

Modified Checklist for Autism in Toddlers: Cross-Cultural Adaptation and Validation in Spain

Ricardo Canal-Bedia · Patricia García-Primo · María Victoria Martín-Cilleros · José Santos-Borbujó · Zoila Guisuraga-Fernández · Lorena Herráez-García · María del Mar Herráez-García · Leticia Boada-Muñoz · Joaquín Fuentes-Biggi · Manuel Posada-de la Paz

Published online: 16 December 2010
© Springer Science+Business Media, LLC 2010

Abstract Early detection and treatment have been shown to be effective in reducing disability severity caused by Autistic Spectrum Disorders (ASDs). As Spanish pediatricians have no detection tool, the Modified Checklist for Autism in Toddlers (M-CHAT) was first translated into and culturally adapted to Spanish. Validity and reliability

studies were carried out in two different geographical areas of Spain, where M-CHAT was administered to two different samples, namely: 2,480 high- and low-risk children; and 2,055 low-risk children. The results obtained were similar to those yielded by the original M-CHAT studies. Differences were found in positive predictive value, due to the low ASD frequency observed in this study. M-CHAT is still being studied in a large population-based screening program in Spain.

On behalf of the MCHAT study (in alphabetical order by surname): Belinchón-Carmona M. (Autonomous University of Madrid); Espinosa A. (Gregorio Marañón University Teaching Hospital); Ferrari-Arroyo MJ. (Madrid Municipal Health Prevention Programs); Franco M. (Zamora Regional Hospital); Hernández J. (ASD Team—Regional Educational Authority); Mar Domingo MJ. (Gregorio Marañón University Teaching Hospital); Martínez-Velarte M. (Zamora Regional Hospital); Martínez-Conejo A. (University of Salamanca); Muñoz A. (*IRIDIA*); Palomo R. (*IRIDIA*); Parra E. (Gregorio Marañón University Teaching Hospital); Velayos L. (*IRIDIA*).

R. Canal-Bedia · M. V. Martín-Cilleros · J. Santos-Borbujó · Z. Guisuraga-Fernández · L. Herráez-García · M. del Mar Herráez-García
University of Salamanca, Salamanca, Spain

P. García-Primo (✉)
National Research Institute of Rare Diseases, Instituto de Salud Carlos III, Monforte de Lemos 5, 28029 Madrid, Spain
e-mail: pgarciaprimo@isciii.es

L. Boada-Muñoz
Gregorio Marañón University Teaching Hospital, Madrid, Spain

J. Fuentes-Biggi
Guipúzcoa Polyclinic and the Guipuzka Autism Association (GAUTENA), San Sebastián, Spain

M. Posada-de la Paz
Natioal Research Institute of Rare Diseases, Instituto de Salud Carlos III and The Centre for Network Research on Rare Diseases (Centro de Investigación Biomédica en Red de Enfermedades Raras – CIBERER), Madrid, Spain

Keywords Autism · Screening · Validation · M-CHAT · Psychometric values

Introduction

It is well known that delayed diagnosis and treatment of Autistic Spectrum Disorders (ASDs) can worsen the prognosis (Corsello 2005; Dawson and Osterling 1997; Kasari et al. 2006). Moreover, despite medical advances, no biological markers exist for establishing an ASD diagnosis, so that diagnosis is still based on detailed behavioral analysis (Lord et al. 2000). Early ASD detection has been based on a combination of different strategies, such as developmental surveillance through casual assessment of communication, social and play skills, and consideration of parent concerns (Filipek et al. 2000). Hence, there is growing consensus as to the need for viable strategies for early ASD identification, using different screening tools that allow for earlier appropriate diagnosis and subsequent intervention (Baird et al. 2000; Baron-Cohen et al. 2000; Canal-Bedia 2001; Dawson et al. 1998; Wetherby et al. 2004). Such early intervention strategies, focused on promoting and enhancing social communication and language skills, combined with social support for families, will have

a positive impact on the development of young children with autism, and also prevent other secondary difficulties, such as depression and family anxiety (Howlin et al. 2000; Rogers and Vismara 2008; Tonge et al. 2006).

To date, several tools have been designed for early ASD detection (Johnson et al. 2007; Mawle and Griffiths 2006; Robins and Dumont-Mathieu 2006). These tools have been classified as “level 1” screening tools because they can be administered to all children in primary care settings, and are designed to differentiate children at risk of ASDs from the general population. They require less time and training to administer, score and interpret than other, so-called “level 2” screening tools which are normally used in diagnostic-service settings. Among these level-1 screening tools are the Checklist for Autism in Toddlers (CHAT) (Baron-Cohen et al. 1992) and the Modified Checklist for Autism in Toddlers (M-CHAT) (Robins et al. 2001). The M-CHAT is a self-administered, 23-item questionnaire that was first developed and validated in the USA as a tool for detecting ASD in children aged under 2 years in a low-risk population, and does not require specialized direct observation.

To date, all M-CHAT studies have yielded similar results and indicate that M-CHAT could be an effective tool for early ASD screening (Inada et al. 2010; Pandey et al. 2008; Robins et al. 2001; Robins and Dumont-Mathieu 2006; Ventola et al. 2007).

In Spain, primary care pediatricians are the only professionals that maintain continuous contact with toddlers under 3 years old (not all children at these ages are cared for in nursery schools). While parents tend to bring their concerns about their children to pediatricians, the latter report that they lack time to conduct a systematic neuro-development assessment of all children (GETEA 2003). Furthermore, children’s behavior in the pediatrician’s office may not represent their typical behavior, and a physician may therefore find it difficult in the brief time available to decide whether a given child’s behavior indicates suspicion of ASD.

