



Spanish Cultural Validation of the Modified Checklist for Autism in Toddlers, Revised

María Magán-Maganto¹ · Ricardo Canal-Bedia¹ · Aránzazu Hernández-Fabián² · Álvaro Bejarano-Martín¹ · Clara J. Fernández-Álvarez¹ · María Martínez-Velarte³ · María V. Martín-Cilleros¹ · Noelia Flores-Robaina⁴ · Herbert Roeyers⁵ · Manuel Posada de la Paz⁶

Published online: 17 October 2018

© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

The Modified Checklist for Autism in Toddlers–revised/follow-up (M-CHAT-R/F) was developed to reduce the number of cases requiring telephone verification. The aim of this study was to validate a Spanish version of the M-CHAT-R/F in the Spanish public health system. The M-CHAT-R/F was translated, culturally adapted, and then administered to 6625 children. Of the 39 positive screening cases, 15 children were diagnosed with autism spectrum disorder (ASD) and 24 with non-ASD disorders or delays. The sensitivity was 0.79 and specificity of 0.99. Positive and negative predictive values were 0.39 and 0.99, respectively. These results are similar to the English equivalent, though observed prevalence was lower. This study supports Spanish National Health System policy makers to consider a universal ASD screening program.

Keywords Autism · Screening tools · Early detection · Cross-cultural validation · M-CHAT-R/F

Autism spectrum disorders (ASD) are neurodevelopmental disorders mainly characterized by difficulties associated to communication and social interaction abilities, accompanied

by restricted interests and repetitive behaviours that significantly affect the daily life activities (APA 2013). The prevalence of ASD has increased in the last decade, with current rates of 1/59 according to the Centers for Disease Control and Prevention (CDC 2018; Baio et al. 2018). European studies report similar rates in Ireland with a prevalence of 1/87 (Boilson et al. 2016) and 10 year old age-specific prevalence estimates in 1999–2001 in Finland (1/167), Sweden (1/147) and Denmark (1/91) have increased since 1990–1992 in these countries by 96%, 354%, and 175%, respectively (Atladdottir et al. 2015).

This continual rise in prevalence highlights the impact on public health and repercussions for National Health Service (NHS) and Social Systems. Improvements in the early

We confirm that this work is original and has not been published elsewhere, nor is it currently under consideration for publication elsewhere. Preliminary data from this study has been presented in the XVIII National Congress AETAPI (Spanish Association for Autism Professionals), León, Spain, in November 2016, in the “XI Autism-Europe International Congress 2016” Edinburgh, UK, in September 2016, and in the “International Meeting for Autism Research (IMFAR)”, Baltimore Convention Center, Maryland, USA in May 2016.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10803-018-3777-5>) contains supplementary material, which is available to authorized users.

✉ Ricardo Canal-Bedia
rcanal@usal.es

¹ Centro de Atención Integral al Autismo (INFOAUTISMO), Faculty of Education, University Institute of Community Integration (INICO), University of Salamanca, Salamanca, Spain

² Hospital Universitario de Salamanca, University of Salamanca, Salamanca, Spain

³ Servicio de Psiquiatría del Complejo Asistencial de Zamora, Zamora, Spain

⁴ Faculty of Psychology, University Institute of Community Integration (INICO), University of Salamanca, Salamanca, Spain

⁵ Department of Experimental Clinical and Health Psychology, Ghent University, Ghent, Belgium

⁶ Institute of Rare Diseases Research & CIBERER, Instituto de Salud Carlos III, Madrid, Spain

detection and intervention will become increasingly important to reduce the ASD's burden and improve the quality of life of these people (Sánchez-Valle et al. 2008). A better awareness and understanding of early flags are needed to improve screening programs not only in early ages, but also later, such as in preschool age (Robins 2016). The American Academy of Pediatrics recommends the use of valid and reliable screening tools at the age of 18 and 24 months in primary care settings (Johnson and Myers 2007), although a recent report issued by the US Preventative Services Task Force (2016; Siu et al. 2016) concluded that there is insufficient evidence to recommend universal screening for ASD. This report has promoted an international debate on the two following key issues: the first question is whether current screening tools are sufficient to accurately identify ASD cases among young children in the general population; and the second is whether interventions exist that are effective for ASD cases identified by early screening (Mandell and Mandy 2015). The report recognizes that screening tools such as Modified Checklist for Autism in Toddlers (M-CHAT; Robins et al. 2001) and Modified Checklist for Autism in Toddlers–Revised/Follow-up (M-CHAT-R/F; Robins et al. 2014) are suitable for detecting ASD in children aged 16–30 months, stating in turn that good quality randomized trials based on intensive behavioural and developmental treatments have shown that early intervention for young children with ASD is effective.

The M-CHAT-R/F (Robins et al. 2014) was developed to reduce the number of cases requiring a follow-up interview (FUI), without decreasing the effectiveness of the tool. The interviews can be a burden on the screening process, as well as a step with high dropout of participants, so reducing the proportion of cases needing this interview is useful. The scoring was also simplified to ease the implementation of ASD population screening.

Even though M-CHAT questionnaires have been translated into many languages, it is important to highlight the fact that cultural differences exist in the different contexts, and proper adaptations should be made in order to cater to the needs of each culture and apply screening tools with accuracy (Soto et al. 2015).

This study aims to validate the Spanish cultural version of the M-CHAT-R/F, a revised and improved version of the previous M-CHAT, consisting of 20 items and a new scoring system.

Methods

Catchment Area

This study was conducted in three provinces in Northwest Spain, namely Salamanca, Zamora and Valladolid, from

April 2014 to December 2016 (data from Valladolid was collected since June 2016). There were 11,184 births in the regions from April 2014 to 2016 (average total population 1,046,050) (INE 2017).

Sample Characteristics

From April 2014 to December 2016, an estimated 9552 children were 2 years old in the areas of Salamanca and Zamora and 2158 children in Valladolid from June to December 2016 (INE 2017), approximating the number of children that should have attended the routine 18 and 24 months “Well Baby Check-up Program” screenings (for further references about description of the Spanish National Health System, see Canal-Bedia et al. 2011). This means that a total of 11,710 children were estimated as eligible for screening. 6625 children were screened by the program, giving a participation rate of 56.6%. Written informed consent to participate in this study was obtained from the parents of all children at the pediatrician centers. Furthermore, for the families of the positive cases that came for further evaluation in the research center, local Ethics Committee approval and another informed consent form for diagnostic assessment were required before the assessment took place.

A total of 93 primary care health centers, 134 pediatricians and 119 pediatric nurses from the above-mentioned regions were involved in the study.

Participants were excluded from the study if they did not complete the FUI or evaluation due to problems of communication with the families (unreachable, $n = 35$; rejected invitation for diagnostic evaluation, $n = 3$; or did not attend evaluation by the time this study concluded, $n = 2$).

Descriptive characteristics of the sample can be found in Table 1.

Instruments

Screening Tool

Children were screened using a translated version of the M-CHAT-R/F (Robins et al. 2014) to Spanish. The original cut-off approach from Robins et al. 2014 was adopted. Three risk levels are differentiated according to the number

Table 1 Samples characteristics

| | Age range 14–22 months | Age range 23–36 months |
|----------------------|------------------------|------------------------|
| Mean age (SD) | 18.22 (0.72) | 24.47 (1.23) |
| Gender male (female) | 1805 (1,621) | 1589 (1500) |
| Screen positive | 31 | 19 |
| Completed evaluation | 19 | 16 |

of failed items obtained in the questionnaire. Low risk is defined by a score from 0 to 2. Medium risk is defined by scores between 3 and 7 failed items. In this case, it would be necessary to verify the questionnaire, for which the FUI interview is administered to obtain additional information about the items at risk. If after the interview there are still three or more failed items, the child will be referred for further assessment. High-risk children are those who fail 8 or more items. In this case, the children are directly referred for diagnostic evaluation to determine the need for early intervention.

