Research Article

Accuracy of the Indonesian Childhood Autism Rating Scale Second Edition (CARS-2) as a Diagnosis Tool for Autism Spectrum Disorder

Achmad Sholeh

Growth Development Unit & Psychology Clinic, Rumah Sakit Dr. Oen Solo Baru, Indonesia

ORCID

Achmad Sholeh https://orcid.org/0000-0003-0408-8786

Abstract.

This study examined the diagnostic accuracy of the Indonesian Childhood Autism Rating Scale, Second Edition, Standard Version (ID-CARS2-ST) in identifying autism spectrum disorder (ASD) and differentiating it from other neurodevelopmental disorders. The ID-CARS2-ST can be used to develop diagnostic hypotheses and design intervention plans for children with ASD, complementing clinical judgment based on the DSM criteria. This study involved a total of 50 subjects in our clinic, consisting of 25 children diagnosed with ASD and 25 children with other neurodevelopmental disorders. We conducted reliability tests and receiver operating characteristic (ROC) curve analysis to assess the sensitivity and specificity of the ID-CARS2-ST to determine the optimal standard cutoff point for diagnosing ASD in children. Cronbach's alpha reliability yielded a result of α = 0.930 and McDonald's omega of ω = 0.937. The optimal ID-CARS2-ST cutoff score was 30, with a sensitivity of 96% and specificity of 88% (AUC = 0.927, P<0.001; Youden's J = 0.84). Our findings are exciting, and a score of 30 is the same as the original cut-off of the CARS2-ST in identifying ASD.

Keywords: CARS2-ST, ID-CARS2-ST, autism spectrum disorder, diagnose, cut-off score

Corresponding Author: Achmad Sholeh; email: zholeh.wei@gmail.com

Published: 27 March 2025

Publishing services provided by Knowledge E

Selection and Peer-review under the responsibility of the ICAP-H Conference Committee.

1. INTRODUCTION

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by persistent deficits in social communication and social interaction along with restricted and repetitive patterns of behavior, interests, or activities. This complex set of symptoms is present from early childhood and limits or interferes with daily functioning (1). The WHO reports that 1 in 100 children have autism (2). At the same time, reports from the Centers for Disease Control and Prevention (CDC) show that in the U.S., 1 in 36 children aged eight years have ASD, and the prevalence of ASD has increased 312% since 2000 with an average diagnosis at age four years (3–5). CDC findings based on race show as many as 33.4 cases per 1000 Asian and Pacific Islander children experience ASD, make it the most ethnic group experiencing ASD (3). In Indonesia,

there are no studies that precisely measure the prevalence of ASD. Still, the Ministry of Health, through its representative, explained that by 2024, it is estimated that as many as 2.4 million Indonesian children have ASD (6). Accurate and early identification of ASD is critical, particularly given the high prevalence, family and societal costs, and recognized importance of early intervention (7).

The Childhood Autism Rating Scale (CARS) is a diagnostic test tool to identify the presence of Autism Spectrum Disorder (ASD) that was first published in 1988 (8). Professionals have widely used CARS for clinical and research purposes worldwide, including in Indonesia. The second edition or CARS-2, was then published in 2010 by presenting various updates which include three forms, CARS2-ST (Standard), CARS2-HF (High Function), and CARS2-QPC (Questionnaire for Parents or Caregivers) (9). As of this edition, CARS-2 can not only be used to develop diagnosis hypotheses but also to design interventions through the 15 areas examined in it (9).

Pediatricians have been relying on first-stage screening tests such as M-CHAT R/F to diagnose ASD, while Clinical Psychologists more commonly use CARS. In contrast to the U.S. and other countries that consider ADOS and ADI-R as gold standards, accessing these tests here tends to be difficult as they require lengthy and expensive training, and there is no process for adapting the tests into Bahasa Indonesia. On the other hand, Kaufman (2022) also criticized the use of ADOS and ADI-R, which are considered the “gold standard” of ASD diagnosis where various scientific evidence proves that the interscorer value of this test tool is always poor, the term “gold standard” embedded in the two tests seems to pressure professionals to always use it in diagnosing ASD because of insurance and other regulatory needs, this seems to ignore other test tools in helping ASD diagnosis more cost-effective, more accessible, and often more time-efficient such as CARS-2. There is no need to limit what assessment tools can be used in ASD diagnosis.