Although a Spanish translation of the M-CHAT was available at the time when this study commenced, this was a version which had been translated into the Spanish spoken in Latin American countries and had not been adapted to the cultural and language differences present in Spain. Moreover, there were no validity studies for this Latin American version. Hence, a cross-cultural validation of the Spanish version to be used in Spain was needed, something that could contribute to comparison of the tool’s psychometric features in other languages and countries.

Cross-cultural adaptations are typically based on the “back translation” method, in order to ensure equivalence between the original questionnaire and its version in another language. Nevertheless, concordance between an

original and a back-translation is no guarantee that the translated questionnaire will be equally effective in the target community where the cultural context is different (Beaton et al. 2000). A pilot study is first required to implement the questionnaire in the new cultural context and compare its psychometric properties with similar statistical procedures that have been used in previous studies of the instrument (International Test Commission 2010).

Thus, the translation, implementation and evaluation of the M-CHAT was tested at pediatricians’ offices belonging to the Spanish National Health System (SNHS) in two different regions before making a final decision on its broader, systematic implementation in the NHS nationwide. Accordingly, this paper sought to describe the procedures used in the translation and validation study of the Spanish version of the M-CHAT intended for use in Spain.

Methods

Setting

This study was conducted in two provinces in Northwest Spain, namely, Salamanca and Zamora, from October 1st, 2005 to April 15th, 2008. This was deemed to be the main study setting for performance of the validity analysis (Stage 1). In addition, a separate geographical area situated in the province of Madrid was chosen for testing M-CHAT reliability in another population, from April 15th, 2006 to April 15th, 2008 (Stage 2). Consequently, the two studies, albeit independent, were analogous in terms of the methods and strategies used. These two geographical areas were selected because their respective health authorities had previously decided to implement population-based, pilot ASD screening programs.

Brief description of the Spanish National Health System. Spain has a national health system that covers 100% of the population, regardless of their level of income and employment status. The system also cares for the legal immigrant population. The Spanish health system is decentralized into autonomous regions (*comunidades autónomas*), and each region is further subdivided into several health areas depending on the size of its population.

Families are assigned to a primary health care pediatrician, who is then responsible for the medical care of all their children from birth to 16 years of age. The SNHS also has a specific “Well Baby Check-up Program” (*Programa del Control del Niño Sano*), which allows for regular standardized examinations that collect data on the developmental milestones of each child (mainly physical condition). Families are not required to pay for any of these health services.

Study Population

Stage 1

All children of both genders aged 18–36 months, whose parents resided in the geographical area during the study period, were selected (children attending the mandatory measles, mumps and rubella vaccination (MMR) program at age 18 months, and/or the general well-baby check-up examination at age 24 months). High-risk children with developmental ages of 18–24 months (maximum chronological age, 48 months) from early intervention centers and/or child-adolescence psychiatric units in the same regions were also accepted to participate in this study, in order to increase the probability of recruiting autistic cases. A total of 2,480 children, 63 of whom were high-risk, were involved at this stage.

Stage 2

A population-based study was conducted covering the entire health-catchment area (Madrid Health Area No.1), extending from the center of Madrid to the city's outskirts. No high-risk children were included in this study, with selection being confined to children aged 18–36 months of either gender who attended the mandatory MMR vaccination program and/or well-baby check-up examination. A total of 2,055 children were included in this stage.

Case Definition

All diagnoses were made as per the Diagnostic and Statistical Manual of Mental Disorders-4th Edition (DSM_IV

TR) criteria, and supported by Autism Diagnostic Observation Schedule-Generic (ADOS-G) scoring. The DSM IV TR was likewise used for classification of type of ASD.

Participants' parents were required to sign an informed consent form that had been formally approved by the appropriate local Ethics Committee.

Procedure

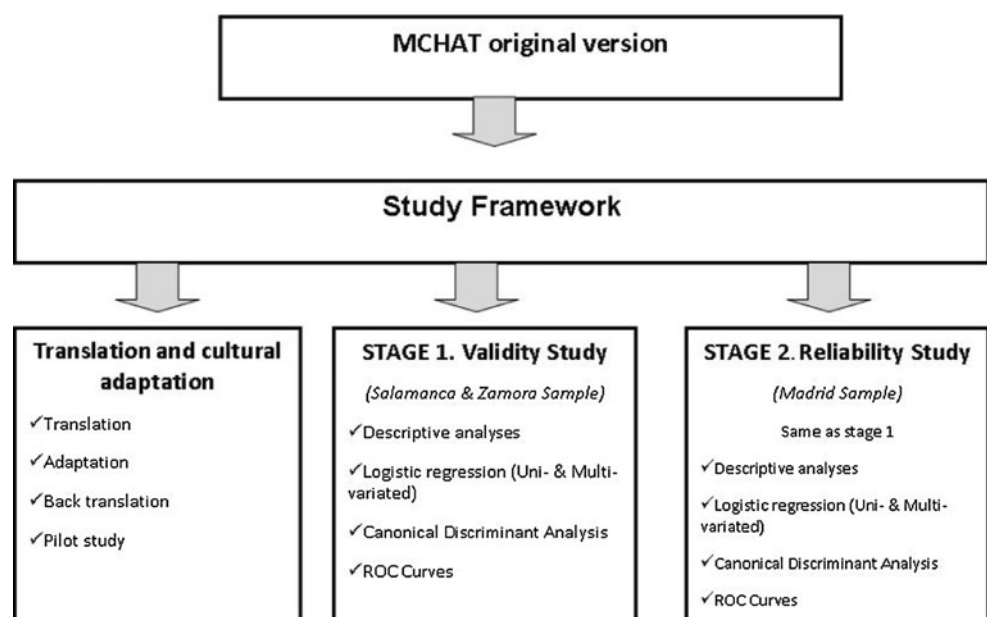
The Spanish M-CHAT validity study was undertaken in the following three phases: (i) translation, back translation, cultural adaptation and a short pilot study to obtain the final M-CHAT version to be used in Spain; (ii) the validity study itself (Stage 1); and (iii) the reliability study (Stage 2) (see Fig. 1).

