J Autism Dev Disord
DOI 10.1007/s10803-016-2876-4

ORIGINAL PAPER

Two-Step Screening of the Modified Checklist for Autism in Toddlers in Thai Children with Language Delay and Typically Developing Children

Pornchada Srisinghasongkram[1,2] · **Chandhita Pruksananonda**[2,3] · **Weerasak Chonchaiya**[2,3]

© Springer Science+Business Media New York 2016

Abstract This study aimed to validate the use of two-step Modified Checklist for Autism in Toddlers (M-CHAT) screening adapted for a Thai population. Our participants included both high-risk children with language delay (N = 109) and low-risk children with typical development (N = 732). Compared with the critical scoring criteria, the total scoring method (failing ≥3 items) yielded the highest sensitivity of 90.7 %; specificity was 99.7 %, positive predictive value 96.1 %, and negative predictive value 99.4 %. The two-step M-CHAT screening is a promising instrument that can be utilized to detect ASD in Thai children in both primary and clinical settings. Moreover, socio-cultural context should be considered when adopting the use and interpretation of the M-CHAT for each country.

Keywords Autism · Delayed language · M-CHAT · Screening · Speech

Introduction

Autism spectrum disorder (ASD) is a common neurodevelopmental disorder with a prevalence of approximately 1–3 % worldwide (Christensen et al. 2016; Kamio et al. 2014; Wingate et al. 2014; Yang et al. 2015). Early identification of individuals with ASD during the early developmental period is extremely important since it leads to proper management and better outcomes for affected children and their families (Warren et al. 2011; Zwaigenbaum et al. 2015). As a result, the American Academy of Pediatrics (AAP) recommends routine specific screening for ASD in all children aged 18 and 24 months at health supervision visits by using ASD-specific screening instruments, including the Modified Checklist for Autism in Toddlers (M-CHAT).

The M-CHAT is one of the most promising parent-completed screening tools commonly used in primary care and clinical settings (Johnson et al. 2007; Robins 2008; Robins et al. 2001; Zwaigenbaum et al. 2015). Furthermore, the M-CHAT can be utilized to screen for ASD in children from 16 months of age and up to 4 years old, which provides desirable sensitivity (0.70–0.97) and a wide range of specificity (0.38–0.99) depending on the settings (Charman et al. 2016; Kleinman et al. 2008; Pandey et al. 2008; Robins 2008; Robins et al. 2001; Snow and Lecavalier 2008; Yama et al. 2012; Zwaigenbaum et al. 2015). Nonetheless, positive predictive value (PPV) of the M-CHAT was reported to be very low to relatively low (0.058–0.43) for low-risk samples and those who were screened at a younger age. The PPV was higher in high-risk individuals and could be improved up to 0.57–0.76 with a follow-up interview (Pandey et al. 2008; Robins 2008; Robins et al. 2001; Zwaigenbaum et al. 2015).

As such, a follow-up interview should be considered in M-CHAT screening in primary care settings. This adjunct process could be more helpful to determine specific risks for ASD and the need for referral in individuals with developmental delay by minimizing the false-positive results (Chlebowski, Robins, Barton, and Fein 2013;

✉ Weerasak Chonchaiya
weerasak.ch@chula.ac.th

[1] Department of Pediatrics, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

[2] King Chulalongkorn Memorial Hospital, The Thai Red Cross Society, Rama IV. Road, Bangkok, Thailand

[3] Division of Growth and Development, Department of Pediatrics, Faculty of Medicine, Chulalongkorn University, Sor Kor Building 11th Floor, Bangkok 10330, Thailand

Published online: 26 July 2016

Kleinman et al. 2008; Robins 2008; Seung et al. 2015; Zwaigenbaum et al. 2015). Moreover, there is a more recent version, the M-CHAT-R/F that appears to better identify ASD at a higher rate compared to the original M-CHAT questionnaire (Robins et al. 2014).

The M-CHAT screening for ASD in both low-risk and high-risk groups of children has also been adopted in many developed countries (Garcia-Primo et al. 2014; Kamio et al. 2014; Zwaigenbaum et al. 2015). However, early detection of ASD in young individuals in developing countries including Thailand can be very challenging since the M-CHAT screening is not widely feasible in the context of limited funding and a lack of human resources, especially staff who are familiar with individuals with ASD. Time constraints and busy clinical settings also impede the M-CHAT screening. This is the case even in individuals who are at risk for ASD, for instance children whose parents or primary physicians have concerns of ASD, siblings of those with ASD, and children with delayed language development.

In addition to these constraints, the clinical utility of the M-CHAT and the best discriminating items for identifying ASD among various cultural backgrounds might differ from research on ASD documented by Robins et al. (2001). Kamio et al. (2015) recently reported six critical items including imitation of action, bringing objects to show, point following, protoimperative pointing, pretend play, and language comprehension that appeared to be most discriminative for 18-month-old Japanese toddlers with ASD from typically developing children. Most of these reported items were clearly different from the critical scores demonstrated in the original M-CHAT research by Robins et al. (2001). Therefore, socio-cultural difference should be taken into consideration when adopting the use and interpretation of the M-CHAT for each country.

In Thailand, there is still no routine screening for ASD in either low-risk or high-risk children. Data gathering on the M-CHAT based solely on parent-completed questionnaires, particularly in low-risk individuals at health supervision visits, may be less reliable as some questions may be misunderstood by parents. In addition to the limitations on resources and time constraints mentioned previously, parents in Thailand are more familiar with interviews or being read the process rather than the careful completion of questionnaires by themselves, as documented in other developing countries like Turkey (Kara et al. 2014).

Moreover, the M-CHAT completion rate was reported to be relatively low at 56.3 % even in high-risk individuals in developed countries like Singapore (Koh et al. 2014). In South Korea, the percentage of changed response for each M-CHAT item from an initial parental report was also noted to be significantly high, up to 65.7 %, after the follow-up interview was performed by a clinician (Seung et al. 2015). As a result, we were likely to encounter difficult issues related to the M-CHAT screening if responses on the M-CHAT were completely dependent on parents without the follow-up interview. Additionally, parents who are less concerned whether their child might have ASD usually do not completely understand all M-CHAT items that are specific and relevant to development and behaviors of their children. Therefore, they tend to provide more false-positive results, particularly if the follow-up interview is not in place (Chlebowski et al. 2013; Robins 2008; Seung et al. 2015; Zwaigenbaum et al. 2015).

To increase the suitability of the M-CHAT in a Thai cultural context, we hypothesized that a screening process that includes both a parent-completed questionnaire followed by a semi-structured interview by trained clinicians, only for cases that initially screen positive based on the total or the critical scoring criteria, could improve overall sensitivity, specificity, PPV, and negative predictive value (NPV) on the M-CHAT. As such, our study aimed at validating the use of two-step M-CHAT screening with a combination of initial parent reporting and interview. A trained clinician conducted a follow-up interview for response items that demonstrated a risk for ASD only for cases that initially screen positive to increase the clinical utility of the M-CHAT in identifying ASD in 18–48 month-old children who were at high- and low-risk for ASD. High-risk individuals were those who presented with language delay, whereas low-risk individuals were children with typical development, who came for routine child health supervision visits or attended events for child health promotion at our center. Furthermore, we compared sensitivity, specificity, PPV, and NPV between the total and the critical scoring methods of the two-step M-CHAT screening with respect to Thai children.

