Flores-Rodríguez, Yassel; Roldán Ceballos, Ofelia; Albores-Gallo, Lilia
Assessing autism with DSM-IV and DSM-5 criteria using the Childhood Autism Rating Scale (CARS)

salud mental

Assessing autism with DSM-IV and DSM-5 criteria using the Childhood Autism Rating Scale (CARS)

Yassel Flores-Rodríguez,[1] Ofelia Roldán Ceballos,[2] Lilia Albores-Gallo[3]

[1] Hospital Psiquiátrico Dr. Samuel Ramírez Moreno, Estado de México, México.

[2] Hospital General Dr. Fernando Quiróz Gutierrez, Ciudad de México, México.

[3] Hospital Psiquiátrico Infantil Dr. Juan N. Navarro, Ciudad de México, México.

Correspondence:
Lilia Albores-Gallo
Hospital Psiquiátrico Infantil Dr. Juan N. Navarro.
Av. San Buenaventura No. 86,
Col. Belisario Domínguez,
Tlalpan, CP. 14080,
Ciudad de México, México.
Phone: 55 2316-2517
Email: lilialbores@gmail.com /
algali.1209@gmail.com

ABSTRACT

Introduction. The Childhood Autism Rating Scale (CARS) is an observational instrument for assessing autism. It is the most widely used instrument for the diagnosis of autism in Mexico. **Objective.** To study the psychometric properties in the Mexican population using DSM-IV and DSM-5 criteria. **Method.** This is an observational, cross-sectional study. Participants (N = 137) were 78.8% male, with an age range between 2 and 18 years, mean 8.5. Parents completed the M-CHAT, the ABC, the ADI-R, and semi-structured clinical interviews with DSM-IV. We inferred DSM-5 criteria by selecting atypical sensory items from the ABC. **Results.** The internal consistency for the total CARS items was α = .88. The concurrent validity of the CARS and the DSM-IV criteria showed a Spearman coefficient of correlation of r^s = .62. The convergent validity with the ADI-R showed moderate correlations, r^s = .32 to .61, with the DSM-IV k = .33 and DSM-5 k = .36. The concordance with the DSM-IV and DSM-5 was 71% and 84.5%, sensitivity was 58.1% and 46-7%, and specificity 76.6% and 90.5%, respectively. The discriminant analysis showed that the CARS correctly classified 97% of children with autism, 70% with pervasive developmental disorders not otherwise specified (PDD-NOS), and 12.5% with Asperger's syndrome. **Discussion and conclusion.** The Mexican version of the CARS is a valid and reliable instrument for diagnosing autism in Mexican children and adolescents.

Keywords: Autism, autism spectrum, validity, reliability, CARS.

RESUMEN

Introducción. La Escala de Evaluación del Autismo Infantil (CARS) es un instrumento observacional para evaluar el autismo. Es la herramienta más utilizada para el diagnóstico del autismo en México. **Objetivo.** Estudiar sus propiedades psicométricas en la población Mexicana utilizando los criterios del DSM-IV y del DSM-5. **Método.** El estudio fue observacional y transversal. Los participantes (N = 137), 78.8%, fueron del sexo masculino, con un rango de edad de entre 2 y 18 años, media = 8.5. Los padres completaron el M-CHAT, el ABC, el ADI-R, y la entrevista clínica semiestructurada según criterios del DSM-IV. Inferimos el diagnóstico según los criterios del DSM-5 seleccionando algunos reactivos sensoriales atípicos de la escala ABC. **Resultados.** La consistencia interna del total de ítems del CARS fue α = .88. La validez concurrente entre el CARS y el diagnóstico según el DSM-IV con el coeficiente de correlación de Spearman fue r^s = .62. La validez convergente con el ADI-R mostró correlaciones moderadas, r^s = .32 a .61, con el DSM-IV k = .33 y el DSM-5 k = .36. La concordancia con el DSM-IV y el DSM-5 fue de 71% y 84.5%; la sensibilidad fue de 58.1% y 46-7%; y la especificidad de 76.6% y 90.5%, respectivamente. El análisis discriminante mostró que el CARS clasificó correctamente al 97% de los niños con autismo, al 70% con trastornos generalizados del desarrollo no especificados de otra manera (PDD-NOS) y al 12.5% con el síndrome de Asperger. **Discusión y conclusión.** La versión mexicana del CARS es un instrumento válido y confiable para el diagnóstico del autismo en niños y adolescentes mexicanos.

Palabras clave: Autismo, espectro autista, validez, confiabilidad, CARS.

INTRODUCTION

Autism Spectrum Disorders (ASDs) DSM-5 criteria (American Psychiatric Association, 2013) include symptoms from one broad domains of deficits in social communication and social interaction which in DSM-IV (American Psychiatric Association, 2000) were two independent domains. Also, DSM-5 requires two symptoms of the last domain of restricted, repetitive patterns of behavior, interests, or activities with the inclusion of hyper- or hyporeactivity to sensory input or unusual interest in sensory aspects of the environmental. In DSM-IV-TR only one criterion was required and there was no consideration on sensory atypical interests.