Translation-Back Translation Method

Both the M-CHAT questionnaire and the 23 items of the M-CHAT telephone follow-up interview (phone FUI) were translated into Spanish by two bilingual persons with experience in assessing child development. They worked independently of each other and with instructions that, rather than being literal, the translation should seek semantic, linguistic and cultural equivalence. The resulting versions were then back-translated by a native English speaker who was bilingual, and compared to the original M-CHAT. The original phone FUI questions for each item were structured in flowchart format, complete with instructions and rules for use by interviewers.

After several exchanges of opinion with the original authors, and making certain amendments, phone FUI flowcharts were approved by the M-CHAT's original

Fig. 1 M-CHAT validity study flow-chart in Spain



authors (Robins et al. 2001) (hereinafter, M-CHAT's original authors). A pilot study of this preliminary, new Spanish M-CHAT version was conducted, using the first 622 children screened in Stage 1 to assess both its comprehensibility and its feasibility in the context of a real scenario in which it was to be subsequently implemented.

Stage 1. Validity Study

A total of 86 primary care pediatricians and nurses were involved in this study. A series of preliminary training courses were given on ASD characteristics, general M-CHAT procedures, implementation of a screening program, and methods and instructions for high-risk case referrals to the specialized diagnostic unit. The M-CHAT was administered to all children at the health care units. Once the study objectives had been explained and the informed consent signed, the nurse and/or physician gave each parent a copy of the M-CHAT.

When completed, the M-CHAT forms were sent to the central research unit and analyzed according to the original cut-off criterion, namely, 3 out of 23 or 2 out of the following 6 critical items: 2, 7, 9, 13, 14 and 15 (see Table 2 for keywords of these items) (Robins et al. 2001). Questionnaires with positive results were then confirmed by phone FUIs conducted by a psychologist with child-development training, who used the specific algorithms described above to re-evaluate each failed item, applying examples and real-life situations to confirm whether the item in question was to be finally regarded as a "fail" or a "pass." In any case where the number of confirmed failed items was still above the cut-off criterion, the child was then referred to the specialized ASD-diagnostic team (Salamanca University ASD unit). In parallel, a second group of children, drawn from the early intervention centers and/or child-adolescence psychiatric units in the same geographical region and displaying behaviors highly indicative of developmental delay or disorder, were also included in this study, following the same process as with the former group. This second group (high-risk children) was clinically evaluated regardless of the M-CHAT results.

An ASD team of psychologists and neurologists applied the algorithm diagnosis developed by the Autism Spectrum Disorder Study Group (*Grupo de Estudio de los Trastornos del Espectro Autista-GETEA*) (Díez-Cuervo et al. 2005), based on the recommendations of the American Academy of Neurology and the Child Neurology Society for the diagnosis of autism (Filipek et al. 2000). A formal interview designed to collect information from parents about their children, the Spanish versions of the Vineland Adaptive Behavior Scales (Sparrow et al. 1987) the Merrill-Palmer Revised Scales of Development (Roid and Sompers 2004) and the ADOS-G module 1 (Lord et al. 2000) were

used to obtain a final diagnosis in each case. Administration of the ADOS-G was videotaped by a student clinician throughout: whenever there was disagreement among the evaluators, the ADOS video was re-examined and the matter decided by consensus. All cases were classified according to the DSM-IV-TR (APA 2000).

Stage 2. Study Reliability

This study used the same methods for the M-CHAT screening program, phone FUI and diagnosis (when applicable) but no ASD high-risk children were included.

Statistical Analyses

A descriptive analysis of frequencies stratified by region, gender and age was made.

M-CHAT item results were considered "failed", when children passed this questionnaire because their cut-off points were below the abovementioned criterion. Otherwise, M-CHAT item results were only considered "failed" after the phone FUI confirmed the information. The percentage of children failing each M-CHAT item was then classified into the following four groups: (a) children needing no follow-up; (b) children whose parents were administered the phone FUI, but whose final outcome was "pass"; (c) children who were referred for diagnostic evaluation but were finally diagnosed with a developmental disorder (DD) other than ASD; and finally, (d) children diagnosed with ASD.

M-CHAT properties, such as sensitivity, specificity and predictive values, were estimated for each stage. Canonical discriminant analysis was also performed to analyze the M-CHAT's ability to distinguish among children with typical development (TD), developmental disorders (DDs) and ASD, when the M-CHAT was applied in a population-based setting.

All 23 M-CHAT items were checked to estimate the probability of predicting ASD cases versus the other two groups (TD and DD jointly), using both uni-variate and multi-variate logistic regression analyses. As complementary information, areas under the response operating curves (ROCs) were also estimated for both stages. All analyses were performed using the SAS version 9.1 servipack 4 software package.