Evaluation Instruments

Using DSM-5 (APA 2013) criteria for ASD as the gold standard, a standardized protocol containing the following instruments was administered to support the diagnosis: clinical history; Merrill-Palmer Revised Scales (Roid and Sampers 2004); Leiter (Roid and Miller 1997); Vineland Scales (Sparrow et al. 2005, 1984); the Autism Diagnostic Observation Schedule-Generic (ADOS-G) module 1 (Lord et al. 2000) and ADOS-2 module T and 1 (Lord et al. 2012). Evaluations were performed by trained and experienced professionals, who applied, scored, and interpreted results from the different instruments to ensure the accuracy of the data and avoid the possible influence of extraneous variables on the results.

Procedure

Screening Procedure

The M-CHAT-R/F was translated and a cross-cultural adaptation was performed. The resulting version was forward and back translated by two bilingual persons and compared with the original version. When translating the questionnaire and FUI, cultural adaptations were particularly needed for the examples given to parents. This was done to make the examples clearer and more understandable in the Spanish context. Question 9 (Does your child like climbing on things? For example, furniture, playground equipment, or stairs) was slightly changed to represent specific furniture (chairs) or playground equipment (slides) that are easier for parents to understand when filling out the questionnaire. In Question 20 (Does your child like movement activities? For example, being swung or bounced on your knee), there is a children's game in Spain that is similar to the action of "being bounced on your knee" where the child pretends your knees are a horse and tries to ride them. In this case, the name of the activity in Spanish is used to describe the action. Experts assessing child development were asked to evaluate the understanding of the questions in the Spanish version or any trouble related to translation, with no problems found.

Furthermore, a random sample of 30 families were asked to complete the questionnaire and report any misunderstanding that they had replying to it. Any suggestions were reported. During all the process, feedback with the original authors was made, with the final version approved by them.

A series of training sessions for health professionals (pediatricians and pediatric nurses together) were organized to explain the new implementations of the screening program in Salamanca and Zamora where the screening has been running since 2005 (see Canal-Bedia et al. 2011) with the M-CHAT. Changes were introduced relating to the use of a new data base system to collect data from the questionnaires (<https://riviere.usal.es/mchatr/>), as well as changes related to the reduction of items in the questionnaire and scoring the M-CHAT-R/F. Additional training courses were given to the health professionals in Valladolid on ASD characteristics, screening procedures and some guidelines for high-risk referrals, thus introducing the screening program. In this area the full process has been implemented within the daily pediatrician and clinical practice since June of 2016.

In the two provinces where the program had already been utilized for more than 10 years, the same procedure was followed as that used in Canal-Bedia et al. (2011), sending the M-CHAT-R questionnaires in paper to the research unit where the team from the research unit entered the questionnaire data into the database system; performed the FUI; and called the family to attend further assessment if screening positive.

The pediatricians and pediatric nurses from Valladolid area implemented the FUI from their own facilities. Once the M-CHAT-R questionnaire was submitted to the database (normally a few days after completing the questionnaire), they would call the family to check the responses to the failed items and, if two or more items are still failed, they would refer the families to the specialized psychiatric units of the NHS.

Evaluation Procedure

Assessments for diagnosis were performed in the central research unit of the University of Salamanca for the positive screen cases detected in the provinces of Salamanca and Zamora, however the positive cases from Valladolid were evaluated by the psychiatric units of the NHS. In all the cases, the evaluation protocol followed that outlined in the instrument section and one of every ten positive cases in Valladolid were reviewed by the research team to check concordance. False negative cases are identified via the public early attention centres that provide their services free of charge in the region. The early attention centres refer all cases that they suspect the child has communicative or social developmental problems. These cases are evaluated by the research team for ASD. Cases that are diagnosed with ASD

are considered to be false negatives if they had previously completed the M-CHAT-R questionnaire in the child health visit at 18 or 24 months with two or less failed items. If they had not completed the M-CHAT-R beforehand, then they are not considered false negatives for the questionnaire and they are not taken into consideration for the analyses of this study.

Data Analyses

Analyses were stratified by age range from 14 to 22 months and from 23 to 36 months to estimate the psychometrical properties of the M-CHAT-R/F at different ages and analyze how potential differences could impact the implementation of the tool. Usually the questionnaire was applied when children attended the routine check-up at 18 and 24 months, but in some cases, physicians applied the screening protocol before or after these ordinary visits due to family or professionals' concerns.

Internal consistency measured with Cronbach's α was performed before and after the FUI. Percentages of children who failed each item of the questionnaire by group were compared between age ranges. In addition, Chi square comparisons between ASD and screen negative, and ASD and Other Neurodevelopmental Disorders (OND)/Language Delay (LDe) groups were performed at a level of significance $p \leq .01$ for the different age ranges.

Values of prevalence, sensitivity, specificity, predictive values and likelihood ratios were calculated for the samples and these values were also estimated taking into account the excluded cases (see "Sample Characteristics" section). This was done because the excluded cases were often part of the higher risk (failed screening) groups and ignoring them in the analyses could misrepresent the real data distribution. The excluded cases were assigned to each group assuming the proportions found in each phase of the screening (see Barbaro and Dissanayake 2010).

Lastly, canonical discriminant analysis was performed to analyze the strength of the M-CHAT-R/F when classifying children with or without ASD in a population based setting.

Statistical analyses were performed using SPSS software version 23 IBM.

Results

Sample from 14 to 22 Months

Internal consistency for the M-CHAT-R/F items was below the threshold for adequate (Cronbach's $\alpha = 0.59$), albeit when analyzing items responses after FUI, internal consistency increased to a Cronbach's α value of 0.62.

Throughout the screening program at this age range, 3529 children were screened and 11 of them were identified with

ASD: nine true positive cases (TP) and two false negatives (FN) (see Fig. 1). Of the 106 that screened positive in the questionnaire, nine cases failed eight or more items and 97 required FUI. After FUI, 22 (22.7%) cases screened positive, 54 (55.7%) cases screened negative and 21 (21.6%) cases were not contactable.

Ten cases were found to be false positives (FP), two of which had a diagnosis of language disorder (LD); three had a global developmental disorder (GDD) and another one had a diagnosis of unspecified neurodevelopmental disorder. One case was given a diagnosis of other systemic diseases (diseases that can affect immune system, organs or tissues of the body). Only three cases did not meet criteria for any other Neurodevelopmental Disorder described in the DSM-5, however two of these children presented LDe.

Table 2 shows the percentage of children who failed each item of the M-CHAT-R/F by group before FUI, along with the results of Chi square tests for each item between ASD and OND/LDe and ASD versus Screen Negative groups. In the ASD group, the most frequently failed items are items: 3 (pretend play) failed by 63.6% of the children; 6 (imperative pointing), 7 (declarative pointing), 9 (show object to share) and 19 (social reference) failed by 54.5% of children (see Table 2). In the OND/LDe group, the items failed at a high percentage were item 7 (61.5%) and item 6 (53.8%). The items that most children failed in the Screen Negative group were item 12 (oversensitive to noise, 9.4%) and item 5 (unusual finger movement, 6.9%). Chi square analyses showed significant statistical differences between ASD and Screen Negative groups in all but five of the items (items 4—like climbing-, 11—social smile-, 12—oversensitive to noise-, 14—eye contact-, and 20—enjoys being swung-; see Table 2). In the comparison between ASD and OND/LDe groups, Chi square did not show significant differences.