Even though we have seen progress in timely recognition, diagnostic delay in countries with medium and low income persist (Montiel-Nava, Chacín, & González-Ávila, 2017; Tapia Guillen et al., 2019; Zavaleta-Ramírez et al., 2020). For example, in México, the prevalence in León, Guanajuato, was .87% (95% CI [.62, 1.1]) (Fombonne et al., 2016), consistent with other recent studies of prevalence (Pedersen et al., 2012). Fombonne et al. (2016) showed that most Mexican children (77.8%) were diagnosed after three years old, even when their parents detected the first symptoms before that age, and most families (61.2%) had earlier contact with health services. Some obstacles for recognizing early signs of autism are lack of training among the health care professionals and sub-optimal performance of the few valid screening instruments'.

Validity studies of the M-CHAT (Modified Checklist for Autism in Toddlers) (Robins, Fein, Barton, & Green, 2001), the SRS (Social Responsiveness Scale) (Constantino & Gruber, 2005), and the ABC (Autism Behavior Checklist) (Krug, Arick, & Almond, 1980) show cultural bias and inconsistent psychometric properties (Albores-Gallo et al., 2012) requiring adjustments in the cut-off scores (Fombonne, Marcin, Bruno, Tinoco, & Marquez, 2012; Living-Varela & Albores-Gallo, 2012). On the other hand, gold standards such as the ADI-R (Le Couteur, Haden, Hammal, & McConachie, 2008) and the ADOS-G (Lord, Rutter, Dilavore, & Risi, 1999) are valuable but expensive and hard to use in busy clinical settings.

CARS (Child Autism Rating Scale) Schopler, Reichler, DeVellis, and Daly (1980) is a diagnostic observational assessment that has consistently shown good psychometric properties. In the original validity study, the internal consistency was very high (Cronbach's alfa α = .94), and the inter-rater reliability showed a concordance of .71, and the concurrent validity with the clinical assessment showed high correlation r = .84, p = < .001 and r = .80, p = < .001.

In 2013, Breidbord and Croudace (2013), performed a meta-analysis of the CARS validity studies published between 1980 and 2012. The inter-item reliability was estimated (k = .19, M = .896, IC 95% [.877, .913], Q = 99.66) as well as the inter-rater reliability measured by the Cohen's

kappa coefficient and the Cochran's Q (k = .33, M = .796, IC 95% [.736, .844], Q = 183.36), and the test-retest (k = 0.8, M = .756, IC 95% [.690, .814], Q = 7.52).

In Mexico, CARS is the most used tool in centers providing therapeutic services for children with autism (Harris & Barton, 2016). Nonetheless, the Mexican version's psychometric properties are unknown. This study aims to compare the instrument performance with the DSM-IV and DSM-5 criteria and investigate some of the psychometric properties of the Mexican version of the CARS (Childhood Autism Rating Scale) in a clinical sample of children and adolescents seeking attention in a child psychiatric hospital.

METHOD

Participants (N = 137) with an age range between 2 and 18 years seeking the diagnosis of Autism Spectrum Disorder in a Psychiatric Hospital. We excluded children with any genetic syndrome associated with Autism Spectrum Disorder.

Instruments

ADI-R (Lord, Rutter, & Le Couteur, 1994)

The Autism Diagnostic Interview, Revised (ADI-R), is a semi-structured interview administered by an experienced clinician for individuals with a mental aged above any genetic syndrome associated with Autism Spectrum Disorder. The interview assesses three domains of symptoms, consistent with the ICD-10 and DSM-IV criteria: language/communication, reciprocal social interactions, restricted, repetitive, and stereotyped behaviors and interests. It also evaluates the period between ages four and five (when symptoms are the most intense). It consists of 93 items, and its results are categorical. The inter-rater reliability was k = .63 to .89 for each item; the intraclass correlation coefficients (ICC) were above .92 for all the domains and subdomains. This study used the algorithm for current behavior, comprising four categories (A, B, C, D), which assess reciprocal social interaction, communication and restricted, repetitive, and stereotyped behavior patterns, and age of developmental abnormalities.

M-CHAT (Robins et al., 2001)

The M-CHAT is a modification to the Checklist for Autism in Toddlers (Baron-Cohen, Allen, & Gillberg, 1992). It consists of 23 dichotomous response items (yes/no) from the parent report. The original validity study used a sample of 1,293 children ages 16 to 30 months of the community. The discriminant function identified six critical items due to the higher coefficients of the discriminating function: items 2, 7, 9, 13, 14, and 15.

Total internal reliability and relation to the critical questions were adequate (α = .85 and α = .83, respectively).

Failure in three of the 23 items of the M-CHAT or two from six critical items reported sensitivity of 87%, an specificity of 99%, a positive predictive value of 80%, and negative predictive value of 99%.