Although the diagnosis of ASD in DSM-5 does not require standardized tests like the diagnosis of Intellectual Disability, tests are often used to help develop hypotheses and improve diagnostic decision-making in ASD (10). All assessments are affected by bias, so mental health professionals need to conduct autism assessments that minimize bias, no matter what tests or tools they do or do not use to address it (10–12). The general principles of assessment are essential to avoid bias, so we need psychological and neuropsychological assessments that have long been devoted to measuring abnormal human behaviour using various methods (13) and followed by clinical decisions from professionals and professional teams (10).

We recognize that diagnosing ASD can be challenging and that professionals require significant experience. The lack of training of professionals in diagnosing ASD, as well as the fact that many mental health professionals only use the first stage of screening tests, makes it difficult to diagnose ASD. Therefore, using the CARS2-ST diagnostic test can be a tool to develop a diagnosis hypothesis and assist in making clinical judgments. This study investigated the optimal cutoff point of the Indonesian Version of CARS2-ST (ID-CARS2-ST) to identify and differentiate ASD from other neurodevelopmental disorders. The findings may help improve the reliability of ASD diagnosis using the relatively easy-to-implement ID-CARS2-ST.

2. METHOD

2.1. Participants

This study was conducted on 50 children with an age range of 20-89 months who were tested using the ID-CARS2-ST in our Outpatient Growth and Development Unit from 1 September 2023 to 23 July 2024 by Clinical Psychologists trained in ASD diagnoses and supported by a team that includes Occupational Therapist and Speech Language Pathologist. The sampling technique used is purposive sampling with quota sampling type, which is sampling by setting a certain amount as a target (quota) that must be met in sampling from the population (14). This technique was chosen so that the number of subjects would be proportional: 50% (25) children diagnosed with ASD based on DSM-V criteria and 50% (25) children with other neurodevelopmental disorders. This study consisted of 25 children diagnosed with ASD based on DSM-V diagnostic criteria and 25 children with other neurological disorders (mean age 51.04 months; 20-89 months).

2.2. Instrument

The Childhood Autism Rating Scale Second Edition (CARS2-ST) (Standard Version) is a diagnostic tool used to identify ASD and design intervention plans. CARS2-ST is used to assess individuals with an estimated overall IQ of 79 or lower, those with communication disorders, or those under the age of 6, regardless of their estimated IQ. This study used the Indonesian Version of CARS2-ST (ID-CARS2-ST), which was adapted into the Indonesian Language by Sary et al. (2021) and produced two factors by DSM-5 criteria, the social communication-interaction factor and the restrictive stereotyped behaviours

Table 1: Demographic characteristics.

Characteristics	n (%)
Sex	
Male	39 (78)
Female	11 (22)
Age (month)	
20 – 48	25 (50)
48 – 72	16 (32)
72 – 89	9 (18)
Diagnosis	
Autism Spectrum Disorder (ASD)	25 (50)
Attention Deficit Hyperactivity Disorder (ADHD)	2 (4)
Global Developmental Disorder (GDD)	5 (10)
Social Communication Disorder (SCD)	8 (16)
Expressive Language Disorder	7 (14)
Specific Learning Disorder	1 (2)
Other Mental & Behavioral Disorder	2 (4)

and sensory sensitivities factor. The original CARS2-ST has an internal consistency of 0.93 (9). It has high reliability and diagnostic validity by DSM-IV and DSM-V diagnosis criteria (15).

The test measures 15 areas, which include 1) Relating to people, 2) Imitation, 3) Emotional response, 4) Body use, 5) Object use, 6) Adaptation to change; 7) Visual response; 8) Listening response; 9) Taste, smell, and touch response and use; 10) Fear or nervousness; 11) Verbal communication; 12) Nonverbal communication; 13) Activity level; 14) Level and consistency of intellectual response; and 15) general impressions (9). The 15 items of CARS-2 are rated using a scale from 1 (within normal limits for that age) to 4 (severely abnormal for that age), and the scores from each question are added to obtain a total score for each patient, ranging from 15 to 60. A total score of 15-29.5 indicates minimal to no symptoms of ASD, a score of 30-36.5 indicates mild to moderate symptoms of ASD, and a score of 37 or above indicates severe symptoms of ASD.

