M-CHAT-R/F Validation as a screening tool for early detection in children with autism spectrum disorder

Validación del M-CHAT-R/F como instrumento de tamizaje para detección precoz en niños con trastorno del espectro autista

María Elisa Coelho-Medeiros\textsuperscript{a}, Jonathan Bronstein\textsuperscript{a}, Karina Aedo\textsuperscript{a}, Jaime A. Pereira\textsuperscript{a,\textgreek{g}}, Verónica Arraño\textsuperscript{b}, Carolina A. Perez\textsuperscript{c}, Patricia M. Valenzuela\textsuperscript{d}, Rosario Moore\textsuperscript{d}, Isabel Garrido\textsuperscript{e}, Paula Bedregal\textsuperscript{f}

\textsuperscript{a}Psychiatry Department, Child and Adolescent Psychiatry Unit. School of Medicine, Pontificia Universidad Católica de Chile
\textsuperscript{b}Occupational Therapist. Psychiatry department. School of Medicine, Pontificia Universidad Católica de Chile
\textsuperscript{c}Pediatric Resident. Department of Pediatrics, School of Medicine. Pontificia Universidad Católica de Chile
\textsuperscript{d}Department of Pediatrics, School of Medicine, Pontificia Universidad Católica de Chile
\textsuperscript{e}Nurse. Psychiatry Department. School of Medicine, Pontificia Universidad Católica de Chile
\textsuperscript{f}Public Health Department, School of Medicine, Pontificia Universidad Católica de Chile
\textsuperscript{g}Brain-Computer interfaces and Neuromodulation Lab and Interdisciplinary Center for Neuroscience, School of Medicine, Pontificia Universidad Católica de Chile

Received: 27-04-2018; Approved: 23-10-2018

Abstract

**Introduction:** Screening for Autism Spectrum Disorders (ASD) using the Modified Checklist for Autism in Toddlers, Revised with Follow-up (M-CHAT-R/F) increases early detection, allowing early interventions and improving prognosis. This tool is part of the management in case of suspected ASD in several clinical guidelines. The objective of this article was to conduct the concurrent and discriminant validation and the reliability analysis of M-CHAT-R/F in the Chilean population. **Patients and Method:** This is the second stage of the cross-cultural adaptation of cross-sectional design. M-CHAT-R/F was applied to a sample of 20 children with suspected ASD and 100 randomly selected healthy control children, aged between 16-30 months. Autism Diagnostic Observation Schedule (ADOS-2), considered as reference, was applied to the 20 patients of the clinical sample, to 20 children of the healthy control sample and to those cases of the healthy control sample with M-CHAT-R/F positive. Cronbach alpha was calculated, as well as M-CHAT-R/F and ADOS-2 correlation, sensitivity, and specificity analyses. **Results:** In the healthy sample, M-CHAT-R/F was positive in two patients, with one of them positive and the other one negative for ASD with ADOS-2 test. In the clinical sample, M-CHAT-R/F was positive in all cases, three of them were negative in the ADOS-2 test. The Alfa reliability of M-CHAT-R/F was 0.889, the discriminant sensitivity and specificity were 100% and 98%, and the concurrent ones were 100% and 87.5% respectively. **Conclusions:** The Chilean M-CHAT-R/F version was reliable, sensitive and specific, similar to the original test, which opens the possibility for its use in clinical samples and for research. Validating M-CHAT-R/F is an ongoing process which must be further developed.

Keywords: Autism; Autism Spectrum Disorder; Screening; Child Psychiatry; Child development

Correspondence: Paula Bedregal pbedregal@uc.cl

How to cite this article: Rev Chil Pediatr. 2019;90(5):492-499. DOI: 10.32641/rcped.v90i5.703
Introduction

Autism Spectrum Disorders (ASD) have become very relevant as a public health problem due to the significant increase in their prevalence over the past two decades. It is estimated that 1 out of every 68 children has ASD, which is four times more frequent in males. Diagnosing ASD is not easy since there are no genetic tests or biological markers that detect it. Another difficulty for early diagnosis is that during the first years of life, children with ASD can achieve normal development in some areas, without parents or health professionals suspecting any alteration. Symptoms and signs suggestive of ASD may be subtle and appear gradually, therefore diagnosis is often made late. Thus, most children are diagnosed after four years of age when it becomes evident that the demands of the environment exceed the child’s ability to respond.