M-CHAT-R/F's estimated properties for detecting ASD cases, taking into account only the assessed cases, showed a prevalence rate (PR) for the sample of 1/318, sensitivity (Sen) of 0.82; a specificity (Spe) of 0.99 and a positive predictive value (PPV) of 0.47. Psychometrical properties were estimated taking into account the excluded cases, leading to roughly nine ASD cases that would be TP. The psychometrical properties of this data were: PR of 1:176, a Sen of 0.90; a Spe of 0.99; and a PPV of 0.49 (see Table 3).

The discriminant analysis shows that when all variables are combined to discriminate between ASD, OND/LDe and Screen Negative, there is an overlap of the groups (see Fig. 2). However, there are some cases that can be easily identified by visual inspection of Fig. 2 as not typical development, and others that can easily be identified as disorders other than autism.

A 83.1% of variance was explained by Function 1, while a 16.9% was explained by Function 2. Together, 98.8% of the cases were classified correctly.

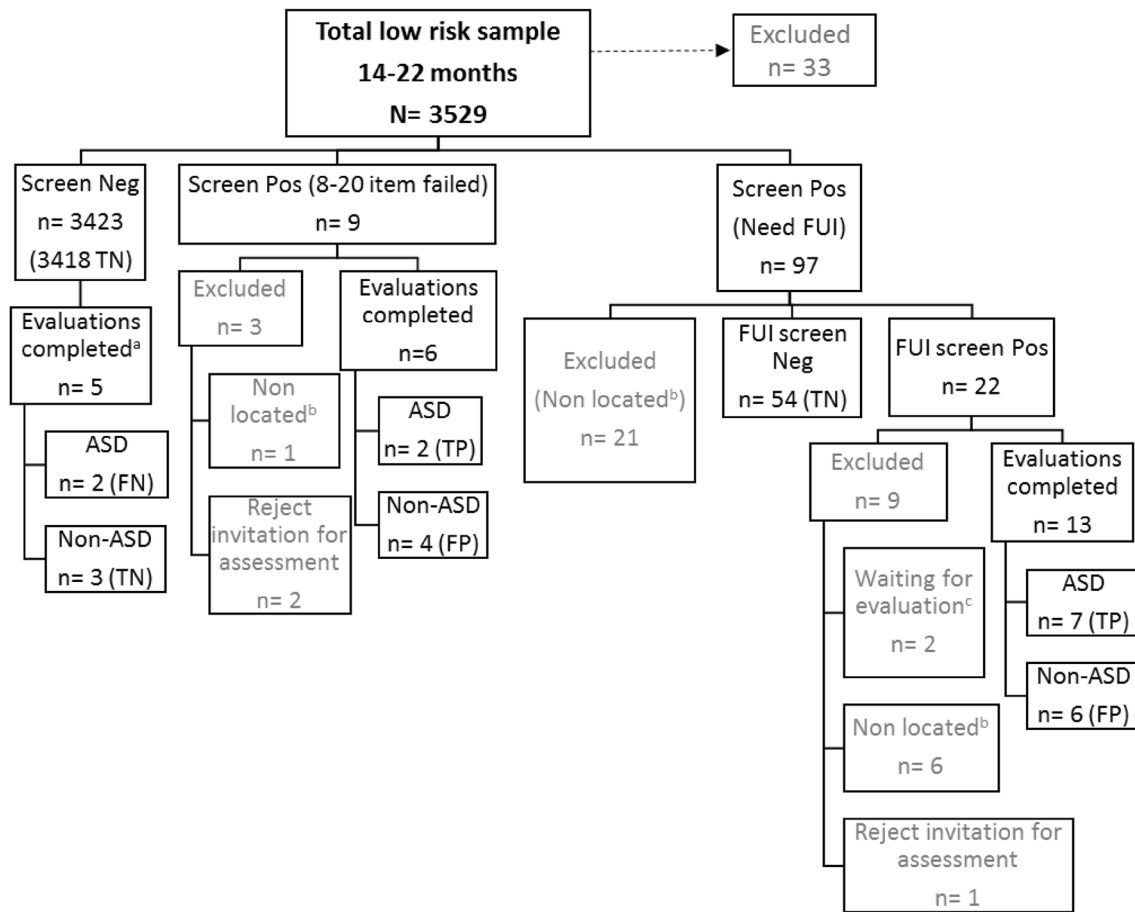


Fig. 1 Flowchart indicating screening results and diagnosis results. ^aEvaluations based on cases detected by the early intervention centers in collaboration with the screening program to better detect false negatives cases. ^bUnable to be contacted for FUI or further evaluations.

^cCases waiting for diagnosis evaluation when data collection was proceeded. *FUI* follow-up interview, *Neg* negative, *Pos* positive, *TP* true positive, *FP* false positive, *TN* true negative, *FN* false negative

Sample from 23 to 36 Months

Internal consistency for the questionnaire was similar to the original values reported in Robins et al. (2014) (Cronbach's $\alpha = 0.63$), increasing to $\alpha = 0.72$ when performing analyses after FUI.

At this age range, 3096 children were screened, eight of them were identified with ASD: six TP cases and two FN (see Fig. 3). Of the 52 that screened positive in the questionnaire, nine cases failed eight or more items and 43 required FUI, of which 10 (23% of those requiring FUI) were positive cases, 25 (58%) cases screened negative and four (9%) cases that could not be contacted. There were 14 FP cases, five of them had a diagnosis of a GDD and another three had a LD. Five of the FP cases did not meet criteria for any of the disorders described as ND, but present LDe. One of the cases had a psychomotor delay (still under diagnostic evaluation).

The items most frequently failed in the ASD group were items 10 (response to name), 16 (gaze following), 18

(understands what is said) and 19 (social reference), failed by 62.5% (before FUI, see Table 2). In the OND/LDe group, the items failed in a high percentage were item 2 (62.5%), 16 (56.3%) and 19 (56.3%). The items that most children failed in the Screen Negative group were items 12 (oversensitive to noise, 10.6%) and 5 (unusual finger movement, 5.3%). Chi square analyses showed significant statistical differences between ASD and Screen Negative groups in all but 5 of the items (items 4—like climbing-, 5—unusual finger movement-, 12—oversensitive to noise-, 13—walking-, and 20—enjoys being swung-; see Table 2). Chi square did not reveal any significant differences when comparing between ASD and OND/LDe groups.