A similar study conducted in South Korea found that the optimal cut-off for CARS2-ST was 30(16). However, another study there recommended a new lower cut-off of 28.5 with sensitivity (0.94) and specificity (1.00), whereas using the conventional cut-off of 30 would result in misdiagnosis based on comparison with ADOS-2 (17). An even lower cut-off score of 26 was also recommended for detecting ASD in Lebanese and Arab

populations for children aged 2-12 years (18,19) A score of 26 was considered as accurate as the DSM-5 in detecting ASD, yielding high sensitivity (0.96) but maintaining moderate specificity (0.70) (19).

2.3. Statistical Analysis

We conducted a Receiver Operating Curve (ROC) analysis to illustrate the trade-off between sensitivity and specificity, determine diagnostic accuracy, and establish the best cut-off point for clinical diagnosis using ID-CARS-2 with a sample of children in Indonesia. The thresholds we used for sensitivity and specificity were $\geq$ 0.60 (high validity), 0.40 – 0.59 (moderate validity), and $<$ 0.40 (low validity) (20) . The classification for the Area Under the Curve (AUC) was 0.5 – 0.6 (very weak accuracy), 0.6 – 0.7 (weak accuracy), 0.7 – 0.8 (moderate accuracy), 0.8 – 0.9 (high accuracy), and 0.9 – 1 (very high accuracy) (21).

3. RESULTS

3.1. Internal Consistency

The internal consistency results of the ID-CARS2-ST demonstrate promising results, with values ranging from 0.462 to 0.809, with a Cronbach's alpha of α = 0.930 and McDonald's omega of ω = 0.937. Item 9, related to sensory issues of taste, smell, and touch response and use, obtained the lowest value of 0.462, while item 2, related to imitation, obtained the highest value of 0.809. The mean of the total ID-CARS2-ST score overall is 29.52, with children with ASD having a mean of 36.38 and children with SCD having a mean of 28.06.

3.2. Optimal Diagnostic Cut-off Using the ID-CARS2 ST

The following are the results of the Receiver Operating Curve (ROC) analysis to determine the optimal cut-off point of the ID-CARS2-ST in diagnosing ASD based on DSM-V criteria and differentiating ASD from non-ASD or other neurodevelopmental disorders.

The ROC analysis results showed that the most optimal cut-off point was a score of 30, with a sensitivity of 0.960 (96%) and a specificity of 0.880 (88%), with an AUC value = 0.927; P>0.001. The accuracy, which measures how often the model can predict

Table 2: Internal consistency.

Item	Mean	Item-Total Correlation
1. Relating to people	2.340	0.793
2. Imitation	2.440	0.809
3. Emotional response	1.970	0.785
4. Body use	1.820	0.673
5. Object use	1.520	0.637
6. Adaptation to change	1.790	0.651
7. Visual response	1.730	0.694
8. Listening response	1.950	0.668
9. Taste, smell, and touch response and use	1.160	0.462
10. Fear or nervousness	1.820	0.485
11. Verbal communication	2.920	0.666
12. Nonverbal communication	1.990	0.733
13. Activity level	2.460	0.489
14. Level and consistency of intellectual response	1.620	0.663
15. General impressions	1.990	0.784

Table 3: Sensitivity and specificity ID-CARS2-ST cut-off scores for diagnoses based on DSM-V.