Methods

Participants and Procedure

From July 2014 to September 2015, there were a total of 841 participants who were enrolled in this study. The participants were divided into two groups; the high-risk and the low-risk groups for ASD as demonstrated in the following section.

High-Risk Group

Eighteen to 48 month-old children with language delay were initially identified by their caregivers who had concerns about their children's language development. Language delayed children in the high-risk group tended to fail

to acquire at least one of the following language milestones: (1) no other words beyond mama and dada by 15 months of age, (2) not following any commands by 18 months of age, (3) no use of phrases by 24 months of age, and (4) no three-word sentences by 36 months of age (Feldman and Messick 2009). These children were primarily seen by our pediatric residents. All participants were then referred for comprehensive evaluation by developmental and behavioral pediatricians.

At a developmental and behavioral evaluation visit, a parent of each individual was thoroughly interviewed about his or her primary concern regarding the child's language development. The interview also included a comprehensive history of all developmental streams and the age at which each developmental milestone was achieved; pre-, peri-, and postnatal risk factors, in particular prenatal complications, maternal illness and drug use during pregnancy, gestational age, birth weight, birth asphyxia (APGAR score of <7 at 5 min) in addition to postnatal complications after birth. We also obtained a family history of any developmental or behavioral problems, including ASD, developmental language disorder, global developmental delay, and attention-deficit/hyperactivity disorder (ADHD) that had to be seen and treated by physicians of such individuals in the child's family (physician-documented developmental and behavioral disorders).

We also screened the medical history of the child and ascertained their main caregiver. Those who were born to mothers with prenatal complications, particularly maternal illnesses or infections during pregnancy, gestational diabetes, pregnancy induced hypertension, and antepartum hemorrhage; a preterm birth; having a birth weight of <2500 g or postnatal complications including birth asphyxia and postnatal conditions that required prolonged hospitalization than usual after birth; having chronic medical illness including epilepsy, congenital heart disease, chronic lung disease, chronic kidney disease, etc.; having congenital anomalies or syndromes; and had previously been diagnosed with ASD before receiving an evaluation at our center, were excluded from this study. As such, there were a total of 109 children aged 18–48 months who were enrolled in the high-risk group in this present study.

The M-CHAT, Thai version, was given to the mother of each individual in the high-risk group on the date of appointment for comprehensive developmental and behavioral evaluation. The mother completed the M-CHAT questionnaire in the waiting area before her child was seen by the developmental and behavioral pediatrician. After finishing the extensive interview mentioned above, the mother was then interviewed again regarding her M-CHAT response by a trained developmental and behavioral pediatrician only if her child screened positive based on the total or the critical scoring criteria. The follow-up interview was important since this process could help mothers understand the exact purpose of each M-CHAT item. For example, mothers answered item 18, "Does your child make unusual finger movements near his/her face?" with a "yes" response frequently. However, upon follow-up interview, the authors determined that Thai mothers intended to indicate that yes, their child wanted to play peek-a-boo or play with his/her fingers. The same pattern was reported in South Korea and Spain, where high rates of endorsement of item 18 were reported in children without ASD, particularly if the parent was not interviewed (Canal-Bedia et al. 2011; Seung et al. 2015).

As a result, we used a semi-structured interview approach for the follow-up interview in this present study. The follow-up interview was performed in person only for items that indicated the risk for ASD to clarify whether the mother understood that particular item on the M-CHAT correctly and to elicit specific examples of the child's behaviors with respect to the validity of the M-CHAT documented in pioneering research by Robins et al. (2001). All items which demonstrated the risk for ASD were read to the mother. Moreover, some items were also demonstrated as actions of such behaviors to aid the mother's understanding, including bouncing on the knee (item1), protoimperative pointing (item 6), protodeclarative pointing (item 7), making eye contact (item 10), plugging ears due to seemingly oversensitivity to noises (item 11), following a point (item 15), following eye gaze (item 17), finger flicking close to the eyes (item 18), and staring episode (item 22). Most mothers generally provided responses immediately after the follow-up interview and demonstration process mentioned above were conducted. However, there were questions that some mothers might still find confusing. As such, those specific items were generally asked as open-ended questions and a range of examples were provided, as displayed in Table 1.

After we completed the second step semi-structured interview, the mother's response was re-scored to determine whether the individual screened positive or negative based on the total or the critical scoring criteria.

On completion of the two-step M-CHAT screening, the clinician performed a physical examination. Each participant's development and behaviors were then evaluated by the developmental and behavioral pediatrician using either the Mullen Scales of Early Learning, the Cognitive Adaptive Test/Clinical Linguistic and Auditory Milestone Scale (CAT/CLAMS) or Denver II (Frankenburg et al. 1992; Hoon et al. 1993; Mullen 1995), depending on the expertise of the particular developmental and behavioral pediatrician and time constraints during a 60-to-90-min period for each new case.

Table 1 The examples of specific M-CHAT items asked as open-ended questions

M-CHAT item	The example of those items
5.	"Does your child have make-believe play? If so, please give an example." If the mother did not know about make-believe play, a few examples including pretending to use a remote control as a telephone or feeding self with a toy spoon would be provided. If the response indicated that the child had ever engaged in make-believe play, the researcher rated this item as a pass, otherwise they rated it as a fail
8.	"What is your child's most favorite toy and how does he/she play with it?" If the response indicated that the child played with his/her favorite toy properly without mouthing or dropping or just carrying it, the researcher rated the item as a pass, otherwise they rated it as a fail
11.	"How does your child react to noise that he/she does not like such as a vacuum cleaner or loud music or doors slamming?" If the response indicated that the child ever seemed oversensitive to such noises by covering his/her ears or screaming, the researcher rated this item as a fail, otherwise they rated it as a pass
13.	"What does your child do when you wave goodbye, clap your hands or blow a kiss?" If the response demonstrated that the child could imitate such actions, the researcher rated this item as a pass, otherwise they rated it as a fail
23.	"How often does your child look at your face to check your reaction when faced with something unfamiliar such as someone new approaching or a slightly scary circumstance?" If the response indicated that the child sometimes looked at the parental reaction when faced with such circumstances, the researcher rated this item as a pass, otherwise they rated it as a fail

Low-Risk Group

Screening for ASD in toddlers and preschool-aged children with typical development by using the M-CHAT is not mandatory during well-child care clinics and health supervision visits in Thailand. Therefore, we decided to invite all 18–48 month-old children who came for child health supervision visits or attended events that were held quarterly to twice a year for child health promotion at our center to participate in this present study. There were 472 participants initially recruited in the low-risk group. These participants' caregivers, mainly mothers, completed the intake questionnaire to verify that their child did not meet the exclusion criteria as previously documented in the high-risk group. To qualify as having typical development, participants' caregivers and primary pediatricians should never have been concerned about the child's development. Each child also had to pass developmental surveillance during health supervision visits.

Considering the exclusion criteria, three participants with underlying epilepsy, one with VACTERL association, two with age more than 48 months, two with preterm birth, and four with incomplete M-CHAT data were finally excluded from this study. As a result, our participants in this particular group were considered very low risk for ASD since the participants' development had never been considered a concern by their caregivers and primary pediatricians. Our participants were therefore not representative of population-based screening.