M-CHAT-R/F's estimated properties for detecting ASD cases taking into account only the monitored cases showed a PR of 1/386; a Sen of 0.75; a Spe of 0.99 and a PPV of 0.3. Psychometrical properties were again estimated taking into account the excluded cases, suggesting the number of TP

Table 2 Percentage of children who failed each item of M-CHAT-R/F by group and age range

| Items | % ASD | | % OND/LDe | | Screen Neg ^a (%) | |
|--------------------------------|----------------|---------------|----------------|----------------|-----------------------------|------------------|
| | 14–22 (n = 11) | 23–36 (n = 8) | 14–22 (n = 13) | 23–36 (n = 16) | 14–22 (n = 3472) | 23–36 (n = 3065) |
| 1. Pointing following | 9.1 | 12.5 | 15.4 | 25 | 0.1 | 0.1 |
| 2. Wondering hearing | 27.3 | 25 | 30.8 | 62.5 | 2.7 | 2.3 |
| 3. Pretend play | 63.6 | 37.5 | 30.8 | 25 | 4.6 | 2.5 |
| 4. Like climbing | 0 | 0 | 15.4 | 6.3 | 0.6 | 0.5 |
| 5. Unusual finger movement | 36.4 | 12.5 | 7.7 | 25 | 6.9 | 5.3 |
| 6. Imperative pointing | 54.5 | 25 | 53.8 | 37.5 | 0.5 | 0.8 |
| 7. Declarative pointing | 54.5 | 37.5 | 61.5 | 50 | 1.1 | 0.5 |
| 8. Interest in other children | 27.3 | 50 | 0 | 37.5 | 0.3 | 0.5 |
| 9. Show object to share | 54.5 | 25 | 46.2 | 25 | 0.5 | 0.2 |
| 10. Response to name | 27.3 | 62.5 | 23.1 | 31.1 | 0.5 | 0.2 |
| 11. Social smile | 0 | 12.5 | 0 | 6.3 | 0.2 | 0.4 |
| 12. Oversensitive to noise | 9.1 | 0 | 38.5 | 12.5 | 9.4 | 10.6 |
| 13. Walking | 18.2 | 0 | 38.5 | 12.5 | 1.3 | 0.2 |
| 14. Eye contact | 0 | 12.5 | 23.1 | 37.5 | 0.3 | 0.7 |
| 15. Imitation of action | 18.2 | 37.5 | 0 | 6.3 | 0.1 | 0 |
| 16. Gaze-following | 45.5 | 62.5 | 46.2 | 56.3 | 2.3 | 1.2 |
| 17. Gaining parent's attention | 36.4 | 25 | 38.5 | 31.3 | 2 | 1 |
| 18. Understands what is said | 45.5 | 62.5 | 30.8 | 50 | 1.2 | 0.3 |
| 19. Social reference | 54.5 | 62.5 | 30.8 | 56.3 | 2 | 2 |
| 20. Enjoys being swung | 0 | 0 | 0 | 0 | 0.2 | 0.2 |

Numbers in bold signifies a Chi square comparison between ASD and Screen negative groups of $p \leq .001$

ASD autism spectrum disorder, OND other neurodevelopmental disorders, LDe language delay, Neg negative, 14–22 age range in months, 23–36 age range in months

^aCases that Screen negative but did not have further assessments

Table 3 Psychometric properties of M-CHAT-R/F by age range

| | 14–22 Months (95% CI) | 14–22 Months (with EC) ^c (95% CI) | 23–36 Months (95% CI) | 23–36 Months (with EC) ^c (95% CI) | 14–36 Months (95% CI) | 14–36 Months (with EC) ^c (95% CI) |
|------------------|------------------------|--|-----------------------|--|------------------------|--|
| PR ^a | 31.46 (15.56–58.08) | 56.67 (35.6–89.09) | 25.9 (12.05–53.16) | 29.07 (14.19–57.26) | 28.85 (17.89–45.95) | 43.77 (29.88–63.68) |
| Sen | 0.82 (0.48–0.97) | 0.9 (0.67–0.98) | 0.75 (0.36–0.96) | 0.78 (0.40–0.96) | 0.79 (0.54–0.93) | 0.86 (0.67–0.96) |
| Spe | 0.99 (0.99–0.99) | 0.99 (0.99–0.99) | 0.99 (0.99–0.99) | 0.99 (0.99–0.99) | 0.99 (0.99–0.99) | 0.99 (0.99–0.99) |
| PPV | 0.47 (0.25–0.71) | 0.49 (0.32–0.65) | 0.3 (0.13–0.54) | 0.29 (0.13–0.51) | 0.39 (0.24–0.55) | 0.41 (0.29–0.54) |
| NPV | 0.99 (0.99–0.99) | 0.999 (0.99–0.99) | 0.99 (0.99–0.99) | 0.99 (0.99–0.99) | 0.99 (0.99–0.99) | 0.99 (0.99–0.99) |
| LR+ ^b | 285.14 (144.64–562.09) | 166.22 (103.72–266.38) | 165.05 (85.46–318.76) | 141.24 (78.39–254.48) | 215.99 (136.08–342.80) | 157.95 (110.55–225.68) |
| LR– ^b | 0.18 (0.05–0.64) | 0.10 (0.03–0.37) | 0.25 (0.08–0.83) | 0.22 (0.07–0.76) | 0.21 (0.09–0.51) | 0.14 (0.06–0.35) |

PR prevalence, Sen sensitivity, Spe specificity, PPV positive predictive value, NPV negative predictive value, LR+ positive likelihood ratio, LR– negative likelihood ratio, CI Confidence Intervals, EC excluded cases

^aNote that the prevalence rate has been expressed in function to a rate of 10,000 people

^bConventional

^cPsychometrical properties were estimated based on the total number of referrals, not just the assessed cases

cases would increase by 1. The psychometrical properties of this data were: PR of 1/344, a Sen of 0.78; a Spe of 0.99; and a PPV of 0.29 (see Table 3).

The discriminant analysis shows similar results to the ones described in the age range 14–22 months (see Fig. 4).

Fig. 2 M-CHAT-R/F scatter plot using canonical discriminant analysis at the age range 14–22 months. *ASD* autism spectrum disorder, *OND* other neurodevelopmental disorders, *LDe* language delay, *Neg* negative