Results

M-CHAT Cultural Adaptation

Questionnaires from 622 children and over 40 phone FUIs were included in this pilot phase. Some modifications were

made to items 3, 5 and 23 after the short pilot study and prior to the validity study, in order to overcome several cultural differences mainly linked to the use of different toys in Spain (including new examples), and Spanish colloquialisms. Items 5, 8 and 17 were also reworded due to parents' lack of understanding or misinterpretation, though only minor cultural nuances were introduced. A final version of the M-CHAT in Spanish, along with its 23 items in flow chart format for the corresponding phone FUIs, were then approved by the M-CHAT's original authors.

Stage 1. Validity Study

Distributions by gender, age and case source for Stages 1 and 2 are shown in Table 1.

A total of 86 children underwent diagnostic assessment, and 23 children were finally identified in Stage 1 as having some type of ASD (Fig. 2). Of these 23 ASD cases, 19 belonged to the high-risk sample.

Table 2 shows the percentage of failed items for each of the above groups, namely: (a) children needing no follow-up; (b) children whose parents were administered the phone FUI, but whose final outcome was "pass"; (c) children who were referred for diagnostic evaluation but were finally diagnosed with a developmental disorder other than ASD; and finally, (d) children diagnosed with ASD.

After confirming the phone FUI, nearly half the M-CHAT items (items 6, 7, 13, 14, 15, 17, 19, 20, 21 and 23) were independently able to identify more than 55% of ASD children. Specifically, items 7, 17 and 21 accounted for the highest percentages (over 70%). Conversely, items 3, 4, and 16 were failed by zero to 9% of ASD cases while items 8, 11, and 22 were failed by 17–35% of ASD cases.

It should also be noted that, in most cases, the relatively higher failure rates for items 11, 18, 22 and 23 in the "No follow-up group" were due to misunderstanding as to the meaning of these items.

Table 1 Demographic data

		Vaccination at age 18 months	Check-up at age 24 months	Diagnosis/EI Units-High-risk sample ^a
Stage 1 (N = 2,480)	Male	754	500	49
	Female	728	435	14
	Total	1,482	935	63
Stage 2 (N = 2,055)	Male	715	391	
	Female	624	325	
	Total	1,339	716	
Age range (in months)		16–20	21–30	30–42

^a Included in Stage 2 only

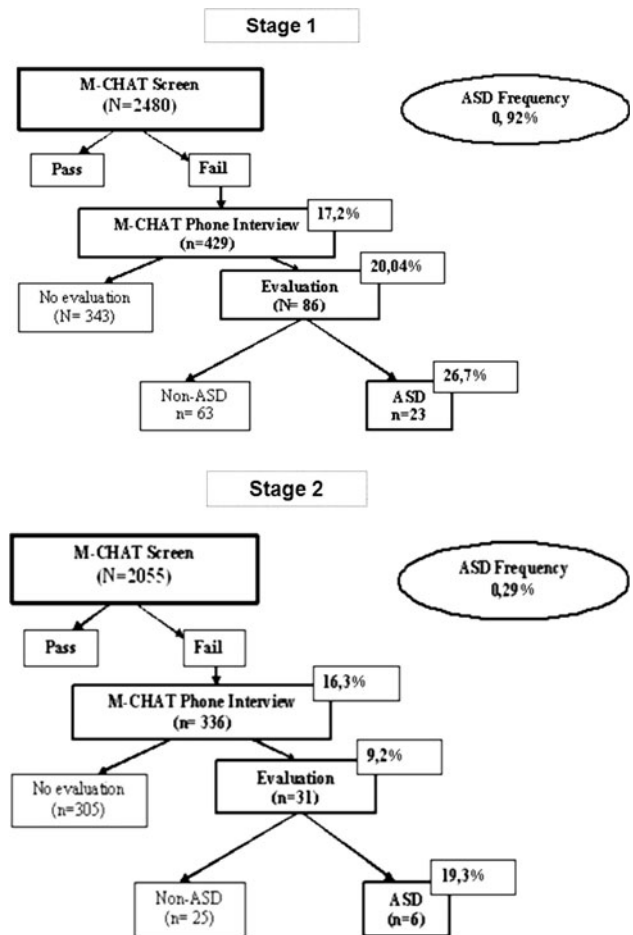


Fig. 2 Flowchart of validity and reliability study results in Spain (Stages 1 and 2)

M-CHAT's estimated properties for detecting ASD cases showed a sensitivity (Sen) of 1; a specificity (Sp) of 0.98; a positive predictive value (PPV) of 0.35; and a negative predictive value (NPV) of 1. The ROC curve also showed that the area under the curve (AUC) was close on 100%. (C = 0.9950).

In terms of the number of failed items, canonical discriminant analysis showed an almost perfect separation between TD and children with ASD, but not such a clear separation between children with ASD and those with DD (Fig. 3).

The logistic regression analysis showed that the estimated probability of being an ASD case only rose exponentially after 5 failed items (Fig. 4).