In Mexico, Albores-Gallo et al. (2012) validated the Mexican version with a mixed sample (clinical and community) of 456 children. The Mexican version of the M-CHAT's internal consistency for the total items (23 items) was KR = .75 and .70 for the six critical items. The convergent validity of the M-CHAT total with the CBCL/1.5-5/PPD reported a Spearman coefficient rho = .66, $p = .01$; correlations with the ADI-R were within a range of rho = .23-.66. The criterion validity between the CARS and the ADI-R (dimensions A, verbal B, and non-verbal B, and C) was $k = .17$ to .61. High concordance with the non-verbal B dimension was found, $k = .61$ $p = .0001$, contrasting with the verbal dimension $k = .29$ $p = .004$.

ABC (Krug, Arick, & Almond, 1993)

The Autism Behavior Checklist (ABC) assesses individuals from 18 months to 35 years of age. It explores 57 atypical behaviors or items organized in five subscales: sensory stimuli, relationships, the use of the body and objects, language, and socialization/self-help. All patients had completed the ABC at the moment of the study, so we used the ABC's sensory items to match the DSM-5 criteria. The selected items of the sensory stimuli subscale were: 1. Misuse of visual discrimination when learning (attention is fixed on parts of the objects, such as size, color, position), 2. Appears not to hear (despite normal hearing tests), 3. Sometimes the individual will not show "startle response" to loud noises, 4. The individual will sometimes not react to painful stimuli (cuts, pinches, contusions), 5. The individual will often do not blink when bright lights are directed to the eyes, 6. Will cover his/her ears to many sounds, 7. Will frown, squint or cover eyes to natural lights, 8. Will frequently not visually react to "new" people, and 9. Stares at no particular object for long periods. The internal consistency reported for this subscale is an alpha coefficient of .59.

Childhood Autism Rating (CARS)
(Schopler, Reichler, & Renner, 1988)

CARS is an observational diagnostic instrument that assesses symptoms by comparing a child likely to be autistic with a neurotypical individual of the same age. The behavior observed in different settings; during a clinical evaluation or in the classroom, through videos recorded by the parents, or by a combination of resources. The scale consist of 15 items: - Relationship with people, mimicking behaviors, emotional response, use of own body, use of objects, adaptation to changes, visual response, auditory response, response to and use of taste, smell, and touch, fear or nervousness, verbal communication, non-verbal communica-

tion, activity levels, level and consistency of intellectual response, general impression.

The scores range from one (the individual behaves within normal limits for the age group) to four (the individual's behavior is severely abnormal for the age group). Mid-point scores are assigned (1.5, 2.5, and 3.5) when the assessed behavior falls between two items. Results reflect normal to severely abnormal punctuation in a continuum. A cut-off score of 30 or above classifies as autism; the range from 30 to 36.5 defines mild to moderate autism, and 37 to 60 defines severe autism. The original study published in 1980 (Schopler et al., 1980) reported internal consistency of $\alpha = .94$ for the instrument's total. Inter-rater reliability demonstrated a concordance of .71. Concurrent validity between the instrument and the clinical judgment were between r = .84, $p = < .001$ and r = .80, $p = < .001$.

Procedure

A convenience sample of patients seeking a diagnosis for autism in the Children's Psychiatric Hospital Juan N. Navarro clinic (Hospital Psiquiátrico Infantil Dr. Juan N. Navarro). Initially, the parents of the children completed the M-CHAT and the ABC instruments. Trained professionals conducted a semi-structured clinical interview based on the DSM-IV criteria (gold standard), and the ADI-R (gold standard) interview was completed in approximately 3.5 hours by an experienced child psychiatrist.

Researchers blinded to the ADI-R, M-CHAT, and the DSM-IV clinical interview results completed the CARS.

We used ADI-R and the DSM-IV interview results to assign patients to three groups: autism, asperger's syndrome, or pervasive developmental disorder not otherwise specified (PDD-NOS).

Statistical analysis

The variables were expressed as means, standard deviations, and proportions. Reliability and internal consistency were analyzed using Cronbach's alfa coefficient for the total sample and the subsamples by gender.

Concurrent validity was calculated with Spearman's Rho coefficients of correlation between the CARS (total score) and the clinical diagnosis score, using DSM-IV criteria.

CARS Concordance with the DSM-IV and DSM-5 criteria

To infer the DSM-5 diagnosis we combined the DSM-IV domains of social interaction and communication deficits to match the DSM-5 social communication domain. We selected the ABC items that best-represented hyper or hypo activity sensory stimuli or hypersensitivity and added restrictive and repetitive behavior to match the DSM-5 criteria.

1. Sometimes an individual will not show "startle response" too loud noises; 2. the individual will sometimes not react to painful stimuli (cuts, pinches, contusions); 3. the individual will frequently not blink when bright lights are directed to the eyes; 4. will cover ears to many sounds; 5. will frown, squint or cover eyes to natural lights; 6. will touch, smell, or taste close objects.

We used the kappa coefficient to calculate the concordance rate of DSM-IV and DSM-5 diagnosis with the CARS (cut-off = 33).

All the instruments' information was gathered in an Excel database and then transferred to the SPSS statistical program, version 17.0.