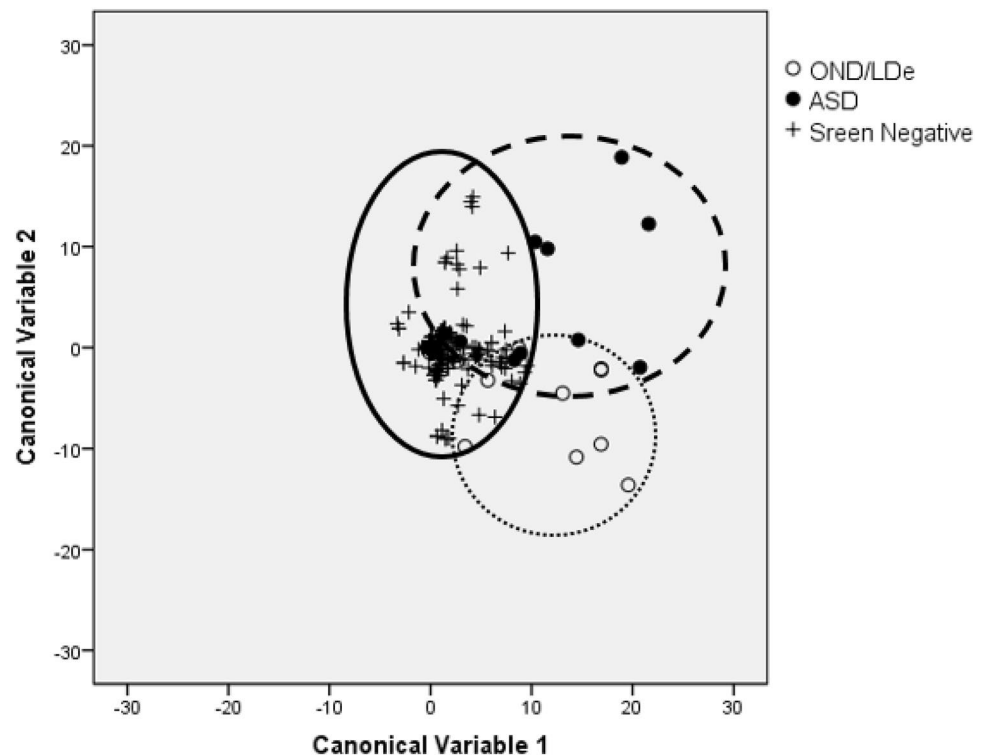
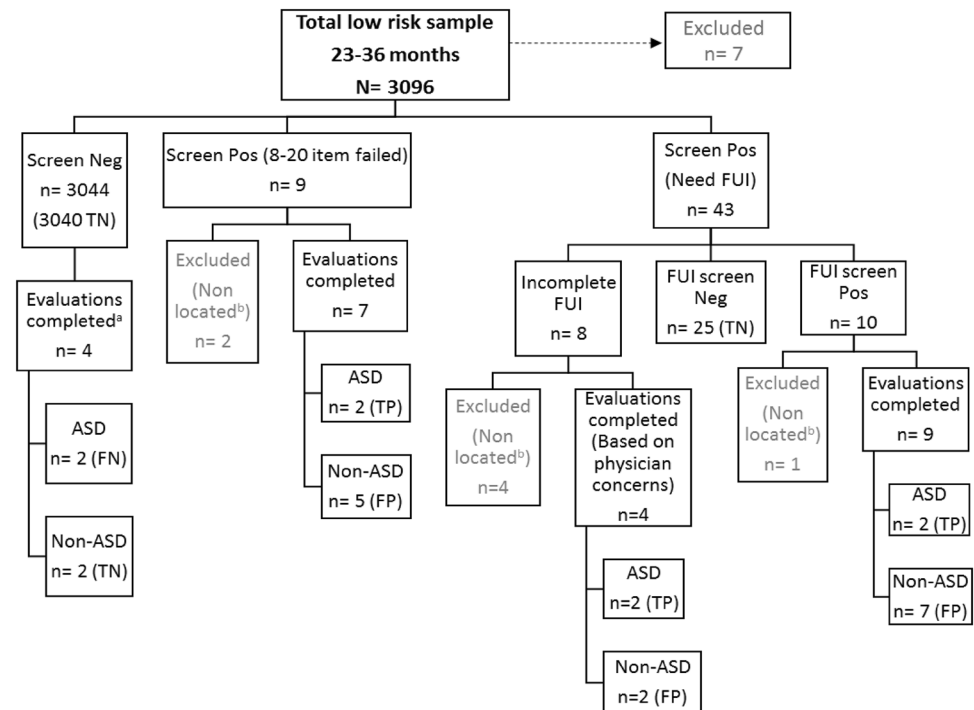


Fig. 3 Flowchart indicating screening results and diagnosis results. ^aEvaluations based on cases detected by the early intervention centers in collaboration with the screening program to better detect false negatives cases. ^bUnable to be contacted for FUI or further evaluations. ^cCases waiting for diagnosis evaluation when data collection was proceeded. *FUI* follow-up interview, *Neg* negative, *Pos* positive, *TP* true positive, *FP* false positive, *TN* true negative, *FN* false negative

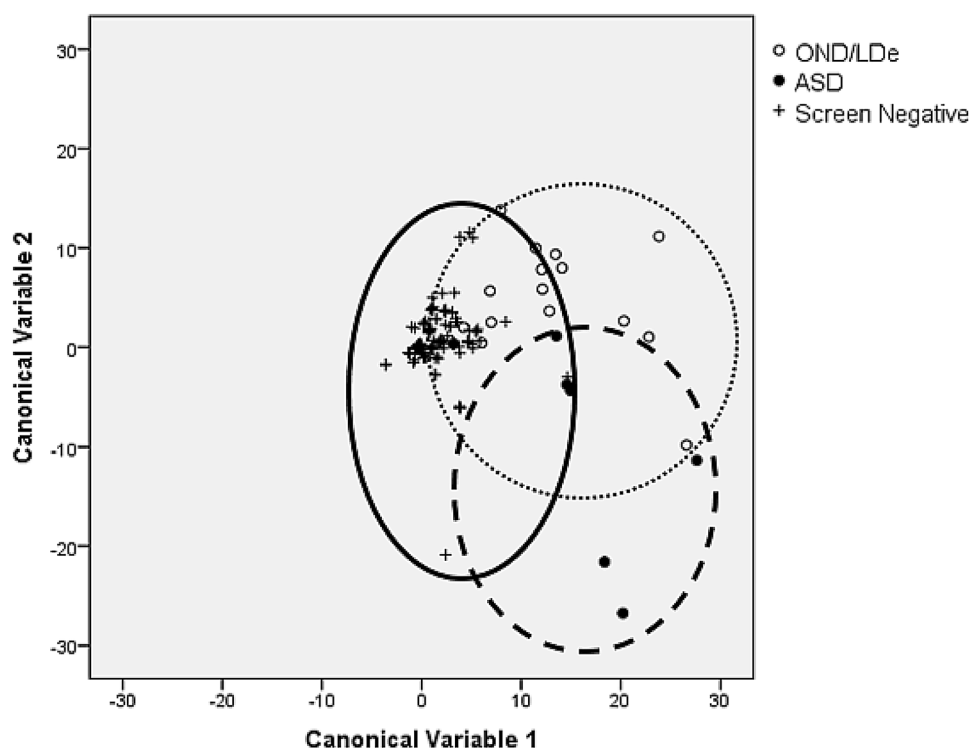


82.3% of variance was explained by Function 1, whereas 17.7% was explained by Function 2. 99.6% of the cases were classified correctly by the model.

Discussion

This study validates the M-CHAT-R/F on a general population sample within the NHS in Spain. The M-CHAT-R/F

Fig. 4 M-CHAT-R/F scatter plot using canonical discriminant analysis at the age range 23–36 months. *ASD* autism spectrum disorder, *OND* other neurodevelopmental disorders, *LDe* language delay, *Neg* negative



was translated, and a cross-cultural adaptation was performed. A total of 6625 questionnaires were collected, with 54 positive screening cases: failing eight or more items in the questionnaire ($n=18$); failing two or more items after FUI ($n=32$); or cause to refer based on physician concerns ($n=4$) (see Figs. 1, 3 for more details). Of the positive screen cases, 72% ($n=39$) completed full assessments. Adding the nine evaluations that were performed to possible FN cases, a total of 48 evaluations were completed, with 19 children receiving a diagnosis of ASD (15 TP and 4 FN). All the evaluated positive screen cases were given a diagnosis of a neurodevelopmental disorder or an LDe, meaning none of the children were identified as having a typical development. This fact shows the accuracy of the screening program in detecting not only ASD but other neurodevelopmental problems as well, supporting the utility of the tool in a NHS for general population.

M-CHAT-R/F psychometrical properties in this research (see Table 3) were similar to original values (except for the prevalence rates) from the validity study carried out by Robins et al. 2014. A previous validation study of the M-CHAT in Spain (Canal-Bedia et al. 2011) also obtained a lower prevalence rate (1/343 in Madrid) than other studies (1/149 in Robins et al. 2014; 1/83 in; Nygren et al. 2012 or 1/147 in; Baduel et al. 2017). Given the age range is similar, the difference may be due to the different size or characteristics of the samples or the methodology applied to implement the different screening programs in different countries. However, prevalence data remains relatively stable over time when

comparing data from the validation study in 2011 and the current study.