Current evidence shows that early ASD diagnosis improves the prognosis and long-term outcome of children with ASD. Early detection allows for individualized, multidisciplinary, and timely treatment that promotes better development of language and social skills, minimizing maladaptive behaviors. For the early ASD detection, the American Academy of Pediatrics (AAP) proposes universal screening for all children aged between 18 and 24 months, thus narrowing the gap between suspicion, diagnosis, and intervention. This reinforces the importance of having reliable screening instruments adapted to the local culture, which can be universally applied to all children in routine health check-ups. Among these screening instruments recommended by the American Academy of Child and Adolescent Psychiatry (AACAP), the Modified Checklist for Autism in Toddlers (M-CHAT) stands out, currently in its version M-CHAT-R/F (R/F: Revised with Follow-up), with sensitivity and specificity over 80%, which incorporates a follow-up interview (Follow-up). The use of this interview greatly reduces cases of false positives, avoiding unnecessary referrals to specialists. Its easy implementation due to a simplified score is also another advantage of the M-CHAT-R/F over the M-CHAT. It is applied to children aged between 16 to 30 months, recruited from the UC Chris tus Health Network (Figure 1). Out of these, 20 children (clinical sample) presented high clinical suspicion of ASD after the evaluation by specialists (pediatric neurologists and psychiatrists, and pediatricians), according to the criteria of the Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-5) of the American Psychiatric Association. Patients with epilepsy or anticonvulsant treatment, severe medical condition or physical disability that made impossible the application of a diagnostic instrument were excluded. The general population sample was made up of 100 children attending well-child visits or general health checks without ASD suspicion. Children attending mental health care for other causes in the 6 months before the administration of the questionnaire were also excluded.

Patients and Method

The study was conducted in two phases. The first was the cultural adaptation from the M-CHAT-R/F in its European Spanish version. A content equivalence, semantic, conceptual, and technical comprehension analysis was made in a purposive sample of UC Health Network users, thus designing the M-CHAT-R/F Chilean version available with free access in the official website of the instrument www.mchatscreen.org. The second phase, and subject of this report, was a cross-sectional design study for the psychometric validation of the instrument M-CHAT-R and R/F. The study required the informed consent of the responsible adult of each participating child and was approved by the ethics committee of the Pontifical Catholic University of Chile.

Participants

The M-CHAT-R and R/F screening were applied on a non-random sample of 120 children aged between 16 and 30 months, recruited from the UC Chris tus Health Network (Figure 1). Out of these, 20 children (clinical sample) presented high clinical suspicion of ASD after the evaluation by specialists (pediatric neurologists and psychiatrists, and pediatricians), according to the criteria of the Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-5) of the American Psychiatric Association. Patients with epilepsy or anticonvulsant treatment, severe medical condition or physical disability that made impossible the application of a diagnostic instrument were excluded. The general population sample was made up of 100 children attending well-child visits or general health checks without ASD suspicion. Children attending mental health care for other causes in the 6 months before the administration of the questionnaire were also excluded.

Instruments

- M-CHAT-R Chilean version: ASD screening questionnaire of 20 dichotomous questions (Yes/No) self-applied by the caregiver: Score < 2 is negative. Scoring between 3-7 points means moderate ASD risk, and indicates follow-up interview (R/F). Score > 8 points, means high ASD risk and indicates direct evaluation by a specialist.
- R/F Chilean version: Structured questionnaire applied by a health professional after brief training. Only items with positive result in M-CHAT-R are
asked, following the flowchart proposed in the instrument. If the R/F score for ‘NOT PASS’ is $> 2$ points, evaluation by a specialist is indicated$^{26}$.