Discriminant validity

To study the instrument's discriminant validity, we conducted a discriminant analysis with all the CARS items to analyze which items predict the membership to the diagnostic categories of Autism, Asperger, and PDDNOS.

Ethical considerations

This study received approval from the ethical committee of the institution (Hospital Psiquiátrico Infantil Dr. Juan N. Navarro) with the following ID number II3-03-1112.

RESULTS

Sample characteristics

An experienced child psychiatrist assessed 137 children; we eliminated 13 cases from the analysis because they were incomplete. The final sample consisted of children ($N = 124$), 78.8% ($N = 99$) male. The children's average age was 8.5 ($SD = 3.8$), ranging from ages 2 to 18 years. The mean age of fathers was 37.9 ($SD = 7.8$), and the mean age of mothers was 34.2 ($SD = 7.2$). The total mean score of the CARS was $M = 32.3$, $SD = 7.3$, ranging between 14 and 50.

Internal consistency

Internal consistency for the total items analyzed by Cronbach's alfa coefficient was α = .88. By gender, internal consistency and mean score for males ($N = 79$) was α = .87, $M = 50.9$, $SD = 13$, and for females ($N = 20$) was α = .90, $M = 49.8$, $SD = 15.8$.

Concurrent validity

Spearman's Rho coefficient between the CARS total score and the DSM-IV diagnosis criteria demonstrated moderate correlation ($r^s = .62$, $p < .0001$).

Table 1

Spearman correlation between CARS and ADI-R symptom domains

	1	2	3	4
1 CARS/Tot	1			
2 ADI-R/SI	.59**	1		
3 ADI-R/CO	.61**	.70**	1	
4 ADI-R/RB	.32**	.38**	.25**	1

Notes: **$p ≤ .01$, *$p < .05$. CARS TOT = Total score. ADI-R/SI = Social Impairment total score. ADI-R/CO = Communication total score. ADI-R/RB = Repetitive behavior total score.

Convergent validity

Spearman's coefficients of correlation between the CARS (total score) and the ADI-R subscales are shown in Table 1.

The highest correlation between the CARS total score and the ADI-R was $r^s = .61$ for the B domain of the communication. The lowest correlation was $r^s = .32$ for the C domain of restricted, repetitive, and stereotyped behavior patterns.

The kappa coefficient between the categorical diagnosis of the CARS (cut-off score of 33) and the M-CHAT (cut-off: 2/6 critical items) demonstrated a kappa coefficient of $k = .66$, $p = < .001$.

Concordance rate with DSM-IV and DSM-5 diagnosis criteria

The concordance rate between the CARS and the categorical diagnosis by DSM-IV and DSM-5 was 71% and 84.5%, respectively (Table 2).

Discriminant analysis

A discriminant analysis was conducted to identify which CARS items could better predict the membership to the diagnostic groups of Autism, Asperger's Syndrome, and PDD-NOS. Two discriminant functions were calculated, with two canonical correlations, r = .627 and r = .563; both

Table 2

Concordance, sensitivity, and specificity of CARS using DSM-IV and DSM 5 diagnostic criteria

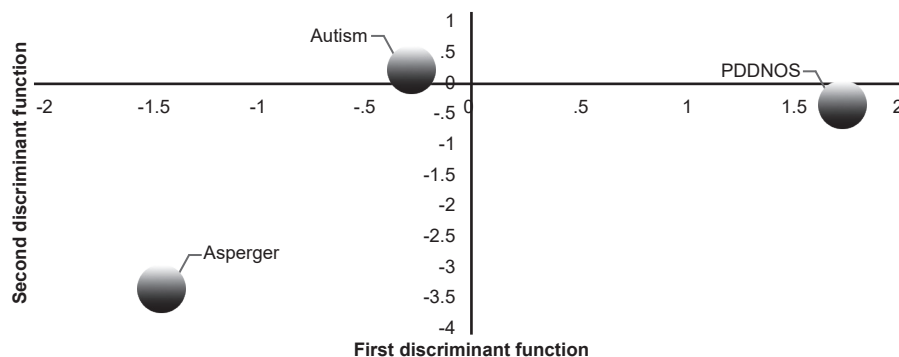|	DSM IV	DSM 5
Kappa coefficient	.33	.36
Concordance %	71%	84.5%
Sensitivity	58.1%	46.7%
Specificity	76.6%	90.5%

Note: CARS cut-off = 33

Figure 1. Centroids of DSM-IV diagnostic categories on the two discriminant functions.

functions explain the 58.2% and the 41.8% of the variance with effect size of 1.0 and .31, respectively. For the combination of both functions (with all the items) the Wilks' lambda was significant $\Lambda = .415$, $\chi^2 = 71.282$, $p < .0001$;

Table 3
Functions at groups centroids

	Function	
Diagnosis DSM IV	1	2
Autism	-.293	.221
Asperger	-1.452	-3.414
PDDNOS	-1.718	-.393

Table 4
Intra-group correlations between discriminant variables and the discriminant functions