Stage 2. Reliability Study

All children were recruited exclusively through the population-based screening program, with no referral cases being added to this stage. The main difference between the results of this study and those of Stage 1 was the lower

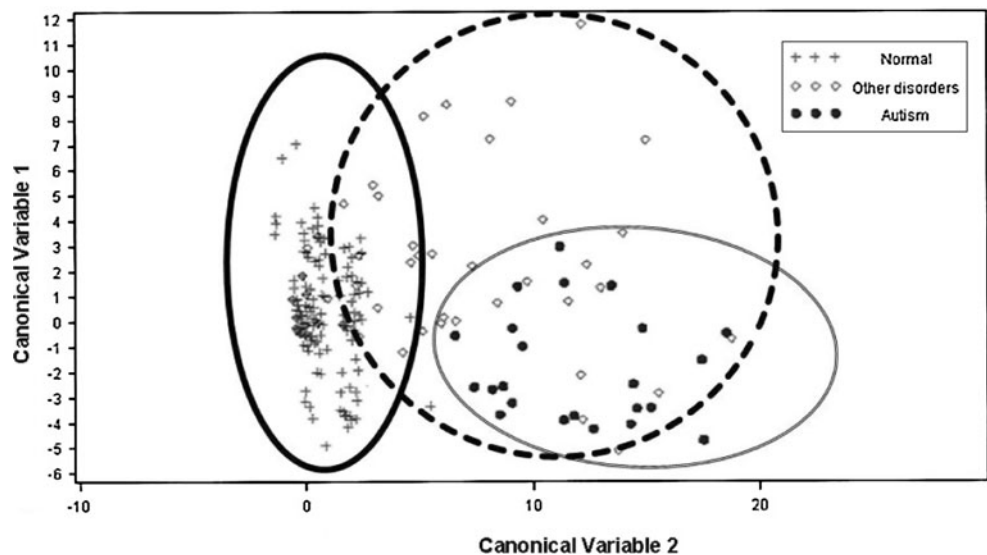
Table 2 Percentage of children in each group who failed each item

Items	M-CHAT item (English key words)	No follow-up (n = 2,032)	OK on follow-up interview (FUI) (n = 362)	Evaluated NON-ASD (n = 63)*	Evaluated ASD (n = 23)
1	Enjoys being swung	0.84	0.83	3.17	8.7
2	Interest in other children	0.25	0.55	23.81	52.2
3	Climbs up stairs	0.34	0.83	6.35	0
4	Enjoys peek-a-boo	0.15	0.28	6.35	8.7
5	Pretend play	1.13	1.10	28.57	47.8
6	Imperative pointing	1.23	1.38	22.22	60.9
7	Declarative pointing	0.54	2.49	28.57	82.6
8	Functional play	6.25	6.63	38.10	34.8
9	Brings objects to show	0.34	0.28	28.57	56.5
10	Eye contact	2.85	1.10	95.24	17.4
11	Oversensitive to noise	12.25	4.70	20.63	17.4
12	Responds to smile	0.34	0	3.17	13
13	Imitation of action	1.03	0.55	20.63	56.5
14	Responds to name	0.54	1.10	15.87	60.9
15	Point following	0.20	0.83	22.22	60.9
16	Walking	1.03	1.10	14.29	4.4
17	Gaze-following	5.61	6.35	44.44	73.9
18	Unusual finger movement	15.75	3.87	19.05	34.8
19	Gaining parent’s attention	2.56	1.38	36.51	69.6
20	Wondering hearing	0.84	0.83	25.40	56.5
21	Understands what is said	5.76	0.55	33.33	73.9
22	Stares at nothing	13.7	2.76	25.40	34.4
23	Social reference	13.7	7.73	26.98	65.2

“n” of groups in this table is different from Fig. 2 due to the different categories used (e.g., M-CHAT screen versus no follow-up; phone interview versus OK on phone screen)

* 4 subjects evaluated as “Typical Development”, were included

Fig. 3 M-CHAT scatter plot using canonical discriminant analysis



estimated PPV (equal to 0.19), likely due to the low ASD frequency observed in this study population (2.9/1,000) (Table 3). Nevertheless, the estimated probability of being

an ASD case was similar to that obtained in Stage 1, and the probability of being an ASD case also rose after 5 failed items. Differences in 95% confidence intervals showed

Fig. 4 M-CHAT response operating curves (ROC) and ASD estimate probability. Logistic regression analysis using PROC LOGISTICS (SAS)