Comparing the M-CHAT-R/F in Spain with the M-CHAT in Spain, the number of FUI needed was reduced from 17.2% in Salamanca and Zamora (see Canal-Bedia et al. 2011) to 2.1% in the same areas in this study (also including Valladolid). This means a considerable reduction in the work load of the professionals doing the evaluations, shortening the waiting list for evaluation and hence increasing the families' satisfaction, while also reducing the screening program's costs.

Slightly more effective results were found using the age range of 14–22 months, though they were not significant (see Table 3). These analyses suggest better accuracy of the test for younger children. This finding was also present when comparing psychometrical properties with excluded cases that were positive in the screening. This may support evidence that the M-CHAT-R/F is an adequate tool for early detection. However, this observation is different from other studies (Toh et al. 2018; Sunita and Bilszta 2012) where the original M-CHAT shows better performance at later ages (24–36 months). This could be due to improvements of the tool with the reduction of items and clarification of the items with examples, together with the implementation of the new scoring based on three different risk levels.

Further analysing the different age groups, different patterns in responses can be seen. Items 3 (pretend play), 6 (imperative pointing), 7 (declarative pointing), 9 (show object to share) and 19 (social reference) were the most failed from the ASD group of 14–22 months, whereas the

group of 23–36 months often failed items 10 (response to name), 16 (gaze-following), 18 (understand what is said) and 19 (see Table 2). Further research is needed to this respect, but if similar results are found, it could be feasible to modify the M-CHAT-R/F to a 2-version questionnaire in function of the different age range of the children with the aim of improving the classification of the cases. Another alternative could be to use different screening tools to make a more precise follow-up of the ASD red flags at different ages or apply a combined screening test and surveillance model as Barbaro and Halder (2016) suggest. This “hybrid” model could also be a good strategy to avoid excluding cases in the telephone FUI stage. When a child scores positive in the M-CHAT-R/F questionnaire administered in the normal baby check-up, a short observational measure like the Social Attention and Communication Study (SACS) (Barbaro and Dissanayake 2010) could take place immediately after scoring the questionnaire, as a verification strategy. This would remove any drop out in this stage, which is important as many of the excluded cases are due to difficulties reaching the family by phone to verify the failed items. To address the problem of families rejecting to attend further evaluations, a strategy could be used to keep interest in the progress of their child and try to contact them after some time has passed (from 6 to 12 months, for example) with the intention of evaluating possible changes. This could prove useful as many families still do not feel concern about the development of their child and as Dereu et al. (2010) suggest, parents might be concerned by the development of their child in later stages of transitioning such as from nurseries to schools.

Regarding the discriminant analyses, results are similar in both age strata, showing that the items of the M-CHAT-R/F questionnaire can discriminate between ASD, OND/LDe and Screen Negative groups, although there are cases in which the answers of those groups overlap (see Figs. 2, 4). It is important to note that some cases can be easily identified as not typical development or screen negative, and others identified as disorders other than autism by inspecting their position on the projection, as these cases fall within only one of the groups' clusters. For the cases that are other disorders, there would be no need to follow up with autism specific interventions.

To analyse this result further, the weights given to each item in the discriminant functions were observed. The higher the weight, the more influential it is to that discriminant function. For example, for the age range sample from 14 to 22 months, if Item 15 (imitation of action) (see Table 4) is failed, it is likely to be ASD. If Item 14 (eye contact) is failed, it is more likely to be another disorder, and if passed it is more likely to be ASD. This by itself may not be indicative of the difference between ASD and OND/LDe, but in combination with the responses to other items it seems important.

Table 4 Canonical discriminant function coefficients for items responses at the age range 14–22 months

| Items | Function | |
|--------------------------------|----------------|----------------|
| | 1 ^a | 2 ^b |
| 1. Pointing following | 1973 | – 2089 |
| 2. Wondering hearing | 345 | – 116 |
| 3. Pretend play | 214 | 469 |
| 4. Like climbing | 651 | – 2747 |
| 5. Unusual finger movement | 180 | 585 |
| 6. Imperative pointing | 5987 | – 334 |
| 7. Declarative pointing | 1206 | – 1154 |
| 8. Interest in other children | 1726 | 8437 |
| 9. Show object to share | 4861 | – 701 |
| 10. Response to name | 2427 | – 894 |
| 11. Social smile | – 3015 | 2346 |
| 12. Oversensitive to noise | 154 | – 486 |
| 13. Walking | 926 | – 2318 |
| 14. Eye contact | 921 | – 8798 |
| 15. Imitation of action | 4264 | 14,443 |
| 16. Gaze-following | 863 | – 235 |
| 17. Gaining parent's attention | 319 | 023 |
| 18. Understands what is said | 1481 | 1319 |
| 19. Social reference | 1,150 | 1149 |
| 20. Enjoys being swung | – 2,388 | – 1529 |
| (Constant) | – 278 | 014 |

Unstandardized coefficients

^aFunction 1 discriminates between the Screen Negative group and the ASD and OND/LDe group

^bFunction 2 discriminates between ASD and OND/LDe

For the sample from 23 to 36 months, visual inspection of function 1 (see Table 5) shows the bigger numbers (items 1—pointing following-, 8—interest in other children-, 10—response to name-, 18—understand what is said-) discriminate more between the Screen Negative group and the ASD and OND/LDe group. Function 2 discriminates between ASD and OND/LDe, so items 1 and 15 seem key to this difference. If item 1 (pointing following) is failed and item 15 (imitation of action) passed, the case is likely OND/LDe, whereas if item 15 is failed and item 1 is passed, it is likely ASD. However, when comparing the results obtained in Table 2, item 15 (imitation of action) is not one of the most failed in any of the age ranges for the ASD group, although it is an item that almost none of the subjects from the OND/LDe group fail, and only one subject from the negative screen group failed this item (at the age range of 14–22 months). This item refers to the imitation of actions, an ability that has been shown to be affected in children with ASD in different studies, although its cause remains unclear (Edmunds et al. 2017; Young et al. 2011; Williams et al. 2004).

Table 5 Canonical discriminant function coefficients for items responses at the age range 23–36 months

| Items | Function | |
|--------------------------------|----------------|----------------|
| | 1 ^a | 2 ^b |
| 1. Pointing following | 5048 | 11,688 |
| 2. Wondering hearing | 1009 | 1667 |
| 3. Pretend play | – 023 | 173 |
| 4. Like climbing | 522 | 2496 |
| 5. Unusual finger movement | 099 | 252 |
| 6. Imperative pointing | 427 | – 936 |
| 7. Declarative pointing | 3348 | 3599 |
| 8. Interest in other children | 4967 | 621 |
| 9. Show object to share | 1601 | – 2669 |
| 10. Response to name | 4093 | – 5964 |
| 11. Social smile | 638 | – 937 |
| 12. Oversensitive to noise | – 123 | 011 |
| 13. Walking | 2697 | 3729 |
| 14. Eye contact | 1308 | 3875 |
| 15. Imitation of action | 2176 | – 23,436 |
| 16. Gaze-following | 2137 | 538 |
| 17. Gaining parent's attention | – 970 | – 459 |
| 18. Understands what is said | 5779 | 1686 |
| 19. Social reference | 1306 | – 092 |
| 20. Enjoys being swung | – 3349 | – 1645 |
| (Constant) | – 245 | – 103 |

Unstandardized coefficients

^aFunction 1 discriminates between the Screen Negative group and the ASD and OND/LDe group^bFunction 2 discriminates between ASD and OND/LDe

Limitations

Although the implementation of the screening program was well accepted by the professionals of the different primary health care centers, there are still many aspects to be improved. One of the biggest limitations was not having access to the data of the children who visit the consultations at 18 and 24 months to estimate the actual scope of the program. The corresponding estimations were made according to the number of children registered in the different provinces participating in the study and expected to be in an age range from 18 to 24 months during the years in which the study was conducted (2014–2016). This data is reported by the Spanish National Institute of Statistics (INE 2017). It was also not possible to collect data about families that did not want to participate in the study.