	Function	
	1	2
Visual response	.746*	-.048
Emotional response	.644*	.124
Sensory response	.543*	.033
Fear or nervousness	.542*	.036
Hearing response	.537*	-.107
Activity level	.435*	-.128
Verbal communication	.431*	.154
Imitation	.399*	.182
Object use	.378*	.288
Body use	.367*	.233
Change adaptation	.360*	.350
Intelectual response level	.182	.483*
Non-verbal communication	.398	.435*
Relationship with people	.248	.381*

Note: Intra-group correlations between discriminant variables and Canonical discriminant functions. * Cut-off > .30 to evaluate the magnitude of the item contribution to discriminate the groups as suggested by Finch (2009) and Schopler et al. (1980).

when removing the first function, the Wilks' lambda remained significant $\Lambda = .683$, $\chi^2 = 30.887$, $p = .006$, so both functions are relevant.

The centroid positions are shown in Figure 1 / Table 3. The CARS sensitivity analysis for diagnosing the subgroups shows that CARS discriminates between Autism and PDD-NOS, correctly identified 97% and 70% of the cases, respectively. However, its effectiveness in assessing Asperger's Syndrome is lower, classifying only 12.5% of the cases (Table 3).

For the second function, only the intellectual response level and consistency, adaptation to changes, non-verbal communication, and relationship with people were significant (Table 4).

DISCUSSION AND CONCLUSION

In this study, we analyzed the CARS Mexican version's psychometric properties as an observational instrument for diagnosing autism.

The sample was predominantly male (78.8%), which is consistent with the higher ratio between male and females and with most of the reliability and validity studies from countries such as Japan (77.2% and 75.8%) (Kurita, Miyake, & Katsuno, 1989; Tachimori, Osada, & Kurita, 2003), United States (82%) (Sturmey, Matson, & Sevin, 1992), and Spain (74.5%) (García-Villamisar & Muela, 2000).

The instrument showed excellent internal consistency for the total of items, demonstrated with a Cronbach alfa coefficient $\alpha = .87$. This result is higher than coefficients reported for children between ages two and six years, $\alpha = .79$, and teenagers 13 to 22 years, $\alpha = .73$ in the US (Garfin, McCallon, & Cox, 1988), and those reported in India, $\alpha = .79$ (Russell et al., 2010), and Brazil, $\alpha = .82$ (Pereira, Riesgo, & Wagner, 2008), both with mixed samples of children and adolescents. Our coefficient is lower than the reported in the original validity study ($\alpha = .95$), but similar to studies in Korea $\alpha = .87$ (Shin & Kim, 1998), and Spain $\alpha = .89$ (García-Villamisar & Muela, 2000).

Most of the validity studies showed excellent internal consistency with Cronbach coefficients similar to the original study, α = .94 (Tachimori et al., 2003; Nordin, Gillberg, & Nydén, 1998; Saemundsen, Magnússon, Smári, & Sigurdardottir, 2003); only Akoury-Dirani, Alameddine, and Salamoun (2013), and Polaino-Lorente, García-Villamisar, and Muela (1991), reported higher Cronbach coefficients, α = .95 and α = .98, respectively.

Convergent validity

We demonstrated CARS has convergent validity with DSM-IV criteria through a high correlation coefficient *Rho* = .62 and *k* = .33, sensitivity (58.1%), specificity (76.6%), and concordance (71%). Our results are lower than others like Chlebowski, Green, Barton, and Fein, (2010), who reported *k* = .57, and Perry, Condillac, Freeman, Dunn-Geier, and Belair, (2005) who showed a concordance rate (88%). These results show good agreement with DSM-IV.

We also demonstrated evidence for convergent validity between CARS and ADI-R. The highest correlation with the ADI-R and domain B (r^s = .61) and moderate for the domains A, B, and C (r^s = .50). Our results are similar to those of Saemundsen et al. (2003), who showed a moderate correlation with domain B (.60), but, contrary to our results they showed a higher correlation with domain A (.81). Also, they found a 66.7% concordance using a cut-off CARS score of 30 with *k* = .40. Other studies compared the autism disorder diagnosis through the ADI-R and the CARS categorical diagnosis, reporting concordance of 85.7%, with *k* = .36 (Ventola et al., 2006), and *k* = .095, *p* = .486 (Rellini, Tortolani, Trillo, Carbone, & Montecchi, 2004).

Differences in the instrument administration, as a structured interview with parents as informants or as an observational tool to code the patient behavior, can explain low correlations between instruments (Saemundsen et al., 2003; Ventola et al., 2006; Pilowsky, Yirmiya, Shulman, & Dover, 1998). For the same reason, higher kappa values *k* = .55 and *k* = .60 are achieved when comparing the CARS and ADOS (both observational instruments) (Chlebowski et al., 2010; Ventola et al., 2006).

Our study demonstrated convergent validity through a high concordance of CARS and the DSM-5 (84.5%), high specificity (90.5%) but low sensitivity (46.7%), and low kappa (.36). These results are inconsistent with other studies showing better kappa coefficient such as Mayes et al. (2014) and Dawkins, Meyer, and Van Bourgondien (2016) and high levels of sensitivity (ranging from 81% to 100%) and specificity (ranging from 70% to 100%) for the CARS when using the DSM-5 compared to DSM-IV criteria. However, both studies were prospective, as Dawkins et al. (2016) included adults, and clinicians were not blind to the clinical diagnoses.