- **ADOS-2**: The Autism Diagnostic Observation Schedule (ADOS) in its second version ADOS-2 is a standardized and semi-structured evaluation of communication, social interaction, interests and imaginative play, which defines the level of concern regarding the possible ASD diagnosis. It is considered the gold standard instrument in the MINSAL Clinical Guideline$^{16}$. Since ADOS-2 is a semi-structured instrument, there is no specific validation in Chile. It has been established, within the rules of its application, that the evaluator makes a cultural adjustment, if necessary, during the session. The Spanish translation of the instrument manual is duly validated$^{28}$. This evaluation instrument consists of a set of precise activities, in a standardized context, where the evaluator observes certain behaviors relevant to the ASD diagnosis. There are four modules that are determined by the language level and age of the child. In this study, module T was applied, which is designed for children under 30 months of age, regardless of their language level. While ADOS-2 is considered the best test for diagnosing autism, the definitive diagnosis is still based on specialized multidisciplinary clinical evaluation$^{29}$.

**Procedure (Figure 1)**

The M-CHAT-R was applied by previously trained specialists doctors or medical residents to the entire clinical and general population sample. The R/F
was applied according to the questionnaire procedure. ADOS-2 was applied to the entire clinical sample and in the general population sample to those patients with an M-CHAT score > 8 points or R/F score > 2 points. In those cases, the parents were informed of the risk involved in a positive screening. Therefore, a non-blinded, standardized assessment (ADOS-2) was conducted simultaneously by three expert evaluators, but with separate instrument coding, discussing discrepancies, if any, in order to reach consensus. For the discriminant analysis, 20 patients were randomly selected from the general population sample with negative M-CHAT-R who also were assessed with ADOS-2.

When other types of clinical diagnoses were detected during the evaluation, suggestions for treatment and referrals were made for each particular case.

Analysis
A sociodemographic descriptive analysis of the sample was performed, and both sub-samples were compared using non-parametric tests (Mann-Whitney U test), and chi-square test for nominal variables. P < 0.05 values were considered statistically significant differences. The internal reliability of the M-CHAT-R/F was calculated using Cronbach’s alpha. Concurrent validity was carried out through correlation analysis, comparing the results of M-CHAT-R/F and ADOS-2 (considered gold standard) in both sub-samples and calculating sensitivity and specificity. Discriminant validity was estimated by comparing the high-risk clinical diagnosis of ASD with the M-CHAT-R/F outcome.

Results
Sociodemographic characteristics
The clinical sample tended to include children older than 24 months, most of them were male (95%), and the caregivers were older than the caregivers of the general sample, with no statistically significant differences in the achieved education level and the health care system (Table 1). There were statistically significant differences between the age of the general and the clinical samples (22 vs 24 months respectively, p < 0.006) as well as differences in the distribution by sex: 58% males in the general sample, and 95% in the clinical sample which is statistically significant (chi-square, p = 0.01).

Table 1. Sample characteristics

<table>
<thead>
<tr>
<th></th>
<th>General Sample</th>
<th>Clinical Sample</th>
<th>Total</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N: 100</td>
<td>N: 20</td>
<td>N: 120</td>
<td></td>
</tr>
<tr>
<td>Mean Age in months (S.D)</td>
<td>22.03</td>
<td>24.85</td>
<td>22.47</td>
<td>*p = 0.007³</td>
</tr>
<tr>
<td></td>
<td>-4.22</td>
<td>-3.43</td>
<td>-4.22</td>
<td></td>
</tr>
<tr>
<td>Male Sex (%)</td>
<td>58.0</td>
<td>94.8</td>
<td>64.2</td>
<td>*p = 0.001²</td>
</tr>
<tr>
<td>Mean caregiver age in years (S.D)</td>
<td>28.87</td>
<td>34.61</td>
<td>29.79</td>
<td>*p = 0.125¹</td>
</tr>
<tr>
<td></td>
<td>-13.06</td>
<td>-4.3</td>
<td>-12.27</td>
<td></td>
</tr>
<tr>
<td>Educational level (caregiver; &gt; 8 years, %)</td>
<td>78.0</td>
<td>85.0</td>
<td>79.2</td>
<td>*p = 0.001²</td>
</tr>
<tr>
<td>W/D: 15.0</td>
<td>W/D: 10</td>
<td>W/D: 14.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health forecast FONASA (%)</td>
<td>38.0</td>
<td>35.0</td>
<td>37.5</td>
<td>*p = 0.762²</td>
</tr>
<tr>
<td>W/D: 19%</td>
<td>W/D: 10%</td>
<td>W/D: 17.5%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Mean difference between general sample and overall clinical sample is statistically significant. W/D: Without Data. S.D: Standard Deviation. 1. Mann-Whitney U test. 2. Chi-square test.

