHAT-23]) for Chinese children. *Pediatrics*. 2004; 114(2):e166–e176. [PubMed: 15286253]
- Zwaigenbaum L, Bauman ML, Fein D, Pierce K, Buie T, Davis PA, et al. Kasari C. Early screening of autism spectrum disorder: recommendations for practice and research. *Pediatrics*. 2015; 136(Supplement 1):S41–S59. [PubMed: 26430169]
- Zweigenbaum L, Bryson S, Brian J, Smith I, Roberts W, Szatmari P, Roncadin C, Garon N, Vaillancourt T. Stability of diagnostic assessment for autism spectrum disorder between 18 and 36 months in a high-risk cohort. *Autism Research*. 2015; doi: 10.1002/aur.1585

Zweigenbaum L, Bryson S, Lord C, Rogers S, Carter A, Carver L, et al. Yirmiya N. Clinical assessment and management of toddlers with suspected autism spectrum disorder: insights from studies of high-risk infants. *Pediatrics*. 2009; 123(5):1383–1391. [PubMed: 19403506]

Figure 1. Flow Chart

Table 1
Sample Demographic Characteristics

Variable	Total
Total Sample	2594
Age at screening, mean (SD), mo	24.16 (2.83)
<i>Sex</i>	
Female	1300
Male	1288
Not reported	6
<i>Maternal Education</i>	
No formal education/incomplete primary school	32
Unfinished secondary school	610
Secondary education	683
Some college	3
College/University Degree	781
Advanced Degree	26
Unknown	459
<i>Person who Completed M-CHAT-R-A</i>	
Mother	2290
Father	211
Other Family Member	42
Unknown	51

Table 2
Total Scores for M-CHAT-R-A and M-CHAT-R/F-A by Screening Outcome

M-CHAT-R-A Total Scores		Mean (SD)
Total sample, <i>n</i> =2594		1.18 (1.96)
Screened Negative on M-CHAT-R-A, <i>n</i> =2341		0.74 (.73)
Screened Positive on M-CHAT-R-A, <i>n</i> =253		5.32 (3.95)
No Evaluation Needed, M-CHAT-R-A, <i>n</i> =101		4.05 (2.29)
Qualified for Evaluation, M-CHAT-R-A, <i>n</i> =26		11.15 (4.63)
Refused Evaluation, <i>n</i> =6		10.83 (4.26)
M-CHAT-R/F-A Total Scores		Mean (SD)
Completed follow-up, <i>n</i> =127		1.72 (3.72)
No Evaluation Needed, M-CHAT-R/F-A, <i>n</i> =101		0.05 (.22)
Qualified for Evaluation M-CHAT-R/F-A, <i>n</i> =26		8.23 (3.77)
Refused Evaluation, <i>n</i> =6		6.33 (4.32)
Evaluated Sample Total Scores	M-CHAT-R-A Mean (SD)	M-CHAT-R/F-A Mean (SD)
ASD Group (ASD+Autism, <i>n</i> =17)	11.71 (4.07)	9.18 (3.59)
ASD, <i>n</i> =3	13.33 (6.03)	6.33 (3.51)
Autism, <i>n</i> =14	11.36 (3.75)	9.79 (3.42)
Non-ASD, <i>n</i> =2	9.50 (2.12)	6.00 (2.83)

Table 3
Demographic Characteristics of Children Evaluated

Variable	Total
Total Sample	20
<i>Sex</i>	
Male	15
Female	5
<i>Maternal Education</i>	
No formal education/incomplete primary school	0
Unfinished secondary school	6
Secondary education	7
Some college	0
College/University Degree	2
Advanced Degree	0
Unknown	5
<i>Person who Completed M-CHAT-R-A</i>	
Mother	20
Age at screening, mean (SD), mo	26.05 (5.30)
M-CHAT-R-A total score, mean (SD), <i>n</i> =20	11.25 (4.83)
M-CHAT-R/F-A total score, mean (SD), <i>n</i> =20	8.80 (3.50)