ID-CARS2-ST *Cut-off*	Sensitivity	Specificity	Youden's J
23.75	0.960	0.680	0.640
25.50	0.960	0.760	0.720
27.50	0.960	0.800	0.760
29.00	0.960	0.840	0.800
30.00*	0.960*	0.880*	0.840*
30.75	0.880	0.880	0.760
31.75	0.800	0.880	0.680
32.75	0.720	0.880	0.600
33.50	0.680	0.880	0.560

** Optimal ID-CARS2-ST cut-off for ASD diagnosis.*

correctly, obtained a value of 0.920 in this model, and the precision value for accurately predicting ASD obtained a value of 0.889. The overall combined accuracy rate, based on the Youden's J Index (22,23) is 0.840 (Youden's J Index [0.960 + 0.880] – 1 = 0.840). This value is the probability of making a correct diagnostic decision about ASD or an 84% overall probability of being accurate. The sensitivity and specificity become

unbalanced if the cut-off is shifted up or down. Therefore, point 30 is considered the most optimal cut-off score.

4. DISCUSSION

This study examined the optimal diagnostic cut-off score for the Indonesian version of the CARS-2 in diagnosing ASD according to DSM-V criteria. Our findings successfully replicated the original cut-off score of 30, which is categorized as the likelihood of having ASD (Autism, Asperger's Disorder, or Pervasive Developmental Disorder Not Otherwise Specified [PDD-NOS]) (9). Our results demonstrate high sensitivity, specificity, and diagnostic accuracy in identifying ASD, with a value of 0.84 based on Youden's J, indicating an 84% accuracy in making correct diagnostic decisions about ASD based on DSM-V criteria.

These results are also consistent with a study conducted in South Korea, where the optimal cut-off for CARS2-ST was found to be 30 (16). However, other studies in South Korea and Lebanese/Arab populations have suggested lower cut-off scores of 28.5 (based on comparison with ADOS-2) and 26, respectively (17,19). These discrepancies may be attributed to differences in the characteristics of the populations studied, highlighting the importance of determining the optimal cut-off score for specific cultural and population contexts. Factors such as ASD prevalence, population characteristics, and cultural norms can influence the manifestation of ASD symptoms.

Our study demonstrates that the ID-CARS2-ST is valid and reliable for identifying ASD. The cut-off score of 30 can be used as a guideline for clinical diagnosis, but a comprehensive clinical assessment based on DSM-5 criteria remains essential. Clinical interviews, behavioral observations, and information from parents or caregivers are still crucial for a holistic understanding of the child's functioning and needs. The ID-CARS2-ST can complement clinical assessment by providing a standardized measurement of ASD symptoms and aiding in designing individualized intervention plans.

Additionally, we recommend that scores of 23 – 29 be considered in identifying SCD in children. This recommendation is based on the movement of scores in our sample and the average score obtained by children with SCD. In the Arab population, it was confirmed that children with SCD obtained an average score of 23 (19). However, further studies need to be conducted specifically on ID-CARS2-ST to differentiate ASD from SCD, given the similarities between the two disorders, with the only difference

being repetitive behaviors and restricted interests. In practice, these two conditions are sometimes very difficult to distinguish, and it is unclear whether SCD is truly separate from ASD.

Although this study provides evidence supporting the use of ID-CARS2-ST in ASD diagnosis, several limitations should be noted. First, the sample size is relatively small (n=50) and may not represent the entire Indonesian child population. Second, this study focused on differentiating ASD from other neurodevelopmental disorders, and future research should investigate the ability of the ID-CARS2-ST to differentiate ASD from other disorders with similar symptoms, particularly Social Communication Disorder (SCD), as we observed difficulties in distinguishing between ASD level 1 and SCD.

5. CONCLUSION AND IMPLICATIONS

In conclusion, this study provides strong evidence that the ID-CARS2-ST is a valid and reliable tool for identifying ASD. The cut-off score of 30 can be used as a guideline in clinical diagnosis. This study contributes to clinical practice in Indonesia by providing an accurate and efficient tool to identify ASD in children. Accurate diagnosis is crucial for early intervention, which can significantly improve long-term outcomes for children with ASD. By utilizing the ID-CARS2-ST in conjunction with comprehensive clinical assessment, professionals can make informed decisions about diagnosis and develop effective individualized intervention plans.

Acknowledgements

The authors would like to express our gratitude to all the children and Rumah Sakit Dr. Oen Solo Baru for their support in this study.

Conflict of interest

The authors declare no conflict of interest.

Funding

The author received financial support for this research and publication from Rumah Sakit Dr. Oen Solo Baru.