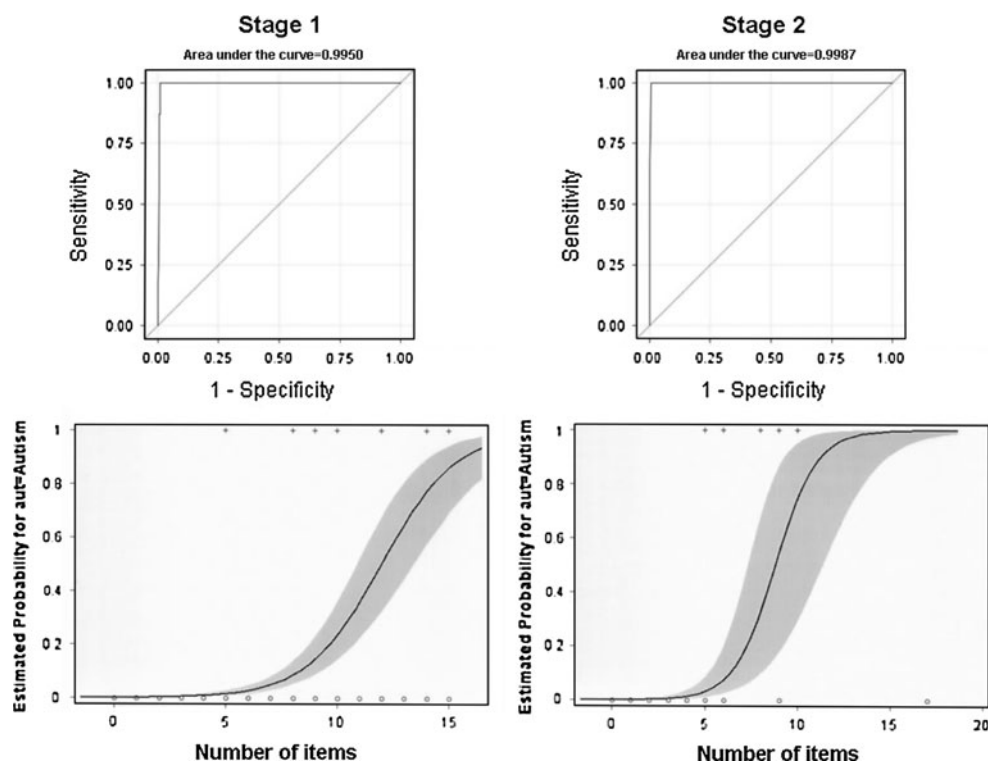


Table 3 Comparison of M-CHAT validity properties (Spain and original study)

	M-CHAT (Stage 1)		M-CHAT (Stage 2)		M-CHAT (Robins 2001)*	
	ASD	NON-ASD	ASD	NON-ASD	ASD	NON-ASD
Fail	23	43	6	25	33	19
Pass	0	2,414**	0	2,024	5	1,236
Validity properties [95% Confidence interval (AUC)]	Prevalence = 0.92%		Prevalence = 0.29%		Prevalence = 2.66%	
	Sen = 1		Sen = 1		Sen = 0.87 [0.76–0.98]	
	Sp = 0.98 [0.98–0.99]		Sp = 0.98 [0.98–0.99]		Sp = 0.99 [0.98–0.99]	
	PPV = 0.35 [0.23–0.46]		PPV = 0.19 [0.05–0.33]		PPV = 0.80 [0.5–0.77]	
	NPV = 1		NPV = 1		NPV = 0.99 [0.99–1]	

All values shown in this table are estimated values because follow-up of negative cases was only partially achieved

* ASD versus non-ASD data have not been published by Robins in this format but have been estimated from the validity properties and other available data from different published papers

** 20 subjects evaluated with developmental disorders

broader bandwidth than that of Stage 1, due to the low number of ASD cases diagnosed in this study.

Discussion

This is the first study to validate the M-CHAT in Spain and our paper outlines the three phases of the validation process: translation; cultural adaptation; validity and reliability analysis. Consequently, this should be regarded as the first official Spanish version of the M-CHAT to be applied in Spain. The previous Spanish version was translated for

Latin American countries and was not valid for Spain due to certain vocabulary nuances and cultural differences. As a result there are now two Spanish language versions of the M-CHAT available to be downloaded from www.mchat-screen.com, i.e., one under the name “Spanish-Western-Hemisphere Version” (for which no validity study has yet been located in the literature), and the version currently presented in this study under the name, “Spanish-Spain version”.

Our results yielded sensitivity and specificity estimates similar to the values of the original M-CHAT validity study (Robins et al. 2001). Moreover, items that best

discriminated between ASD and non-ASD cases generally coincided with items identified as critical in the original M-CHAT study (Robins et al. 2001). The only major difference resided in the positive predictive values, owing to the fact that this parameter relies on prevalence rather than internal M-CHAT properties. This difference can be clearly justified by the frequency of ASD cases observed in our study (1 case in 108 children in Stage 1, and 1 case in 300 children in Stage 2), as opposed to other M-CHAT studies with higher prevalence rates, e.g., 1 case in 33 children (30/1,000) (Kleinman et al. 2007), and 1 case in 50 children (20.1/1,000) (Pandey et al. 2008). Nevertheless, this limitation would affect the PPV but not the remaining the properties of the questionnaire, since sensitivity and specificity are intrinsic test properties regardless of the ASD prevalence observed. Furthermore, the area under the ROC curve was close to 100% in both stages (Fig. 4), which means that the M-CHAT is a good instrument for detecting all children with ASD. Although it was not our stated intention to seek a new cut-off point for the M-CHAT, the estimated probability distribution of being an ASD case when logistic regression analysis is used, indicates that some false positives cases would be reduced if the M-CHAT were only deemed to be positive after 5 failed items. While this criterion is not a real cut-off point, it nevertheless supports the idea that the certainty of detecting an ASD case could be increased among children from a population-based setting, only after a certain number of items were failed. In order for us to be able to estimate a new cut-off point, we would have had to take into account, not only the different slopes of the ROCs, but also the specific benefits of treating a true ASD case, the costs of treating a false positive, and the different ASD prevalences in each setting (Metz 1978).

Our study also highlighted the utility of using specific flowchart forms for each item during the phone FUI. In our study, this M-CHAT procedure showed that, from the 429 children with a failed phone-FUI result, only 86 screened positive.

Some other limitations of M-CHAT administration were detected in this study, such as some items being misunderstood, and (time-consuming) difficulties in locating families by telephone in order to administer the phone FUI.

To complete the M-CHAT validity process, cases that screened negative need to be followed up to ensure that non-ASD cases were not misclassified as ASD or vice versa. These types of studies are very expensive because the majority of negative cases in a population will be real negative cases, and the probability of detecting an ASD case among them will be very low. Nevertheless, this type of follow-up is intended to be conducted in a population sample drawn from the total of negative cases detected in Stage 2. In addition, a monitoring system based on early intervention units in the education and welfare system, was

established in 2006 for all children living in the area targeted by Stage 1. Under this system, all positive cases underwent a follow-up assessment after 6 and 12 months, thereby rendering it highly unlikely that autism would not yet have been identified.