Though low values of internal consistency were found in this study, similar results were found by Robins et al. 2014, suggesting it is a characteristic of the tool. This is likely because the items in the tool do not measure just one dimension.

Another important limitation is the number of excluded cases due to different reasons (not located during the telephone interview, refusal to attend further evaluations, etc). The psychometric properties depend also on the attrition rates, so the values could be under or over-estimated. These properties can vary greatly depending on the positive diagnosed cases found by the screening. To minimize this problem, the M-CHAT-R/F properties were also calculated using the excluded cases (see “Data Analyses” section).

The methods used for detecting false negative cases could be improved for older ages. Once children are attending school, normally at 3 years of age, the children are then supported/evaluated by the educational system instead of the early intervention centres. As the educational system did not collaborate in this study, the detection of false negative cases at later ages was more difficult. With this in mind, the psychometrical properties of the M-CHAT-R/F are likely to be weaker than those reported in this study, and there is possibly a larger difference between the younger and older age strata.

Also, it is important to note that the ASD groups are quite small, so results of the failed items patterns found in this sample should be interpreted with caution because these results may not be significant in bigger samples.

Future Research

For future research it would be good to have a procedure that allows early detection programs to systematically collect data on the scope of universal screenings, to be able to make more realistic estimations of prevalence at early ages. For example, implementing a functionality in the data collection application where each professional could register the number of check-up visits would allow future studies to compare this data with the families participating in the study.

Longitudinal studies of screening, with an evolutionary perspective, could improve the early detection of screening programs (Magán-Maganto et al. 2017). Follow-up of cases at different critical ages, such as 36 months (transition age at which most children in Spain are enrolled to schools) or 48–60 months (age when children have been in schools for a while and their progress and social adaptation to different contexts can be assessed better), would help to improve understanding of the early signs of ASD and its development over time. In the case of children diagnosed with ASD and having a less severe symptomatology or high cognitive functioning, the diagnosis becomes more evident at later ages, when the social demand becomes higher (Christensen et al. 2016; Kleinman et al. 2008).

Future studies are needed to replicate the results found, to confirm that the M-CHAT-R/F has a better performance at younger ages (14–22 months).

Conclusions

This study provides new evidence for the use of the M-CHAT-R/F as a screening tool for neurodevelopmental disorders in the general population. This work is a relevant contribution to current research, identifying ASD at a young age. The new M-CHAT-R/F version has showed a decrease of 15.1% in the number of children needing the FUI in similar areas in comparison with the Spanish validation of the M-CHAT. This fact has an important impact regarding a better and simpler implementation and reduction of the cost for universal ASD screening, in this case within a NHS as a standard practice for pediatricians within the well-child visits. Although there have been many improvements in the last years related to the screening programs, the tools used, or strategies to reduce the false negative and false positive cases, further research is still required to continue improving the early detection of ASD.

The results suggest that the training of pediatric teams in the recognition of the early signs of ASD related to difficulties in the development of social communication, combined with the use of a standardized tool, are fundamental to advance in the early detection of these disorders. The feasibility results of this program can be useful for future strategies for early detection of ASD in the framework of prevention policies.

Acknowledgments We wish to thank all families and professionals from Salamanca, Zamora and Valladolid health areas that have collaborated to make this study possible. This research was conducted as a part of the first author's PhD project.

Author Contributions MMM participated in the study design and coordination, performed the statistical analysis and interpreted the data, and drafted the manuscript; RCB conceived the study, participated in its design and coordination and drafted the manuscript; AHF participated in the design and coordination of the study and performed the measurement; ABM participated in the coordination of the study and helped to draft the manuscript; CFA helped to draft the manuscript; MMV participated in the coordination of the study and performed the measurement; MVMC participated in the coordination of the study; NFR helped to draft the manuscript; HR helped to draft the manuscript; MPP participated in the design and interpretation of the data and helped to draft the manuscript. All authors read and approved the final manuscript.

Funding This work was supported by the State Program of Research, Development and Innovation Oriented towards Social Challenges from *Ministerio de Economía y Competitividad* [Grant No. PSI2013-47840-R] and [Grant No. PSI2016-80575-R]; and DG-SANTÉ, European Commission Project Autism Spectrum Disorders in the European Union—ASDEU [Grant No. SANCO/2014/C2/035].

Conflict of interest The authors have no conflict of interest to declare.

References

- American Psychiatric Association, et al. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. Washington DC: American Psychiatric Pub.
- Atladdottir, H. O., Gyllenberg, D., Langridge, A., Sandin, S., Hansen, S. N., Leonard, H., Gissler, M., Reichenberg, A., Schendel, D. E., Bourke, J., Hultman, C. M., Grice, D. E., Buxbaum, J. D., & Parner, E. T. (2015). The increasing prevalence of reported diagnoses of childhood psychiatric disorders: A descriptive multinational comparison. *European Child & Adolescent Psychiatry*, 24(2), 173–183. <https://doi.org/10.1007/s00787-014-0553-8>.
- Baduel, S., Guillon, Q., Afzali, M. H., Foudon, N., Kruck, J., & Rogé, B. (2017). The french version of the modified-checklist for autism in toddlers (M-CHAT): A validation study on a french sample of 24 month-old children. *Journal of Autism and Developmental Disorders*, 47(2), 297–304. <https://doi.org/10.1007/s10803-016-2950-y>.
- Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., Kurzius-Spencer, M., Zahorodny, W., Rosenberg, R. C., White, T., Durkin, M. S., Imm, P., Nikolaou, L., Yeargin-Allsopp, M., Lee, L.-C., Harrington, R., Lopez, M., Fitzgerald, R. T., Hewitt, A., Pettygrove, S., Constantino, J. N., Vehorn, A., Shenouda, J., Hall-Lande, J., Braun, K. V. N., & Dowling, N. F. (2018). Prevalence of autism spectrum disorder among children aged 8 years—Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2014. *MMWR Surveillance Summaries*, 67(6), 1. <https://doi.org/10.15585/mmwr.ss6706a1>.
- Barbaro, J., & Dissanayake, C. (2010). Prospective identification of autism spectrum disorders in infancy and toddlerhood using developmental surveillance: The social attention and communication study. *Journal of Developmental & Behavioral Pediatrics*, 31(5), 376–385. <https://doi.org/10.1097/DBP.0b013e3181df7f3c>.
- Barbaro, J., & Halder, S. (2016). Early identification of autism spectrum disorder: Current challenges and future global directions. *Current Developmental Disorders Reports*, 3(1), 67–74. <https://doi.org/10.1007/s40474-016-0078-6>.
- Boilson, A. M., Staines, A., Ramirez, A., Posada, M., & Sweeney, M. R. (2016). Operationalisation of the European Protocol for Autism Prevalence (EPAP) for autism spectrum disorder prevalence measurement in Ireland. *Journal of Autism and Developmental Disorders*, 46(9), 3054–3067. <https://doi.org/10.1007/s10803-016-2837-y>.
- Canal-Bedia, R., García Primo, P., Martín Cilleros, M. V., Santos Borbujo, J., Guisuraga Fernández, Z., Herráez García, L., Herráez García, M., Fuentes Biggi, J., Boada Muñoz, L., de la Paz, P. M. V. (2011). Modified Checklist for Autism in Toddlers: Cross-cultural adaptation and validation in Spain. *Journal of Autism and Developmental Disorders*, 41, 1342–1351. <https://doi.org/10.1007/s10803-010-1163-z>.
- Centers for Disease Control and Prevention, CDC. (2018). Autism Spectrum Disorder (ASD). Retrieved from <https://www.cdc.gov/ncbddd/autism/data.html>.
- Christensen, D. L., Bilder, D. A., Zahorodny, W., Pettygrove, S., Durkin, M. S., Fitzgerald, R. T., Rice, C., Kurzius-Spencer, M., Baio, J., & Yeargin-Allsopp, M. (2016). Prevalence and characteristics of autism spectrum disorder among 4-year-old children in the autism and developmental disabilities monitoring network. *Journal of Developmental & Behavioral Pediatrics*, 37(1), 1–8. <https://doi.org/10.1097/DBP.0000000000000235>.
- Dereu, M., Warreyn, P., Raymaekers, R., Meirsschaut, M., Pattyn, G., Schietecatte, I., & Roeyers, H. (2010). Screening for