After meeting the inclusion criteria, each participant's mother initially completed the M-CHAT questionnaire by herself. The mother was then re-interviewed only if her child screened positive as per the two-step screening, mainly by the first author or developmental and behavioral pediatricians. To help ensure the clarity and reliability of provided data, this follow-up interview was done either immediately in person or within a 2-month-period via a telephone interview, depending upon the workload of the researchers. The re-interview generally lasted for 2–10 min and was focused on the specific responses that were at risk for ASD as documented in the high-risk group mentioned previously. In addition, 272 participants, who were originally enrolled as typically developing infants at 6 months of age in another separate study of a longitudinal cohort of the impact of electronic media exposure on young children, were screened for ASD using the M-CHAT at 18 and 24 months of age. As such, there were a total of 732 participants enrolled in the low-risk group.

Those who actually screened positive on the M-CHAT following both initial parent report and follow-up interview process on either the total or the critical scoring method were comprehensively evaluated by developmental and behavioral pediatricians in the same procedure as documented in the high-risk group. Participants who screened negative on the M-CHAT were also followed-up to elucidate whether they still had not been diagnosed with ASD. The follow-up method for those who screened negative was one of the following: (1) appointed to be evaluated by developmental and behavioral pediatricians; (2) had their medical records reviewed; or (3) a telephone follow-up.

A developmental evaluation was performed by a developmental and behavioral pediatrician by using the Mullen Scales of Early Learning (Mullen 1995) on the same date as the completed two-step M-CHAT screening in each individual who was previously enrolled in our longitudinal cohort described above. The telephone

interview for those who screened negative on the two-step M-CHAT included (1) "Do you have any concerns about your child's development and behaviors particularly social and language development in addition to restricted, repetitive, and stereotyped behaviors or interests that make you concerned that your child might be autistic? Restricted, repetitive, and stereotyped behaviors or interests included lining up toys, echolalia, extreme distress to small changes, and fixed interest to unusual objects."; (2) "Has your child's primary pediatrician ever diagnosed him/her with ASD?"; (3) "Does your child attend daycare, nursery or kindergarten? If so, have any care providers or teachers of your child ever been concerned whether he or she might have ASD?" The medical records review and telephone follow-up interview for those who screened negative on the M-CHAT generally occurred within a 6-to-12-month period after the two-step M-CHAT screening completion.

Materials

M-CHAT, Thai Version

We used the original version of the M-CHAT in this study. The M-CHAT, Thai version, was back translated and finally developed with kind permission from Dr. Robins DL, based on the principles of good practice for adaptation of the M-CHAT under her recommendations to retain its originality and integrity (Robins et al. 2001). This original questionnaire has 23 yes/no questions which can generally be completed by a parent within 5 min and scored by a trained clinician within 2 min. To be documented as failing M-CHAT screening with regard to the original M-CHAT research, such individuals have to fail three or more items (total scoring) or two or more critical items on the M-CHAT, which are items number 2, 7, 9, 13, 14, and 15 respectively (critical scoring).

The Diagnosis of ASD and Other Neurodevelopmental Disorders

In addition to a thorough interview, physical examination, and developmental evaluation using developmental tests mentioned previously, participants' development and behaviors were comprehensively observed by experienced developmental and behavioral pediatricians. These behaviors included joint attention both initiated and responded to by the child, affective reciprocity and shared enjoyment with both examiner and the child's caregiver during developmental evaluation. Other behaviors observed were free play, make-believe play, bubble play with the examiner, direct imitation, eye contact, and response to name. The pediatricians noted the overall language and communication used both verbally and non-verbally in the session, as well as any restricted, repetitive, and stereotyped behaviors, and also hyper- or hyporeactivity to sensory stimuli. The diagnosis of ASD and other neurodevelopmental disorders, including global developmental delay and language disorder, were based on the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5; American Psychiatric Association 2013) in addition to a team consensus agreement to verify the diagnosis of such disorders in participants.

Statistical Analysis

Statistical analysis was performed using Stata version 13.1 (Stata Corp., College Station, Texas). For descriptive analysis, frequencies of categorical variables were calculated, while median and interquartile ranges (IQR) were calculated for continuous variables. The non-parametric Wilcoxon rank-sum test was used to compare continuous variables between the two groups. Chi-square test was used in case of categorical variables. The internal consistency of the M-CHAT for each scoring method was calculated. Sensitivity, specificity, PPV, and NPV with 95 % confidence interval (CI) were calculated for each scoring method of the M-CHAT using the final diagnosis of ASD based on the DSM-5 criteria (American Psychiatric Association 2013) as the gold standard. Moreover, item-level data for each M-CHAT question that helps discriminate ASD from non-ASD were also performed and demonstrated as sensitivity and specificity with 95 % CI to uniquely identify those with ASD in this population. Receiver operating characteristic (ROC) analysis was used in the evaluation of the M-CHAT for each scoring method. All p values reported are two-sided in which statistical significance was defined as p of $<.05$.

Results

Demographic Characteristics of Participants

Participants in the high-risk group were significantly older than those in the low-risk group [median age 30 (IQR 26–35) vs. 21 (IQR 18–30) months, $p < .001$], although they had similar age ranges (18–48 months). Children whose parents had concerns whether their children's language development might be delayed were predominantly male compared with those whose parents did not have any concerns during the time of the M-CHAT screening (77.1 vs. 48.8 %, $p < .001$). The median age at which those in the high-risk group had their first word was 24 months whereas those in the low-risk group had the median age of the first word at 12 months. As expected, those in the high-risk group appeared to have a family history of

developmental and behavioral problems compared with those in the low-risk group (36.7 vs. 7.5 %, $p < .001$). There were no significant differences in other demographic characteristics including main caregiver of the child, paternal and maternal age, father's education, and family income between both groups of participants, except for mother's education where the number of mothers with at least a bachelor's degree were significantly lower in the high-risk group compared with those in the low-risk group (39.5 vs. 60.2 %, $p = .001$). Moreover, comparisons of demographic variables between those who screened positive and negative on the M-CHAT in each group of participants, and also between those who screened positive on the M-CHAT in both high-risk and low-risk groups are demonstrated in Table 2.

Diagnosis of ASD and Other Neurodevelopmental Disorders

Of 109 participants in the high-risk group, 45 (41.3 %) individuals were finally diagnosed with ASD, 53 (48.6 %) with language disorder, 8 (7.3 %) with global developmental delay, and 3 (2.8 %) participants were classified as typically developing children based on comprehensive developmental and behavioral evaluation by developmental and behavioral pediatricians with respect to the DSM-5 criteria (American Psychiatric Association 2013).