Since then, every child over 24 months of age identified as a possible ASD case by such units has been immediately referred to the Salamanca University ASD diagnostic unit which was tasked with ASD diagnosis during this stage. To date, no referred case has been classed as negative by the M-CHAT at an earlier age. It is clear, moreover, that this type of screening study is really based on a cross-sectional study design, and that all M-CHAT-property considerations have to be made bearing this limitation in mind (the sample followed up is not regularly included). The only well-known follow-up study of M-CHAT-screened children was published by Kleinman et al. (2007), several years after the first validity study.

The use of standardized-tool-based procedures for early ASD detection is a widely recognized recommendation, since, by reducing diagnostic delay, these instruments facilitate early intervention. Diagnostic delay is a problem that has, not only been documented in Spain (GETEA 2003), but is also acknowledged as being one of the factors that reduces expectations of a favorable prognosis. By supplying evidence that ASD cases could be detected by the SNHS at approximately 24 months of age, the M-CHAT data in our study strengthen the hope that the problem of diagnostic delay can be overcome in Spain. Inasmuch as it facilitates early detection and intervention, this instrument could also be of great value in reducing parents' anxiety when possible concerns need to be confirmed or clarified. Nevertheless, we are still far from demonstrating that the application of the M-CHAT in a population-based setting would be effective from a cost-effective stance.

M-CHAT administration starts with pediatrician and/or nurses having to introduce the characteristics of this screening program to parents. Parents, pediatricians and nurses alike displayed great interest in systematically checking toddlers' communicative and social development levels, which is undoubtedly a sign of significant progress here in Spain.

Apart from the fact that the health services are responding to a social demand, this study has identified a clear need for coordination between the health services and ASD-specific early intervention units in Spain. The SNHS is substantially different from other national and local health services in which the M-CHAT has been studied (Atlanta, Canada, Middle Eastern countries, etc.) (Eaves et al. 2006; Robins et al. 2001; Seif et al. 2008). It covers 100% of the population of all income levels and employment status, and also cares for legal immigrants. These

differences should be taken into account by the health authorities, when it comes to assessing all the necessary requirements for implementing these types of standardized population-based screening programs.

Conclusions

Spain now has a specific ASD-case detection tool available, which is easy to use in the pediatrician's office. The major challenge to be faced now, however, is to demonstrate that an ASD early-detection program using M-CHAT within a population-based framework is cost-effective and can be widely implemented in the SNHS. If this can be done, such a program's ability to detect cases with any social and/or communicative development delay could then be regarded as an added value.

Efficiency could be also increased if pediatricians and nurses could play a dual role in the procedure of implementing the M-CHAT in Spain, which would entail, not only distributing the questionnaire during office visits, but also confirming failed items using the phone FUI flow-charts. This has also been suggested in other studies (Robins 2008). Since the literature has reported that exclusive reliance by pediatricians on clinical criteria has yielded very poor results in terms of the number of high-risk cases detected, (Gillberg 1990; Rapin 1995), using the M-CHAT, professionals' knowledge and skills in this area would thus be further reinforced.

Acknowledgments This research was supported by grants: SAN/1015/2005 from the Castile-León Regional Health Authority (*Consejería de Sanidad de Castilla y León*); TPY1289/05 (Carlos III Institute of Health); SEJ2006-11536 from the Spanish Ministry of Science & Innovation (*Ministerio de Ciencia e Innovación*); SA043A06 from the Castile-León Regional Educational (*Consejería de Educación de la Junta de Castilla y León*), Castile-León Regional Autism Association (*Federación de Autismo de Castilla-León*); EDU/894/2009 Excellence Group of Research (GR197), Junta de Castilla y León (Spain) and Spanish Royal Society for the Handicapped (*Real Patronato de Discapacidad de España*). We should like to thank all those nurses and pediatricians in the Spanish National Health System who participated in the screening program in Salamanca, Zamora and Madrid Health Area 1. Needless to say, we should also like to express our heartfelt gratitude to all the early-intervention providers, the children, and their families, without whom this study would not have been possible.