Table 4

Response Stability from M-CHAT-R-A to M-CHAT-R/F-A

Item #	M-CHAT-R-A (Screener)				% Changed after Follow up (Passed)	% Changed after Follow up	Missing	Total	M-CHAT-R/F-A (Follow up)			
	Passed	Failed	Changed after Follow up	Failed					Pass	Fail	#Pass of Missing	# Fail of Missing
1	103	23	10				1	127	113	14	1	0
2 ^{**}	82	43	42		43%		2	127	126	1	2	0
3	78	44	28		64%		5	127	109	18	5	0
4	109	17	13		76%		1	127	123	4	1	0
5 ^{**}	62	61	54		89%		4	127	119	8	4	0
6	84	11	5		45%		32	127	114	13	31	1
7	86	37	19		51%		4	127	107	20	4	0
8	99	26	11		42%		2	127	109	18	2	0
9	87	36	18		50%		4	127	109	18	4	0
10	105	18	11		61%		4	127	119	8	4	0
11	110	15	8		53%		2	127	119	8	2	0
12 ^{**}	40	85	83		98%		2	127	124	3	2	0
13	120	4	2		50%		3	127	125	2	3	0
14	106	18	9		50%		3	127	117	10	2	1
15	98	23	16		70%		6	127	117	10	5	1
16	88	35	19		54%		4	127	109	18	3	1
17	84	34	18		53%		9	127	109	18	8	1
18	96	28	17		61%		3	127	114	13	2	1
19	90	33	23		70%		4	127	115	12	3	1
20	114	11	8		73%		2	127	124	3	2	0

^{**} Reverse scored items.

Table 5
DFA Results: M-CHAT-R-A Items in Function 1

Item #	Question	Function 1
1	If you point at something across the room, does your child look at it?	0.629
10	Does your child respond when you call his or her name?	0.525
11	When you smile at your child, does he or she smile back at you?	0.491
9	Does your child show you things by bringing them to you or holding them up for you to see – not to get help, but just to share?	0.472
18	Does your child understand when you tell him or her to do something?	0.448
16	If you turn your head to look at something, does your child look around to see what you are looking at?	0.427
17	Does your child try to get you to watch him or her?	0.413
19	If something new happens, does your child look at your face to see how you feel about it?	0.400
14	Does your child look you in the eye when you are talking to him or her, playing with him or her, or dressing him or her?	0.385
15	Does your child try to copy what you do?	0.360
8	Is your child interested in other children?	0.346
7	Does your child point with one finger to show you something interesting?	0.337
DFA Results: M-CHAT-R-A Items Not Included in Function		
Item #	Question	Function 1
3	Does your child play pretend or make-believe?	0.222
2	Have you ever wondered if your child was deaf? *	0.150
4	Does your child like climbing on things? **	0.136
20	Does your child like movement activities? **	0.108
5	Does your child make unusual finger movements near his or her eyes? *	0.086
6	Does your child point with one finger to ask for something or to get help?	0.084
13	Does your child walk? **	0.037
12	Does your child get upset by everyday noises? *	-0.035

* Indicates Reverse scored item, e.g., “Yes” response indicates ASD symptom presence

** Indicates foil items

Table 6
M-CHAT-R/F-A with Three Items Removed, a Comparison with Original

Screening Category	Original M-CHAT-R-A		M-CHAT-R-A Without Three Reverse Scored Items	
	<i>n</i>	Mean (SD)	<i>n</i>	Mean (SD)
M-CHAT-R-A	2594	1.18 (1.96)	2594	0.49 (1.68)
Screened Negative on M-CHAT-R-A	2341 (90%)	0.74 (0.73)	2487 (96%)	0.21 (0.47)
Screened Positive on M-CHAT-R-A	253	5.32 (3.95)	107	7.07 (4.29)
No Evaluation Needed, M-CHAT-R-A	101	4.05 (2.29)	35	4.91 (3.01)
Qualified for Evaluation, M-CHAT-R-A	26	11.15 (4.63)	24	10.42 (3.91)
Lost to Follow up/Unable to Contact	126	5.13 (3.82)	48	6.96 (4.28)
Refused Evaluation	6	10.83 (4.26)	4	10.75 (6.08)