With regard to participants in the low-risk group, 10 (1.4 %) individuals who screened positive on the two-step M-CHAT were referred to be comprehensively evaluated by developmental and behavioral pediatricians. There were nine children who were finally diagnosed with ASD, whereas one individual was diagnosed with global developmental delay. Among 722 individuals in the low-risk group who screened negative on the M-CHAT, 276 (38.2 %) were evaluated by developmental and behavioral pediatricians on the same date as the two-step M-CHAT screening since most of them (97.1 %) were originally recruited in our group's cohort. We reviewed the medical records of 105 (14.5 %) individuals and 269 (37.3 %) were followed-up on the telephone with their mothers to confirm whether those who screened negative on the two-step

Table 2 Demographic characteristics of participants

Variable	Low-risk group			p^a	High-risk group			p^a	p^b	p^c
	Total (n = 732)	M-CHAT screening			Total (n = 109)	M-CHAT screening				
		Negative (n = 722)	Positive (n = 10)			Negative (n = 68)	Positive (n = 41)			
Age (months)	21 (18–30)	21 (18–29)	24 (24–42)	.09	30 (26–35)	30 (25–35)	30 (27–35)	.54	.18	<.001
Male gender	357 (48.8)	350 (48.5)	7 (70.0)	.18	84 (77.1)	51 (75.0)	33 (80.5)	.51	.47	<.001
Age with first word (months)	12 (12–12)	12 (12–12)	17 (12–20)	.01	24 (17–24)	20 (14–24)	24 (18–26)	.01	.12	<.001
Family history of developmental and behavioral problems	55 (7.6)	53 (7.4)	2 (20.0)	.13	40 (36.7)	26 (38.2)	14 (34.1)	.67	.39	<.001
Main caregiver										
Mother	450 (61.5)	443 (61.4)	7 (70.0)	.86	72 (66.1)	45 (66.2)	27 (65.9)	.59	.66	.20
Grandmother/grandfather	194 (26.5)	192 (26.6)	2 (20.0)		26 (23.9)	18 (26.5)	8 (19.5)			
Father's age (years)	35 (31–39)	35 (31–39)	37 (32–42)	.24	36 (30–40)	35 (30–40)	36 (31–43)	.87	.44	.39
Mother's age (years)	33 (30–37)	34 (30–37)	33 (32–36)	.88	33 (28–37)	34 (29–38)	33 (29–37)	.49	.70	.50
Father's education										
Bachelor's degree and higher	388 (56.7)	385 (57.1)	3 (30.0)	.09	47 (46.5)	29 (47.5)	18 (45.0)	.80	.39	.05
Mother's education										
Bachelor's degree and higher	440 (64.1)	436 (64.5)	4 (40.0)	.11	43 (42.6)	26 (42.6)	17 (42.5)	.99	.89	<.001
Family income >30,000 Baht/month[d]	431 (62.8)	426 (63.0)	5 (50.0)	.40	63 (62.4)	44 (72.1)	19 (47.5)	.01	.89	.93

Categorical variables are presented as number and percentage (in parentheses) whereas continuous variables are presented as median and interquartile ranges (in parentheses)

The national income per capita per month (Baht) based on the Office of the National Economic and Social Development Board, Office of the Prime Minister, Thailand in 2015 was 16,778.50 Baht

[a] p value for comparison of demographic characteristics between those who screened negative and positive on the M-CHAT in each group of participants

[b] p value for comparison of demographic characteristics between those who screened positive on the M-CHAT in the low- and high-risk groups

[c] p value for comparison of demographic characteristics between total samples of both groups of participants

[d] Family income was derived from the summation of father's and mother's income

M-CHAT still had not been diagnosed with ASD during the period of this study. However, 72 (10.0 %) participants who screened negative on the M-CHAT could not be contacted to recheck their developmental and behavioral status. Among those with negative M-CHAT screening who could be contacted (90 %), none were classified as having the diagnosis of ASD. The flowchart of participants who underwent the two-step M-CHAT screening is displayed (Fig. 1) and the demographic characteristics of children with and without ASD are demonstrated in Table 3.

The M-CHAT Screening in Participants

In the high-risk group, there were 41 (37.6 %) and 68 (62.4 %) individuals out of 109 participants who screened positive and negative on the two-step M-CHAT, respectively. Of the 41 participants who screened positive on the M-CHAT, 40 (97.6 %) were diagnosed with ASD (PPV), whereas 63 out of 68 (92.6 %) individuals who screened negative on the M-CHAT had not been diagnosed with ASD (NPV).

Among 45 participants who were diagnosed with ASD by developmental pediatricians in the high-risk group, 40 (88.9 %) were defined as failing the two-step M-CHAT screening (true positive rate or sensitivity) while 5 (11.1 %) individuals were classified as false negative on the M-CHAT based on the total scoring system. Most participants without ASD (63/64, 98.4 %) screened negative on the M-CHAT (true negative rate or specificity) based on the total scoring method, whereas there was only one participant (1/64, 1.6 %) without ASD who screened positive on the two-step M-CHAT (false positive). This particular individual who screened positive on the M-CHAT screening with a total score of 3 was finally diagnosed with global developmental delay.

Of 732 participants in the low-risk group, 207 (28.3 %) children initially screened positive on the M-CHAT by only parent reporting. Of 207 participants with positive screening on the M-CHAT, 10 (4.8 %) continued to screen positive after the semi-structured follow-up interview, whereas 197 (95.2 %) were false-positive cases after the follow-up interview was conducted. None of the screen positive cases declined the evaluation. Among 207 participants with positive screening on the M-CHAT that required the second step M-CHAT screening, 122 (58.9 %) received the follow-up interview in person at the visit whereas 85 (41.1 %) received the follow-up interview on the phone within 2 months. None of the latter group was unable to be reached to complete the follow-up questions.

As described previously, 10 (1.4 %) out of 732 participants in the low-risk group screened positive on the two-step M-CHAT whereas the remainder of participants (98.6 %) in this group screened negative on the two-step M-CHAT (Fig. 1). There were no participants with ASD having false negative M-CHAT screening in the low-risk

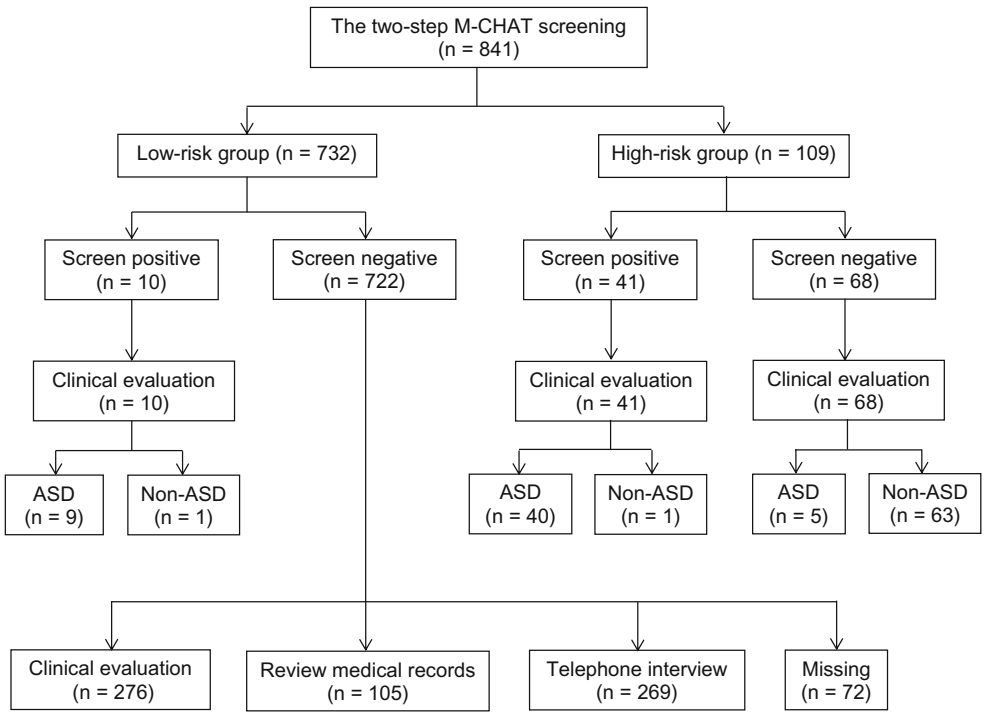

Fig. 1 Flowchart of participants underwent the two-step M-CHAT screening

Table 3 Demographic characteristics of children with ASD and without ASD diagnosis