Overall our results show that even though CARS was developed before the DSM classification, it has convergent validity with both DSM-IV and DSM-5 classifications and provides evidence for an increased specificity with DSM-5 compared to the DSM-IV, which was a primary goal of the recent DSM-5 classification.

The discriminant analysis demonstrated the instrument better performance identifying autism and PDD-NOS groups. However, it is less useful in identifying the Asperger's syndrome. This result is consistent with Rellini et al. (2004) and García-López and Narbona, (2014) as it shows that the CARS did not discriminate between children without autism and children with Asperger's syndrome and a has a less discriminant power for PDD-NOS. However, the analysis of discriminant functions' centroids of our study supports the capacity to discriminate between autism and PDD-NOS, reporting that three items make the difference: non-verbal communication, level, and consistency of intellectual response and relationship with people.

Some authors have demonstrated that the total score can help distinguish between groups. For example, Chlebowski et al. (2010) demonstrated that children with classic autism scored higher than the PDD-NOS group *M* = 35.1, *SD* = 4.2 and *M* = 29, SD = 4, respectively. In turn, Tachimori et al. (2003) suggest that cut-off scores of 27 and 27.5 are better discriminating PDD-NOS from Asperger's syndrome, with a sensitivity of .65 and .71, and a cut-off CARS score of 30 to 30.5 differentiates autism well from PDD-NOS, with a sensitivity of .68 and .75 for each, for adolescents and adults. Mayes et al. (2009) suggested that a cut score of 25.5 is useful discriminating high functioning autism, with 95% sensitivity and 97% specificity.

A limitation of our study was its retrospective design, so the generalizability of the results is not possible.

In addition, the lack of IQ for the participants, decreased our ability to study the discriminant validity between severe autism and intellectual disabilities. Future studies should study the interobserver reliability.

This study shows the validity evidence of the Mexican Spanish version of the CARS, which is the most widely used instrument in Mexico, making this scale a valuable tool in the often-busy clinical environment. A new version of CARS2 (Schopler, Van Bourgondien, Wellman, & Love, 2010) is available and has the same items and scoring standards as the original CARS (Schopler et al., 1988). Also, Sanchez and Constantino (2020) provided a protocol to standardize the observation showing incremental validity of the instrument making it suitable for telehealth applications with potential use for assessing children in remote areas or during the pandemia of COVID-19.

Conflict of interest

Acknowledgements

REFERENCES

Akoury-Dirani, L., Alameddine, M., & Salamoun, M. (2013). Validation of the Lebanese Childhood Autism Rating Scale-Second Edition-Standard Version. *Research in Autism Spectrum Disorders*, *7*(9), 1097-1103. doi: 10.1016/j.rasd.2013.05.004

Albores-Gallo, L., Roldán-Ceballos, O., Villarreal-Valdes, G., Betanzos-Cruz, B. X., Santos-Sánchez, C., Martínez-Jaime, M. M., ... Hilton, C. L. (2012). M-CHAT Mexican Version Validity and Reliability and Some Cultural Considerations. *ISRN Neurology*, *2012*, 408694. doi: 10.5402/2012/408694

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders DSM-5* (5th Edition). Washington, DC: American Psychiatric Publishing.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: American Psychiatric Association.

Baron-Cohen, S., Allen, J., & Gillberg, C. (1992). Can autism be detected at 18 months? The needle, the haystack, and the CHAT. *The British Journal of Psychiatry: The Journal of Mental Science*, *161*, 839-843. doi: 10.1192/bjp.161.6.839

Breidbord, J., & Croudace, T. J. (2013). Reliability generalization for Childhood Autism Rating Scale. *Journal of Autism and Developmental Disorders*, *43*(12), 2855-2865. doi: 10.1007/s10803-013-1832-9

Chlebowski, C., Green, J. A., Barton, M. L., & Fein, D. (2010). Using the Childhood Autism Rating Scale to Diagnose Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *40*(7), 787-799. doi: 10.1007/s10803-009-0926-x

Constantino, J. N., & Gruber, C. (2005). *Social Responsiveness Scale (SRS) Manual*. Western Psychological Services.