References

[1] American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-5-TR) (Fifth Edit). 2022.

[2] World Health Organization. Autism. 2023.

[3] Maenner MJ, Warren Z, Williams AR, Amoakohene E, Bakian AV, Bilder DA, et al. Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2020. MMWR Surveill Summ. 2023 Mar;72(2):1–14.

[4] Zablotsky B, Black LI, Maenner MJ, Schieve LA, Danielson ML, Bitsko RH, et al. Prevalence and Trends of Developmental Disabilities among Children in the United States: 2009-2017. Pediatrics. 2019 Oct;144(4):e20190811.

[5] Shaw KA, Maenner MJ, Bakian AV, Bilder DA, Durkin MS, Furnier SM, et al. Early Identification of Autism Spectrum Disorder Among Children Aged 4 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2018. MMWR Surveill Summ. 2021 Dec;70(10):1–14.

[6] Stefanni DM. detikHealth. [Juta Anak di Indonesia Idap Autisme.]. Wamenkes Ungkap. 2024;2:4.

[7] Zero to Three. DC:0–5TM: Diagnostic classification of mental health and developmental disorders of infancy and early childhood. Washington DC; 2016.

[8] Schopler E, Reichler R, Renner B. Childhood Autism Rating Scale (CARS) [Manual]Western Psychological Services; 1988.

[9] Schopler E. Bourgondien ME, Van, Wellman GJ, Love SR. Childhood Autism Rating Scale, Second Edition (CARS2) [Manual]. Torrance: Western Psychological Services.; 2010.

[10] Kaufman NK. Rethinking “gold standards” and “best practices” in the assessment of autism. Appl Neuropsychol Child. 2022;11(3):529–40.

[11] Kaufman N, Bush S. Validity Assessment in Military Psychology. In 2020. p. 211.

[12] Kaufman NK, Bush SS. Concussion in Pediatric Neuropsychology. J Pediatr Neuropsychol. 2020 Mar;6(1):14–26.

[13] Wolff M, Bridges B, Denczek T. The complexity of autism spectrum disorders. Routledge. Routledge; 2018. https://doi.org/10.4324/9780429454646.

[14] Kasiram M. Metode Penelitian Kuantitatif. UINPress; 2008.

[15] Dawkins T, Meyer AT, Van Bourgondien ME. The Relationship Between the Childhood Autism Rating Scale: Second Edition and Clinical Diagnosis Utilizing the DSM-IV-TR and the DSM-5. J Autism Dev Disord. 2016 Oct;46(10):3361–8.

[16] Lee S, Yoon SA, Shin MS. Validation of the Korean Childhood Autism Rating Scale-2. Res Autism Spectr Disord. 2023 May;103:102128.

[17] Ji SI, Park H, Yoon SA, Hong SB. A Validation Study of the CARS-2 Compared With the ADOS-2 in the Diagnosis of Autism Spectrum Disorder: A Suggestion for Cutoff Scores. J Korean Acad Child Adolesc Psychiatry. 2023 Jan;34(1):45–50.

[18] Akoury-Dirani L, Alameddine M, Salamoun M. Validation of the lebanese childhood autism rating scale-second edition-standard version. Res Autism Spectr Disord. 2013;7(9):1097–103.

[19] Alotaibi B, Alotaibi A. Exploratory and confirmatory factor analyses of the Arabic version of the Childhood Autism Rating Scale. Res Autism Spectr Disord. 2021 Aug;86:86.

[20] Nelson DE, Holtzman D, Bolen J, Stanwyck CA, Mack KA. Reliability and validity of measures from the behavioral risk factor surveillance system (BRFSS). Soz Praventivmed. 2001;46(1 Suppl 1):S3–42.

[21] Metz CE. Basic principles of ROC analysis. Semin Nucl Med. 1978 Oct;8(4):283–98.

[22] Biggersta BJ. Comparing diagnostic tests: a simple graphic using likelihood ratios. Volume 19. STATISTICS IN MEDICINE Statist. Med; 2000.

[23] Youden WJ. Index for rating diagnostic tests. Cancer. 1950 Jan;3(1):32–5.