Variable	Total (n = 841)	ASD		*p*
		With (n = 54)	Without (n = 787)	
Age (months)	24 (18–30)	30 (25–35)	23 (18–30)	<.001
Male gender	441 (52.4)	44 (81.5)	397 (50.4)	<.001
Age with first word (months)	12 (12–18)	24 (18–27)	12 (12–17)	<.001
Family history of developmental and behavioral problems	95 (11.3)	15 (27.8)	80 (10.2)	<.001
Main caregiver				
Mother	522 (62.1)	37 (68.5)	485 (61.6)	.64
Grandmother/grandfather	220 (26.2)	10 (18.5)	210 (26.7)	
Father's age (years)	36 (32–40)	37.5 (32–45)	36 (32–40)	.15
Mother's age (years)	34 (30–37)	33.5 (29–37)	34 (30–38)	.96
Father's education				
Bachelor's degree and higher	435 (51.7)	25 (46.3)	410 (52.1)	.17
Mother's education				
Bachelor's degree and higher	483 (57.5)	25 (46.3)	458 (58.3)	.28
Family income >30,000 Baht/month	493 (58.7)	28 (51.9)	465 (59.2)	.27

Categorical variables are presented as number and percentage (in parentheses) whereas continuous variables are presented as median and interquartile ranges (in parentheses)

group. In other words, all individuals who screened negative on the two-step M-CHAT screening were classified as non-ASD after the developmental and behavioral status of the child was known. Likewise, the individual who had false positive on the M-CHAT screening in the low-risk group was diagnosed with global developmental delay as mentioned above. In our whole sample, the total scoring method of the two-step M-CHAT screening appeared to provide relatively high sensitivity of 90.7 %, specificity of 99.7 %, PPV of 96.1 %, and NPV of 99.4 % based on the total scoring criteria (Table 4).

Item-Level Data of Each M-CHAT Question that Helps Discriminate ASD from Non-ASD

Based on the item-level data of each M-CHAT question, there were key questions including interest in other children, protoimperative pointing, protodeclarative pointing, functional play, bringing objects to show, imitation of actions, and following eye gaze (items 2, 6, 7, 8, 9, 13, 17) that tended to provide acceptable sensitivity and higher specificity to help discriminate ASD from non-ASD in this population (Table 5). These items were still the better key questions to differentiate those with ASD from individuals without ASD in the high-risk group except for item 13. The question item 21, "Does your child understand what people say?" was more likely to better identify those with ASD for high-risk individuals since it yielded higher sensitivity and specificity compared with item 13 (data were not shown for the high-risk group).

Comparison of Sensitivity, Specificity, PPV, and NPV Between the Total and the Critical Scoring Methods of the Two-Step M-CHAT Screening

With respect to internal consistency of the scoring methods on the M-CHAT, it was 0.81 for the total scoring and 0.78 for the critical scoring method. Regarding comparison of sensitivity, specificity, PPV, and NPV between the total and the critical scoring methods in all participants, the total scoring criteria (failing ≥3 items) was more likely to yield highest sensitivity, specificity, PPV, and NPV compared

Table 4 Comparison of sensitivity, specificity, PPV, and NPV between the total and the critical scoring methods of the M-CHAT, Thai version in all participants

Parameter	Total scoring method	95 % CI	Critical scoring method	95 % CI
Sensitivity	90.7	[79.7, 96.9]	75.9	[62.4, 86.5]
Specificity	99.7	[98.9, 99.9]	99.7	[99.1, 100.0]
PPV	96.1	[84.1, 98.8]	95.3	[84.2, 99.4]
NPV	99.4	[98.5, 99.8]	98.4	[97.2, 99.1]
AUC	95.2	[91.3, 99.1]	87.8	[82.1, 93.6]

CI confidence interval, *PPV* positive predictive value, *NPV* negative predictive value, *AUC* area under the curve

Table 5 Item-level data for each M-CHAT question that helps discriminate ASD from non-ASD in all participants

M-CHAT item	Sensitivity	95 % CI	Specificity	95 % CI
1. Enjoy being swung	0.0	[0.0, 6.6]	98.1	[96.9, 98.9]
2. Interest in other children	46.3	[32.6, 60.4]	99.1	[98.2, 99.6]
3. Climbing on things	1.9	[0.0, 9.9]	98.6	[97.5, 99.3]
4. Enjoy playing peek-a-boo	11.1	[4.2, 22.6]	99.6	[98.9, 99.9]
5. Pretend play	25.9	[15.0, 39.7]	99.5	[98.7, 99.9]
6. Protoimperative pointing	46.3	[32.6, 60.4]	99.5	[98.7, 99.9]
7. Protodeclarative pointing	57.4	[43.2, 70.8]	98.7	[97.7, 99.4]
8. Functional play	42.6	[29.2, 56.8]	95.6	[93.9, 96.9]
9. Bringing objects to show	68.5	[54.4, 80.5]	96.1	[94.5, 97.3]
10. Eye contact	24.1	[13.5, 37.6]	99.5	[98.7, 99.9]
11. Oversensitive to noise	16.7	[7.9, 29.3]	82.3	[79.5, 84.9]
12. Response to smile	1.9	[0.0, 9.9]	99.9	[99.3, 100.0]
13. Imitation of actions	44.4	[30.9, 58.6]	97.8	[96.6, 98.7]
14. Response to name	33.3	[21.1, 47.5]	99.4	[98.5, 99.8]
15. Following a point	35.2	[22.7, 49.4]	99.9	[99.3, 100.0]
16. Walking	0.0	[0.0, 6.6]	100.0	[99.5, 100.0]
17. Following eye gaze	53.7	[39.6, 67.4]	96.3	[94.8, 97.5]
18. Unusual finger movements	29.6	[18.0, 43.6]	81.1	[78.2, 83.7]
19. Getting parents' attention	37.0	[24.3, 51.3]	99.0	[98.0, 99.6]
20. Deafness	29.6	[18.0, 43.6]	95.6	[93.9, 96.9]
21. Language comprehension	38.9	[25.9, 53.1]	99.2	[98.3, 99.7]
22. Staring episodes	33.3	[21.1, 47.5]	95.6	[93.9, 96.9]
23. Social reference	27.8	[16.5, 41.6]	95.7	[94.0, 97.0]

CI confidence interval

with the critical scoring method (Fig. 2). A comparison of sensitivity, specificity, PPV, and NPV between the total and the critical scoring methods in both groups of participants is shown in Table 4.