- autism spectrum disorders in Flemish day-care centres with the checklist for early signs of developmental disorders. *Journal of Autism and Developmental Disorders*, 40(10), 1247–1258. <https://doi.org/10.1007/s10803-010-0984-0>.
- Edmunds, S. R., Ibañez, L. V., Warren, Z., Messinger, D. S., & Stone, W. L. (2017). Longitudinal prediction of language emergence in infants at high and low risk for autism spectrum disorder. *Development and Psychopathology*, 29(1), 319–329. <https://doi.org/10.1017/S0954579416000146>.
- Instituto Nacional de Estadística (INE) (2017). Retrieved from: <http://www.ine.es>.
- Johnson, C. P., & Myers, S. M. (2007). Identification and evaluation of children with autism spectrum disorders. *Pediatrics*, 120(5), 1183–1215. <https://doi.org/10.1542/peds.2007-2361>.
- Kleinman, J. M., Robins, D. L., Ventola, P. E., Pandey, J., Boorstein, H. C., Esser, E. L., Wilson, L. B., Rosenthal, M. A., Sutera, S., Verbalis, A. D., Barton, M., Hodgson, S., Green, J., Dumont-Mathieu, T., Volkmar, F., Chawarska, K., Klin, A., & Fein, D. (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. <https://doi.org/10.1007/s10803-007-0450-9>.
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H. Jr., Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). The autism diagnostic observation schedule-generic: A standar measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, 30(3), 205–223. <https://doi.org/10.1023/A:1005592401947>.
- Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism diagnostic observation schedule—(ADOS-2)* (2nd ed.). Los Angeles, CA: Western Psychological Corporation.
- Magán-Maganto, M., Bejarano-Martín, Á., Fernández-Alvarez, C., Narzisi, A., García-Primo, P., Kawa, R., Posada, M., & Canal-Bedia, R. (2017). Early detection and intervention of ASD: A European Overview. *Brain Sciences*, 7(12), 159. <https://doi.org/10.3390/brainsci7120159>.
- Mandell, D., & Mandy, W. (2015). Should all young children be screened for autism spectrum disorder? *Editorial. Autism*, 19(8), 895–896. <https://doi.org/10.1177/1362361315608323>.
- Nygren, G., Sandberg, E., Gillstedt, F., Ekeröth, G., Arvidsson, T., & Gillberg, C. (2012). A new screening programme for autism in a general population of Swedish toddlers. *Research in Developmental Disabilities*, 33(4), 1200–1210. <https://doi.org/10.1016/j.ridd.2012.02.018>.
- Robins, D. L. (2016). Prevalence counts: commentary on “prevalence and characteristics of autism spectrum disorder among 4-year-old children in the autism and developmental disabilities monitoring network”. *Journal of Developmental & Behavioral Pediatrics*, 37(1), 80–82. <https://doi.org/10.1097/DBP.0000000000000245>.
- Robins, D. L., Casagrande, K., Barton, M., Chen, C. M. A., 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. <https://doi.org/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. <https://doi.org/10.1023/A:1010738829569>.
- Roid, G. H., & Miller, L. J. (1997). *Leiter international performance scale-revised: Examiners manual*. Wood Dale: Stoelting.
- Roid, G. H., & Sampers, J. L. (2004). *Merrill-Palmer-revised: Scales of development*. Wood Dale: Stoelting.
- Sánchez-Valle, E., Posada, M., Villaverde-Hueso, A., Touriño, E., Ferrari-Arroyo, M. J., Boada, L., Martín-Arribas, M. C., Canal, R., & Fuentes-Biggi, J. (2008). Estimating the burden of disease for autism spectrum disorders in Spain in 2003. *Journal of Autism and Developmental Disorders*, 38(2), 288–296. <https://doi.org/10.1007/s10803-007-0393-1>.
- Siu, A. L., US Preventive Services Task Force (USPSTF), Bibbins-Domingo, K., Grossman, D. C., Baumann, L. C., Davidson, K. W., Ebell, M., García, F. A., Gillman, M., Herzstein, J., Kemper, A. R., Krist, A. H., Kurth, A. E., Owens, D. K., Phillips, W. R., Phipps, M. G., & Pignone, M. P. (2016). Screening for autism spectrum disorder in young children: US Preventive Services Task Force recommendation statement. *Journal of the American Medical Association*, 315(7), 691–696. <https://doi.org/10.1001/jama.2016.0018>.
- Soto, S., Linas, K., Jacobstein, D., Biel, M., Migdal, T., & Anthony, B. J. (2015). A review of cultural adaptations of screening tools for autism spectrum disorders. *Autism*, 19(6), 646–661. <https://doi.org/10.1177/1362361314541012>.
- Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *Vineland Adaptive Behavior Scales: Interview Edition: Survey Form*. Circle Pines: American Guidance Service.
- Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland adaptive behavior scales: (Vineland II), survey interview form/caregiver rating form*. Livonia: Pearson Assessments.
- Sunita, & Bilszta, J. L. C. (2012). Early identification of autism: A comparison of the checklist for autism in toddlers and the Modified Checklist for Autism in Toddlers. *Journal of Paediatrics and Child Health*, 49(6), 438–444.
- Toh, T. H., Tan, V. W. Y., Lau, P. S. T., & Kiyu, A. (2018). Accuracy of Modified Checklist for Autism in Toddlers (M-CHAT) in detecting autism and other developmental disorders in community clinics. *Journal of Autism and Developmental Disorders*, 48(1), 28–35.
- US Preventative Services Task Force. Final Recommendation Statement: Autism Spectrum Disorder in Young Children: Screening, (2016). Retrieved from: <https://www.uspreventiveservicestaskforce.org/Page/Document/RecommendationStatementFinal/autism-spectrum-disorder-in-young-children-screening>.
- Williams, J. H., Whiten, A., & Singh, T. (2004). A systematic review of action imitation in autistic spectrum disorder. *Journal of Autism and Developmental Disorders*, 34(3), 285–299. <https://doi.org/10.1023/B:JADD.0000029551.56735.3a>.
- Young, G. S., Rogers, S. J., Hutman, T., Rozga, A., Sigman, M., & Ozonoff, S. (2011). Imitation from 12 to 24 months in autism and typical development: A longitudinal Rasch analysis. *Developmental Psychology*, 47(6), 1565–1578. <https://doi.org/10.1037/a0025418>.