References

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: American Psychiatric Association.
- Baird, G., Charman, T., Baron-Cohen, S., Cox, A., Swettenham, J., Wheelwright, S., et al. (2000). A screening instrument for autism at 18 months of age: A 6-year follow-up study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 39(6), 694–702.
- Baron-Cohen, S., Allen, J., & Gillberg, C. (1992). Can autism be detected at 18 months? The needle, the haystack, and the CHAT. *British Journal of Psychiatry*, 161, 839–843.
- Baron-Cohen, S., Wheelwright, S., Cox, A., Baird, G., Charman, T., Swettenham, J., et al. (2000). Early identification of autism by the checklist for autism in toddlers (CHAT). *Journal of the Royal Society of Medicine*, 93(10), 521–525.
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25, 3186–3191.
- Canal-Bedia, R. (2001). Referencia Conjunta y Autismo. In A. Riviere & J. Martos (Eds.), *Autismo: Comprensión y explicación actual* (pp. 57–72). Madrid: APNA_IMSERSO.
- Corsello, C. (2005). Early intervention in autism. *Infants and Young Children*, 18, 74–85.
- Dawson, G., Meltzoff, A. N., Osterling, J., & Rinaldi, J. (1998). Neuropsychological correlates of early symptoms of autism. *Child Development*, 69(5), 1276–1285.
- Dawson, G., & Osterling, J. (1997). Early intervention in autism. The effectiveness of early interventions. In M. J. Guralnick (Ed.), *The effectiveness of early intervention* (pp. 307–326). Baltimore: Brookes.
- Díez-Cuervo, A., Muñoz-Yunta, J. A., Fuentes-Biggi, J., Canal-Bedia, R., Idiazábal-Aletxa, M. A., Ferrari-Arroyo, M. J., et al. (2005). Guía de buena práctica para el diagnóstico de los trastornos del espectro autista. *Revista Neurologica*, 41, 299–310.
- Eaves, L. C., Wingert, H. D., Ho, H. H., & Mickelson, E. C. (2006). Screening for autism spectrum disorders with the social communication questionnaire. *Journal of Developmental and Behavioral Pediatrics*, 27, S95–S103.
- Filipek, P. A., Accardo, P. J., Ashwal, S., Baranek, G. T., Cook, E. H., Jr., Dawson, G., et al. (2000). Practice parameter: Screening and diagnosis of autism: Report of the Quality Standards Subcommittee of the American Academy of Neurology and the Child Neurology Society. *Neurology*, 55, 468–479.
- GETEA. (2003). Análisis del Diagnóstico de afectados de Autismo y Planificación del Registro Español de TEA. http://iier.isciii.es/autismo/pdf/aut_if04.pdf.
- Gillberg, C. (1990). Infantile autism: Diagnosis and treatment. *Acta Psychiatrica Scandinavica*, 81, 209–215.
- Howlin, P., Mawhood, L., & Rutter, M. (2000). Autism and developmental receptive language disorder—A follow-up comparison in early adult life. II: Social, behavioural, and psychiatric outcomes. *The Journal of Child Psychology and Psychiatry*, 41(5), 561–578.
- Inada, N., Koyama, T., Inokuchi, E., Kuroda, M., & Kamio, Y. (2010). Reliability and validity of the Japanese version of the modified checklist for autism in toddlers (M-CHAT). *Research in Autism Spectrum Disorders*. doi:10.1016/j.rasd.2010.04.016.
- International Test Commission. (2010). Test adaptation guidelines. Retrieved June 20, 2010, from <http://www.intestcom.org/>.
- Johnson, C. P., Myers, S. M., & American Academy of Pediatrics Council on Children With Disabilities. (2007). Identification and evaluation of children with autism spectrum disorders. *Pediatrics*, 120(5), 1183–1215.
- Kasari, C., Freeman, S., & Paparella, T. (2006). Joint attention and symbolic play in young children with autism: A randomized controlled intervention study. *The Journal of Child Psychology and Psychiatry*, 47(6), 611–620.
- Kleinman, J. M., Robins, D. L., Ventola, P. E., Pandey, J., Boorstein, H. C., Esser, E. L., et al. (2007). The modified checklist for autism in toddlers: A follow-up study investigating the early detection of autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 38, 827–839.
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., et al. (2000). The autism diagnostic observation

- schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism of Developmental Disorders*, 30, 205–223.
- Mawle, E., & Griffiths, P. (2006). Screening for autism in pre-school children in primary care: Systematic review of English Language tools. *International Journal of Nursing Studies*, 43(5), 623–636.
- Metz, C. E. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine*, 8, 283–298.
- Pandey, J., Verbalis, A., Robins, D. L., Boorstein, H., Klin, A. M., Babitz, T., et al. (2008). Screening for autism in older and younger toddlers with the modified checklist for autism in toddlers. *Autism*, 12, 513–535.
- Rapin, I. (1995). Physicians' testing of children with developmental disabilities. *Journal of Child Neurology*, 10(Suppl), 11–15.
- Robins, D. L. (2008). Screening for autism spectrum disorders in primary care settings. *Autism*, 12(5), 537–556.
- Robins, D. L., & Dumont-Mathieu, T. M. (2006). Early screening for autism spectrum disorders: Update on the modified checklist for autism in toddlers and other measures. *Journal of Developmental and Behavioral Pediatrics*, 27(2 Suppl), S111–S119.
- 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 of Developmental Disorders*, 31(2), 131–144.
- Roid, G. H., & Sampers, J. (2004). *Merrill-Palmer Developmental Scale, Revised manual*. Wood Dale, IL: Stoelting.
- Rogers, S. J., & Vismara, L. A. (2008). Evidence-based comprehensive treatments for early autism. *Journal of Clinical Child and Adolescent Psychology*, 37(1), 8–38.
- Seif Eldin, A., Habib, D., Noufal, A., Farrag, S., Bazaid, K., Al-Sharbati, M., et al. (2008). Use of M-CHAT for a multinational screening of young children with autism in the Arab countries. *International Review of Psychiatry*, 20, 281–289.
- Sparrow, S., Balla, D., & Cicchetti, D. (1984). *Vineland adaptive behavior scales: Interview edition*. Circle Pines, MN: American Guidance Services.
- Tonge, B., Brereton, A., Kiomall, M., Mackinnon, A., King, N., & Rinehart, N. (2006). Effects on parental mental health of an education and skills training program for parents of young children with autism: A randomized controlled trial. *Journal of American Academy of Child and Adolescent Psychiatry*, 45(5), 561–569.
- Ventola, P., Kleinman, J., Pandey, J., Wilson, L., Esser, E., Boorstein, H., et al. (2007). Differentiating between autism spectrum disorders and other developmental disabilities in children who failed a screening instrument for ASD. *Journal of Autism of Developmental Disorders*, 37, 425–436.
- Wetherby, A. M., Woods, J., Allen, L., Cleary, J., Dickinson, H., & Lord, C. (2004). Early indicators of autism spectrum disorders in the second year of life. *Journal of Autism of Developmental Disorders*, 34(5), 473–493.