The total scoring method still produced the highest sensitivity and NPV in those in the high-risk group when compared to the critical scoring method. Although the critical scoring criteria yielded the highest specificity and PPV of 100 for those in the high-risk group, it had relatively lower sensitivity of 75.6 (95 % CI 60.5–87.1) compared with the total scoring method. If we excluded those in the low-risk group who were unable to be reached after screening negative on the two-step M-CHAT screening, our results remain unchanged. Moreover, the total scoring method was likely to have higher sensitivity, specificity, PPV, and NPV than the critical scoring criteria used for both younger (18–30 months old) and older individuals (>30–48 months old).

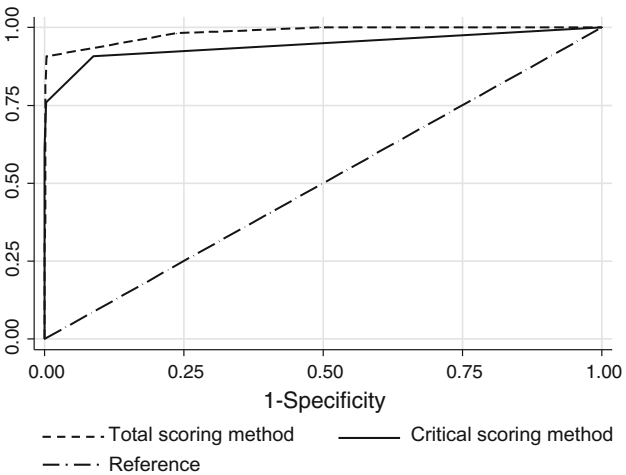

Fig. 2 Receiver operating characteristic (ROC) curves for the total and the critical scoring methods of the M-CHAT, Thai version

Discussion

To our knowledge, this is the first study attempting to validate the use of the two-step M-CHAT screening relevant to Thai culture. This was conducted by implementing a combination of an initial parent report and the semi-structured follow-up interview on site for response items that indicated a risk for ASD only for cases that initially screen positive. A trained clinician conducted a follow-up interview immediately or within a 2-month period so the clinician was able to identify ASD in 18–48 month-old, both high-risk and

low-risk groups of Thai children. The two-step M-CHAT screening in this present study appeared to yield good sensitivity (75.9–90.7 %), high PPV (95.3–96.1 %), very high specificity (99.7 %), and NPV (98.4–99.4 %) according to the scoring criteria used for the whole sample. However, it is important to note that using a two-stage screening process does not improve sensitivity. The primary goal is to reduce the false positive rate, which most affects PPV, but the sensitivity may actually decrease slightly.

Sensitivity, specificity, and predictive power of the two-step M-CHAT screening used in this present study were consistent with discriminant function analysis documented in the first M-CHAT research of its type by Robins et al. (2001) and also other literature (Kleinman et al. 2008; Pandey et al. 2008; Robins 2008). However, the PPV demonstrated in our study was generally higher than what was reported in previous studies (Kamio et al. 2015; Pandey et al. 2008; Robins 2008; Robins et al. 2001). This particular finding might be due to the fact that the follow-up interview for response items that demonstrated a risk for ASD was entirely performed by clinicians who have expertise in screening both high- and low-risk individuals for ASD. This has previously been demonstrated in Kara et al.'s study in which the M-CHAT was conducted by health care providers (Kara et al. 2014), thereby reducing false-positive rates and ultimately resulting in improvement in the PPV.

Moreover, the screen-positive rate following the two-step M-CHAT screening for low-risk individuals in our study was 1.4 %, which was relatively similar to the rate of screen-positive after the follow-up interview in a large population study (272 out of 18,989 children) by Chlebowski, Robins, Barton, and Fein (2013). As such, an additional 2-to-10-min follow-up interview by the clinician after positive screening by initial parent response on the M-CHAT should be considered even in the context of limitations on resources and time constraints. Relying solely on parent-completed questionnaires, especially from a mother with lower education, could be inaccurate, particularly in Thai culture where parents are more familiar with the interviewing process of being read the questions rather than completing questionnaires on their own. Furthermore, this second step screening should be incorporated into the general practice of ASD screening due to the fact that it is valuable in helping to minimize the false-positive results as documented in previous studies (Robins 2008; Seung et al. 2015). Otherwise, a waiting list of referral for comprehensive developmental and behavioral evaluation, which is generally limited in Thailand, in children with false-positive screening, will be too long and may unnecessarily provoke parental anxiety of such individuals for a period of time. Those with positive screens were also likely to lose follow-up.

With respect to the scoring methods of the two-step M-CHAT screening, the total scoring criteria (failing $\geq$3 of any items) was more likely to yield the highest sensitivity, specificity, PPV, and NPV than the critical scoring method as also documented in the studies by Robins et al. (2014) and Chlebowski et al. where the total scoring cut-off on the M-CHAT could identify more screen-positive cases than the critical scoring method. Furthermore, the total scoring criteria still hold promise for both diagnostic and predictive power of the two-step M-CHAT screening for younger (18–30 months of age) and older (> 30–48 months of age) individuals in our study, which was in contrast to the study by Koh et al. (2014) where the critical (failing $\geq$2 critical items) scoring criteria were far better than the total scoring method in detection of true-positive cases of ASD in young Singaporean toddlers. The varying findings between ours and Koh et al.'s study plausibly resulted from differences in study populations (high- and low-risk groups vs. only high-risk group), M-CHAT screening method used (initial parent completion followed by re-interview for response items being at risk for ASD versus only parent completion of the M-CHAT), and expectations on child behaviors and development with regard to each cultural background.

Regarding item-level data analysis, key questions that were likely to help discriminate ASD from non-ASD in young Thai participants included interest in other children (item 2), protoimperative pointing (item 6), protodeclarative pointing (item 7), functional play (item 8), bringing objects to show (item 9), imitation of actions (item 13), and following eye gaze (item 17). There were three items including protoimperative pointing, functional play, and following eye gaze that appeared to yield better sensitivity and specificity than other critical items including response to name and following a point, originally reported by Robins et al. (2001). Nonetheless, bringing objects to show, protodeclarative pointing, interest in other children, and imitation of actions were common critical items that were sensitive for children with ASD both in the US and Thailand.

The difference in key questions between our study and the study by Kamio et al. (2015) from Japan was also observed where pretend play, point following, and language comprehension were noted to be more helpful to differentiate 18-month-old Japanese toddlers with ASD from typically developing children than those reported in our study. This dissimilarity of key questions among studies was likely due to varying degrees of parental sensitivity and familiarity in observation of certain child behavior which is specific and relevant to each culture. For instance, protoimperative pointing and functional or pretend play were preverbal social behaviors which were more sensitive and were likely to be easier observed by the parent of an individual with ASD than "response to name"

behavior in Asian culture when compared to Western culture. We speculated that children in Asian culture generally live with other family members rather than only their parents. They were likely to have their names called often by everyone in the household and in turn, these children may choose not to respond when their names are called because they are accustomed to hearing their names often, even though they are typically developing. However, "bringing objects to show" and "imitation of actions" seem to be common items that are the most sensitive to discriminate individuals with ASD from those without in various ethnicities (Kamio et al. 2015; Robins 2008; Robins et al. 2001; Stenberg et al. 2014; Wong et al. 2004). The difference in the critical items or key questions relatively relevant to each cultural context is interesting; however, this topic goes beyond the scope of this study.