Dawkins, T., Meyer, A. T., & Van Bourgondien, M. E. (2016). The Relationship Between the Childhood Autism Rating Scale: Second Edition and Clinical Diagnosis Utilizing the DSM-IV-TR and the DSM-5. *Journal of Autism and Developmental Disorders*, *46*, 3361-3368. doi: 10.1007/s10803-016-2860-z

Finch, H. (2009). Identification of Variables Associated With Group Separation in Descriptive Discriminant Analysis: Comparison of Methods for Interpreting Structure Coefficients. *The Journal of Experimental Education*, *78*(1), 26-52. doi: 10.1080/00220970903224602

Fombonne, E., Marcin, C., Bruno, R., Tinoco, C. M., & Marquez, C. D. (2012). Screening for Autism in Mexico. *Autism Research*, *5*(3), 180-189. doi: 10.1002/aur.1235

Fombonne, E., Marcin, C., Manero, A. C., Bruno, R., Diaz, C., Villalobos, M., ... Nealy, B. (2016). Prevalence of Autism Spectrum Disorders in Guanajuato, Mexico: The Leon survey. *Journal of Autism and Developmental Disorders*, *46*(5), 1669-1685. doi: 10.1007/s10803-016-2696-6

García-López, C., & Narbona, J. (2014). Inventario del espectro autista y childhood autism rating scale: Correspondencia con criterios DSM-IV-TR en pacientes con trastornos generalizados del desarrollo. *Anales de Pediatría*, *80*(2), 71-76. doi: 10.1016/j.anpedi.2013.05.012

García-Villamisar, D., & Muela, C. (2000). Propiedades Psicométricas de la Childhood Autism Rating Scale (CARS) como instrumento diagnóstico de los adultos autistas en el ámbito laboral. *Revista de Psicología General y Aplicada*, *53*(3), 515-521.

Garfin, D. G., McCallon, D., & Cox, R. (1988). Validity and reliability of the Childhood Autism Rating Scale with autistic adolescents. *Journal of Autism and Developmental Disorders*, *18*(3), 367-378. doi: 10.1007/BF02212193

Harris, B., & Barton, E. E. (2016). Autism services in Mexico: A qualitative survey of education professionals. *International Journal of School & Educational Psychology*, *5*(1), 1-13. doi: 10.1080/21683603.2016.1155514

Krug, D., Arick, J., & Almond, P. (1980). Behavior checklist for identifying severely handicapped individuals with high levels of autistic behavior. *Journal of Child Psychology and Psychiatry*, *21*(3), 221-229. doi: 10.1111/j.1469-7610.1980.tb01797.x

Krug, D., Arick, J., & Almond, P. (1993). *Autism Screening Instrument for Educational Planning-ASIEP-2*. PRO-ED.

Kurita, H., Miyake, Y., & Katsuno, K. (1989). Reliability and validity of the Childhood Autism Rating Scale-Tokyo version (CARS-TV). *Journal of Autism and Developmental Disorders*, *19*(3), 389-396. doi: 10.1007/BF02212937

Le Couteur, A., Haden, G., Hammal, D., & McConachie, H. (2008). Diagnosing Autism Spectrum Disorders in Pre-school Children Using Two Standardised Assessment Instruments: The ADI-R and the ADOS. *Journal of Autism and Developmental Disorders*, *38*(2), 362-372. doi: 10.1007/s10803-007-0403-3

Living-Varela, K., & Albores-Gallo, L. (2012). *Confiabilidad y Validez del Inventario de comportamiento autista (ABC) en niños con trastornos del espectro autista*. UNAM.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24*(5), 659-685. doi: 10.1007/BF02172145

Lord, C., Rutter, M., Dilavore, P., & Risi, S. (1999). *Autism Diagnostic Observation Schedule*. Western Psychological Services.

Mayes, S. D., Calhoun, S. L., Murray, M. J., Morrow, J. D., Yurich, K. K. L., Mahr, F., ... Petersen, C. (2009). Comparison of Scores on the Checklist for Autism Spectrum Disorder, Childhood Autism Rating Scale, and Gilliam Asperger's Disorder Scale for Children with Low Functioning Autism, High Functioning Autism, Asperger's Disorder, ADHD, and Typical Development. *Journal of Autism and Developmental Disorders*, *39*(12), 1682-1693. doi: 10.1007/s10803-009-0812-6

Mayes, S. D., Calhoun, S. L., Murray, M. J., Pearl, A., Black, A., & Tierney, C. D. (2014). Final DSM-5 under-identifies mild Autism Spectrum Disorder: Agreement between the DSM-5, CARS, CASD, and clinical diagnoses. *Research in Autism Spectrum Disorders*, *8*(2), 68-73. doi: 10.1016/j.rasd.2013.11.002

Montiel-Nava, C., Chacín, J. A., & González-Ávila, Z. (2017). Age of diagnosis of autism spectrum disorder in Latino children: The case of Venezuelan children. *Autism*, *21*(5), 573-580. doi: 10.1177/1362361317701267

Nordin, V., Gillberg, C., & Nydén, A. (1998). The Swedish version of the Childhood Autism Rating Scale in a clinical setting. *Journal of Autism and Developmental Disorders*, *28*(1), 69-75. doi: 10.1023/a:1026067104198

Pedersen, A., Pettygrove, S., Meaney, F. J., Mancilla, K., Gotschall, K., Kessler, D. B., ... Cunniff, C. (2012). Prevalence of Autism Spectrum Disorders in Hispanic and Non-Hispanic White Children. *Pediatrics*, *129*(3), e629-e635. doi: 10.1542/peds.2011-1145