Average scores of tests studied
The clinical sample showed significantly higher average scores than the general sample, both in M-CHAT-R/F (9.1 vs 0.81) and in ADOS-2 (15 vs 2.7). (Table 2).

M-CHAT-R and R/F Behavior (See Figure 1):
In 90% of the general population sample, the M-CHAT-R was negative and in 10% it was positive. The R/F was applied to this group of patients according to the flowchart proposed by the authors of the instrument. Out of these, seven were negative (score < 1) and three patients continued with positive results. These patients were referred for evaluation by a specialist and application of ADOS-2 test, module T. In one case, the result was categorized as High Risk for ASD and in another case, the diagnosis was ruled out. The third case abandoned the protocol before the diagnostic evaluation.

To detect false negatives, 20 cases were randomly selected from the general population sample and the ADOS-2 test was applied to them, all with Low or No Risk for ASD results.

In the clinical sample with suspected ASD, both M-CHAT-R and R/F were positive in all subjects. This
sample was clinically evaluated and analyzed with the ADOS-2 test. In three cases, the result was Low or No Risk for ASD, and in the remaining 17, the result was High risk for ASD. These three false-positive cases corresponded to patients with delayed language development, that is, another neurodevelopmental disorder, but which does not interfere with pragmatic communication or social reciprocity.

Reliability: The internal consistency was analyzed for all items of the M-CHAT-R and for the R/F with a result considered adequate (Cronbach’s alpha = 0.889). This result is above the reliability level of the original instrument in this same version (Alpha = 0.79)14.

Validity
Discriminant validity: The contrast of M-CHAT-R/F with clinical suspicion of ASD showed 100% discriminant validity. (Table 3)

Concurrent Validity: The M-CHAT-R/F with the ADOS-2 diagnostic test showed a high correlation between their scores (Pearson r = 0.849, p = 0.0001) and the sensitivity and concurrent specificity was 100% and 83.3% respectively (Table 3).

Discussion
In this work, we continue with the validation process of the M-CHAT-R/F, already initiated with the cultural adaptation to Chilean Spanish25,26. Having a screening instrument that facilitates early screening for ASD may allow for early interventions and better prognosis in the development of children with ASD. The M-CHAT-R/F has been demonstrated to fulfill this purpose worldwide30.

In general terms, the Chilean instrument showed adequate internal consistency, and successfully distinguishes the healthy population from the one with a developmental disorder (Table 3).

When comparing this instrument with a standardized one for the ASD diagnosis such as ADOS-2, it presented high levels of sensitivity and specificity, similar to the original M-CHAT-R/F created by Robins et al14.

It is worth mentioning that in the clinical population, false positives detected with the M-CHAT-R/F turned out to have other types of neurodevelopmental disorders, similar to that described in the validation of the original instrument by Robins et al31. Therefore, these are children who likewise require evaluation and interventions by specialized child development teams32.

The male:female difference ratio observed in our study exceeds the one described in the literature, which indicates a ratio of 4 males per 1 female in ASD patients, even in the last prevalence studies33. However, in children younger than 30 months, this ratio is not as consistent due to a trend towards later detection of ASD cases in females34. The biases identification involved in this difference regarding the early diagnosis of ASD among boys and girls supports the development of current lines of research35.

It is observed that the average age of the children’s caregivers in the clinical sample is considerably higher than the general sample. This data is consistent with that described in the literature, where the age of the pa-

<table>
<thead>
<tr>
<th>Test</th>
<th>General Sample</th>
<th>Clinical Sample</th>
<th>Total</th>
<th>Valor p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n = 100</td>
<td>n = 20</td>
<td>n = 120</td>
<td></td>
</tr>
<tr>
<td>M-CHAT-R/F:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Mean</td>
<td>0.81</td>
<td>9.10</td>
<td>2.19</td>
<td>0.000</td>
</tr>
<tr>
<td>- SD</td>
<td>-1.21</td>
<td>-3.12</td>
<td>-3.52</td>
<td></td>
</tr>
<tr>
<td>ADOS-2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Mean</td>
<td>2.70</td>
<td>15.00</td>
<td>8.42</td>
<td>0.000</td>
</tr>
<tr>
<td>- SD</td>
<td>3.36</td>
<td>5.81</td>
<td>7.73</td>
<td></td>
</tr>
</tbody>
</table>

1. Mann-Whitney U test
rents is a risk factor for ASD\textsuperscript{36}. Finally, in this sample, we have a population that mostly has caregivers with higher technical or university education, something that is not necessarily representative of the Chilean reality\textsuperscript{37}, but that could characterize the user population of the UC-Christus Health Network.