There have been no updated studies on the rate of ASD in Thailand for more than a decade. The rate of ASD in the low-risk sample in our study was 12.3 in 1000 or approximately one in 81 (9 out of 732 cases), which was comparable to what has been reported in the literature to date (Christensen et al. 2016; Kamio et al. 2014; Wingate et al. 2014; Yang et al. 2015). Although ASD is prevalent in many countries around the world, there have been tremendous recent debates on universal screening for ASD in an era where ASD has become an important health issue, particularly if young affected individuals and those who have limited access to resources are still undetected and do not receive appropriate intervention. As such, we argue more strongly that universal screening should be incorporated into routine health supervision visits, particularly in communities where the health-care system regarding early ASD screening, referral for appropriate diagnosis and management are well established, as demonstrated in our setting.

There are some strengths of this present study in that most follow-up interviews for response items being at risk for ASD occurred immediately on site after initial parent report. As such, each particular item could be thoroughly explained in person by a clinician who is knowledgeable about ASD, which resulted in reducing the time between initial parent screening and the follow-up interview. Moreover, those who screened negative on the two-step M-CHAT were mostly evaluated by developmental and behavioral pediatricians or had their medical records reviewed or were followed-up on the telephone to verify the status of non-ASD diagnosis.

Although the two-step M-CHAT screening used in this study was likely to be a useful instrument in the detection of ASD in both high-risk and low-risk individuals in Thailand, there are still some limitations in that this study was conducted in only one university-based hospital in Bangkok. Therefore, there could be a referral bias for those in the high-risk group. However, children who had been diagnosed with ASD before receiving an evaluation at our center were excluded from this study. Moreover, those in the low-risk group, not only in the high-risk sample, were also enrolled to expand the clinical utility of the two-step M-CHAT screening in a Thai context.

In addition, we used a semi-structured interview approach, as opposed to a more structured and scripted format, for the follow-up interview in this present study, which could result in considerable variable reliability across participants. Changes in the wording and nature of the items described in our semi-structured interview may also result in some discrepancy from the original structured M-CHAT follow-up interview recommended by Robins (2008). However, all interviewers were trained and closely supervised by the corresponding author to establish reliability among interviewers to ensure that the follow-up interview process provided a range of examples and reworded questions to support comprehension.

Furthermore, we speculated that a less structured and scripted approach of our semi-structured interview could improve the feasibility of the follow-up interview in general pediatric practice. Our results demonstrated that almost 60 % of those with positive M-CHAT screening who required the second step M-CHAT screening received a follow-up interview in person at the visit. Nonetheless, this feasibility might be diminished in case of busier general pediatric practice settings and exhibited less reliability across informants particularly if the interviewers do not receive appropriate training on the follow-up interview.

Children with significant medical or known genetic histories or syndromes were excluded from our study. As a result, how the Thai M-CHAT would perform to identify autism with this particular population cannot be known. Moreover, our participants in the low-risk group were considered very low risk for ASD compared with those ascertained from population based screening of children at well-child visits enrolled in the original study by Robins et al. (2001), because children whose caregivers or pediatricians had concerns about development were excluded. This limitation may be addressed in future studies that include children with a broader range of prenatal/postnatal medical complications and the child's development.

As a referral center for evaluation and management for individuals with ASD and the fact that our participants' parents had high educational levels, our findings, especially sensitivity, specificity, PPV, and NPV, should be cautiously interpreted. These findings may not be applicable to other health-care settings in Thailand, particularly in

disadvantaged communities where caregivers of at-risk children for ASD tended to have lower levels of education and there were likely to be more limitations on resources. Moreover, the clinical diagnosis of ASD in this present study was made by developmental and behavioral pediatricians who are experts in the field based on the DSM-5 criteria (American Psychiatric Association 2013) without using the Autism Diagnostic Observation Schedule (ADOS) and the Autism Diagnostic Interview-Revised (ADI-R) (Lord et al. 1994, 1999), despite the fact that the corresponding author has received clinical training in both standardized ASD diagnostic assessments. This was mainly due to time constraints. However, the corresponding author and a senior author of this study have experience in the evaluation and management of ASD for almost 10 and 30 years, respectively.

In summary, the two-step M-CHAT screening, which consisted of an initial parent-completed response on the questionnaire, followed by the interview for response items that were at risk for ASD only for cases that initially screen positive, appeared to be a promising useful tool to screen for ASD in both high-risk and low-risk individuals in Thailand. Compared with the critical scoring method, the total scoring or failing $\geq$3 of any items on the M-CHAT was the best scoring criteria that provided the highest sensitivity of 90.7 %, specificity of 99.7 %, PPV of 96.1 %, and NPV of 99.4 % respectively. To achieve appropriate clinical utility of the M-CHAT relevant to each population and setting, the socio-cultural context should be seriously considered when adopting the use and interpretation of the M-CHAT for each country.

Acknowledgments We are very grateful to all study participants and their families who participated in this study; Associate Professor Chitsanu Pancharoen and Assistant Professor Therdpong Tempark for their help with participant enrollment; Dr. Dhanika Boonma, Dr. Nattaporn Tassanakijpanich, and Dr. Jarujan Subchartanan, developmental and behavioral pediatric fellows at the Division of Growth and Development, Department of Pediatrics for their assistance on the telephone follow-up for those who passed the M-CHAT screening to verify whether participants still have not had the diagnosis of ASD during the period of this study; Ms. Jirachaya Sophonphan for her help on statistical analysis; Dr. Juthamas Worachotekamjorn and Assistant Professor Sureelak Sutchritpongsa for their helpful comments and suggestions on our manuscript; and Sam Ormond who professionally edited this entire manuscript. This study was supported by a grant from Ratchadapiseksompotch Fund, Faculty of Medicine, Chulalongkorn University (Grant Number: RA 58/004), and partially funded by a grant from the Thailand Research Fund (Grant Number: IRG5780015), Bangkok, Thailand.

Author Contribution PS participated in the design of the study, data collection, coordination and drafted the manuscript; CP participated in the design, performed the measurement and helped to revise the manuscript; WC conceived of the study, participated in its design, data collection and coordination, performed the measurement, the statistical analysis, interpretation of the data, helped to draft and revise the manuscript critically for important intellectual content. All authors read and approved the final manuscript.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical Standards All procedures performed in this study involving human participants were in accordance with the ethical standards of the institutional review board and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders, fifth edition (DSM-5)*. Arlington, VA: American Psychiatric Association.

Canal-Bedia, R., Garcia-Primo, P., Martin-Cilleros, M. V., Santos-Borbujo, J., Guisuraga-Fernandez, Z., Herraez-Garcia, L., et al. (2011). Modified checklist for autism in toddlers: Cross-cultural adaptation and validation in Spain. *Journal of Autism and Developmental Disorders, 41*(10), 1342–1351. doi:10.1007/s10803-010-1163-z.

Charman, T., Baird, G., Simonoff, E., Chandler, S., Davison-Jenkins, A., Sharma, A., et al. (2016). Testing two screening instruments for autism spectrum disorder in UK community child health services. *Developmental Medicine and Child Neurology, 58*(4), 369–375. doi:10.1111/dmcn.12874.

Chlebowski, C., Robins, D. L., Barton, M. L., & Fein, D. (2013). Large-scale use of the modified checklist for autism in low-risk toddlers. *Pediatrics, 131*(4), e1121–e1127. doi:10.1542/peds.2012-1525.