Pereira, A., Riesgo, R. S., & Wagner, M. B. (2008). Childhood autism: Translation and validation of the Childhood Autism Rating Scale for use in Brazil. *Jornal de Pediatria*, *84*(6), 487-494. doi: 10.2223/JPED.1828

Perry, A., Condillac, R. A., Freeman, N. L., Dunn-Geier, J., & Belair, J. (2005). Multi-site Study of the Childhood Autism Rating Scale (CARS) in Five Clinical Groups of Young Children. *Journal of Autism and Developmental Disorders*, *35*(5), 625-634. doi: 10.1007/s10803-005-0006-9

Pilowsky, T., Yirmiya, N., Shulman, C., & Dover, R. (1998). The Autism Diagnostic Interview-Revised and the Childhood Autism Rating Scale: Differences Between Diagnostic Systems and Comparison Between Genders. *Journal of Autism and Developmental Disorders*, *28*(2), 143-151. doi: 10.1023/A:1026092632466

Polaino-Lorente, A., García-Villamisar, D., & Muela, C. (1991). The internal consistency and concurrent validity of Spanish traslation of CARS and AAPEP. *He Journal of Child Psychology and Psychiatry Abstracts Booklet*, *121*.

Rellini, E., Tortolani, D., Trillo, S., Carbone, S., & Montecchi, F. (2004). Childhood Autism Rating Scale (CARS) and Autism Behavior Checklist (ABC) Correspondence and Conflicts with DSM-IV Criteria in Diagnosis of Autism. *Journal of Autism and Developmental Disorders*, *34*(6), 703-708. doi: 10.1007/s10803-004-5290-2

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. doi: 10.1023/a:1010738829569

Russell, P. S., Daniel, A., Russell, S., Mammen, P., Abel, J. S., Raj, L. E., … Thomas, N. (2010). Diagnostic accuracy, reliability and validity of Childhood Autism Rating Scale in India. *World Journal of Pediatrics*, *6*(2), 141-147. doi: 10.1007/s12519-010-0029-y

Saemundsen, E., Magnússon, P., Smári, J., & Sigurdardóttir, S. (2003). Autism Diagnostic Interview-Revised and the Childhood Autism Rating Scale: Convergence and Discrepancy in Diagnosing Autism. *Journal of Autism and Developmental Disorders*, *33*(3), 319-328. doi: 10.1023/A:1024410702242

Sanchez, M. J., & Constantino, J. N. (2020). Expediting clinician assessment in the diagnosis of autism spectrum disorder. *Developmental Medicine & Child Neurology*, *62*(7), 806-812. doi: 10.1111/dmcn.14530

Schopler, E., Reichler, R. J., DeVellis, R. F., & Daly, K. (1980). Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). *Journal of Autism and Developmental Disorders*, *10*(1), 91-103. doi: 10.1007/BF02408436

Schopler, E., Reichler, R. J., & Renner, B. R. (1988). *The Childhood Autism Rating Scale (CARS)*. Western Psychological Services.

Schopler, E., Van Bourgondien, M. E., Wellman, J., & Love, S. (2010). *Childhood autism rating scale-second edition (CARS2): Manual*. Los Angeles: Western Psychological Services.

Shin, M. S., & Kim, Y.-H. (1998). Standardization study for the Korean version of the Childhood Autism Rating Scale: Reliability, validity and cut-off score. *Korean Journal of Clinical Psychology*, *17*(1), 1-15.

Sturmey, P., Matson, J. L., & Sevin, J. A. (1992). Brief report: Analysis of the internal consistency of three autism scales. *Journal of Autism and Developmental Disorders*, *22*(2), 321-328. doi: 10.1007/BF01058159

Tachimori, H., Osada, H., & Kurita, H. (2003). Childhood Autism Rating Scale-Tokyo Version for screening pervasive developmental disorders. *Psychiatry and Clinical Neurosciences*, *57*(1), 113-118. doi: 10.1046/j.1440-1819.2003.01087.x

Tapia Guillen, L. G., Solis Bravo, M. A., Albores-Gallo, L., Flores-Rodriguez, Y., Roldán-Ceballos, O., Aguerre-Estevez, R., & Martin-Salas, J. (2019). Autism Spectrum Disorders in an Upper-Middle Income Country. *Global Journal of Intellectual & Developmental Disabilities*, *6*(2), 1-7. doi: 10.19080/GJIDD.2019.06.555684

Ventola, P. E., Kleinman, J., Pandey, J., Barton, M., Allen, S., Green, J., … Fein, D. (2006). Agreement Among Four Diagnostic Instruments for Autism Spectrum Disorders in Toddlers. *Journal of Autism and Developmental Disorders*, *36*(7), 839-847. doi: 10.1007/s10803-006-0128-8

Zavaleta-Ramírez, P., Rosetti, M. F., Albores-Gallo, L., Vargas-Soberanis, M. A., López, O. N., & Medina-Mora, M. E. (2020). Pathways to a Diagnosis of Autism Spectrum Disorder. *Psychiatric Services*, *71*(11), 1120-1126. doi: 10.1176/appi.ps.201900518