Within the limitations of the study, the small number of the sample stands out, limited mainly by the availability of patients under 30 months of age with clinical suspicion of ASD in our network. Due to this difficulty, we designed a study with a 5:1 ratio between the general sample and the clinical one. Therefore, it is suggested to incorporate other national contexts to describe the behavior of the M-CHAT-R/F in a broader population that is representative of the Chilean reality.

Another methodological problem we faced in the design was the difficulty of blinding it to the ADOS-2 application since parents, being aware of the result of the previous screening, came to this evaluation with diagnostic doubt, identifying as bias the over-interpretation of the relationship between the two instruments. Both difficulties are a common barrier in the validation studies of screening/diagnostics instruments of ASD in this age range, which is an important methodological challenge to overcome\textsuperscript{29}.

Despite the difficulties mentioned above, this work has resulted in the implementation of training for health professionals in primary care to implement this adapted version of the M-CHAT-R/F in the Well-Child Care Program for 18-month-old children in those who present risk factors for ASD such as children with a developmental delay in the language and/or social areas in the Psychomotor Development Evaluation Scale (Escala de Evaluación del Desarrollo Psicomotor-EEDP) test, siblings and children of patients with ASD, as well as cases where there is a direct clinical suspicion. On the other hand, this study is also considered an initial but important step for conducting studies of ASD prevalence in Chile in a population younger than 30 months. Determining the ASD prevalence, which is still unknown in our country, will allow us to measure the need and be able to distribute the necessary resources from the implementation of specialized health systems, as well as public policies that contribute to the well-being of these patients and their families.

Conclusion

The Chilean version of the M-CHAT-R/F was reliable, sensitive and specific similarly to that described in the original study, which allows Chile to have a validated screening instrument for Autism Spectrum Disorder. Certain methodological limitations were observed, such as the small sample and the difficulty of blinding in the diagnostic evaluation. Although the validation of clinical instruments is a continuous process which is constantly improved, this first step provides two very significant contributions: at the level of clinical practice, it allows improving the diagnostic capacity in the early detection of ASD, and at the level of public health, the application of the same validated screening instrument achieves the standardization of early detection in primary care. In this way, an early diagnosis can be made, improving the prognosis and therefore the health burden.

Ethical Responsibilities

**Human Beings and animals protection:** Disclosure the authors state that the procedures were followed according to the Declaration of Helsinki and the World Medical Association regarding human experimentation developed for the medical community.

**Data confidentiality:** The authors state that they have followed the protocols of their Center and Local regulations on the publication of patient data.

**Rights to privacy and informed consent:** The authors have obtained the informed consent of the patients and/or subjects referred to in the article. This document is in the possession of the correspondence author.

**Conflicts of Interest**

Authors declare no conflict of interest regarding the present study.

**Acknowledgments**

To the departments of Psychiatry, Pediatrics and Public Health of the Pontificia Universidad Católica de Chile for their support in carrying out the work.
To Dr. Manuel Arriaza, Pediatric Neurologist at Sótero del Río Hospital and Red Salud UC-Christus for his active referral of patients for the study.
To the hundreds of parents and multiple professionals who collaborated in the different phases of the study.

**Financial Disclosure**

Special seed-interdisciplinary research competition 2015, Project PS 08/15 Dr. Coelho-Medeiros - Dr.
Bedregal. Research Directorate, School of Medicine. Pontificia Universidad Católica de Chile. Entity that financed the work had no influence on any aspect of it (design, collection, analysis or interpretation of data, preparation, revision or approval of the final manuscript).

References

30. Zwaigenbaum L, Bauman ML, Fein D,