Christensen, D. L., Bilder, D. A., Zahorodny, W., Pettygrove, S., Durkin, M. S., Fitzgerald, R. T., et al. (2016). Prevalence and characteristics of autism spectrum disorder among 4-year-old children in the autism and developmental disabilities monitoring network. *Journal of Developmental and Behavioral Pediatrics, 37*(1), 1–8. doi:10.1097/DBP.0000000000000235.

Feldman, H. M., & Messick, C. (2009). Language and speech disorders. In W. B. Carey, A. C. Crocker, W. L. Coleman, E. R. Elias, & H. M. Feldman (Eds.), *Developmental-behavioral pediatrics* (4th ed., pp. 717–729). Philadelphia, PA: Saunders Elsevier.

Frankenburg, W. K., Dodds, J., Archer, P., Shapiro, H., & Bresnick, B. (1992). The Denver II: A major revision and restandardization of the Denver Developmental Screening Test. *Pediatrics, 89*(1), 91–97.

Garcia-Primo, P., Hellendoorn, A., Charman, T., Roeyers, H., Dereu, M., Roge, B., et al. (2014). Screening for autism spectrum disorders: State of the art in Europe. *European Child and Adolescent Psychiatry, 23*(11), 1005–1021. doi:10.1007/s00787-014-0555-6.

Hoon, A. H, Jr., Pulsifer, M. B., Gopalan, R., Palmer, F. B., & Capute, A. J. (1993). Clinical Adaptive Test/Clinical Linguistic Auditory Milestone Scale in early cognitive assessment. *Journal of Pediatrics, 123*(1), S1–S8.

Johnson, C. P., Myers, S. M., Lipkin, P. H., Cartwright, J. D., Desch, L. W., Duby, J. C., et al. (2007). Identification and evaluation of children with autism spectrum disorders. *Pediatrics, 120*(5), 1183–1215. doi:10.1542/peds.2007-2361.

Kamio, Y., Haraguchi, H., Stickley, A., Ogino, K., Ishitobi, M., & Takahashi, H. (2015). 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, 45*(12), 4147–4153. doi:10.1007/s10803-015-2527-1.

Kamio, Y., Inada, N., Koyama, T., Inokuchi, E., Tsuchiya, K., & Kuroda, M. (2014). Effectiveness of using the Modified Checklist for Autism in Toddlers in two-stage screening of autism spectrum disorder at the 18-month health check-up in Japan. *Journal of Autism and Developmental Disorders, 44*(1), 194–203. doi:10.1007/s10803-013-1864-1.

Kara, B., Mukaddes, N. M., Altinkaya, I., Guntepe, D., Gokcay, G., & Ozmen, M. (2014). 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, 18*(3), 331–338. doi:10.1177/1362361312467864.

Kleinman, J. M., Robins, D. L., Ventola, P. E., Pandey, J., Boorstein, H. C., Esser, E. L., et al. (2008). 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), 827–839. doi:10.1007/s10803-007-0450-9.

Koh, H. C., Lim, S. H., Chan, G. J., Lin, M. B., Lim, H. H., Choo, S. H., et al. (2014). 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, 44*(2), 405–416. doi:10.1007/s10803-013-1880-1.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (1999). *Autism diagnostic observation schedule*. Los Angeles, CA: Western Psychological Services.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*(5), 659–685.

Mullen, E. M. (1995). *Mullen scales of early learning*. Circle Pines: American Guidance Service.

Pandey, J., Verbalis, A., Robins, D. L., Boorstein, H., Klin, A., Babitz, T., et al. (2008). Screening for autism in older and younger toddlers with the Modified Checklist for Autism in Toddlers. *Autism, 12*(5), 513–535. doi:10.1177/1362361308094503.

Robins, D. L. (2008). Screening for autism spectrum disorders in primary care settings. *Autism, 12*(5), 537–556. doi:10.1177/1362361308094502.

Robins, D. L., Casagrande, K., Barton, M., Chen, C. M., Dumont-Mathieu, T., & Fein, D. (2014). Validation of the modified checklist for Autism in toddlers, revised with follow-up (M-CHAT-R/F). *Pediatrics, 133*(1), 37–45. doi:10.1542/peds.2013-1813.

Robins, D. L., Fein, D., Barton, M. L., & Green, J. A. (2001). The Modified Checklist for Autism in Toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 31*(2), 131–144.

Seung, H., Ji, J., Kim, S. J., Sung, I., Youn, Y. A., Hong, G., et al. (2015). Examination of the Korean modified checklist of autism in toddlers: Item response theory. *Journal of Autism and Developmental Disorders, 45*(9), 2744–2757. doi:10.1007/s10803-015-2439-0.

Snow, A. V., & Lecavalier, L. (2008). Sensitivity and specificity of the Modified Checklist for Autism in Toddlers and the Social Communication Questionnaire in preschoolers suspected of having pervasive developmental disorders. *Autism: The International Journal of Research and Practice, 12*(6), 627–644. doi:10.1177/1362361308097116.

Stenberg, N., Bresnahan, M., Gunnes, N., Hirtz, D., Hornig, M., Lie, K. K., et al. (2014). Identifying children with autism spectrum disorder at 18 months in a general population sample. *Paediatric and Perinatal Epidemiology, 28*(3), 255–262. doi:10.1111/ppe.12114.

Warren, Z., McPheeters, M. L., Sathe, N., Foss-Feig, J. H., Glasser, A., & Veenstra-Vanderweele, J. (2011). A systematic review of early intensive intervention for autism spectrum disorders. *Pediatrics, 127*(5), e1303–e1311. doi:10.1542/peds.2011-0426.

Wingate, M., Kirby, R. S., Pettygrove, S., Cunniff, C., Schulz, E., Ghosh, T., et al. (2014). Prevalence of autism spectrum disorder among children aged 8 years—Autism and Developmental Disabilities Monitoring Network, 11 sites, United States, 2010. *Morbidity and Mortality Weekly Report. Surveillance Summaries, 63*(2), 1–21.

Wong, V., Hui, L. H., Lee, W. C., Leung, L. S., Ho, P. K., Lau, W. L., et al. (2004). A modified screening tool for autism (Checklist for Autism in Toddlers [CHAT-23]) for Chinese children. *Pediatrics, 114*(2), e166–e176.

Yama, B., Freeman, T., Graves, E., Yuan, S., & Karen Campbell, M. (2012). Examination of the properties of the Modified Checklist for Autism in Toddlers (M-CHAT) in a population sample. *Journal of Autism and Developmental Disorders, 42*(1), 23–34. doi:10.1007/s10803-011-1211-3.

Yang, W., Xia, H., Wen, G., Liu, L., Fu, X., Lu, J., et al. (2015). Epidemiological investigation of suspected autism in children and implications for healthcare system: A mainstream kindergarten-based population study in Longhua District, Shenzhen. *BMC Pediatrics, 15*(1), 207. doi:10.1186/s12887-015-0531-4.

Zwaigenbaum, L., Bauman, M. L., Fein, D., Pierce, K., Buie, T., Davis, P. A., et al. (2015). Early screening of autism spectrum disorder: Recommendations for practice and research. *Pediatrics, 136*(Suppl 1), S41–S59. doi:10.1542/peds.2014-3667D.