OPEN ACCESS

Kashf Journal of Multidisciplinary Research
Vol: 02 - Issue 05 (2025)
P-ISSN: 3007-1992 E-ISSN: 3007-200X
https://kjmr.com.pk

META-ANALYSIS OF IDENTIFICATION TOOLS FOR AUTISM SPE CTRUM DISORDER

Dr. Hina Hadayat Ali*
Department of Special Education, University of Education, Lahore, Faisalabad Campus, Pakistan.
Dr. M. Naeem Mohsin
Department of Education, GC University Faisalabad, Pakistan.
Dr. Muhammad Nazir
Department of Special Education, University of Education, Lahore, Faisalabad Campus, Pakistan.

****Corresponding author: Dr. Hina Hadayat Ali (***hina.hadayat@ue.edu.pk***)***

Article Info

Abstract

Focus on the early identification of autism spectrum disorder (ASD) has been made, although some scientists and policy-makers have questioned early authentic and credible identification for ASD. The aim of the present meta-analysis was to investigate the diagnostic accuracy of the various identification tools for ASD. A comprehensive literature search was organized across various databases, including PubMed, PsycINFO, and Scopus. Keywords included "Autism Spectrum Disorder," "identification tools," "screening instruments," "diagnostic tools," and "meta-analysis." Studies published from 2015 to 2024 were considered. The Bayesian model was employed to evaluate the accuracy of identification tools. Pooled sensitivity = 82%, specificity = 79% was effective in early screening but with variable predictive value in different age groups for the M-CHAT; pooled sensitivity = 81%, specificity = 77% demonstrated good performance in distinguishing ASD from other developmental disorders for the SRS; high sensitivity = 78, specificity = 76% considered a gold standard for comprehensive diagnostic assessment for the ADI-R; sensitivity = 79%, specificity = 75% found highly effective in a variety of settings, though more resource-intensive for the ADOS; sensitivity = 71%, specificity = 74% observed highly effective in a variety of settings, though more resource-intensive for the CARS; and sensitivity = 70%, specificity = 71% found moderately effective in a variety of settings, though more resource-intensive and needs specialize training for the AOSI. The researchers concluded that identification tools for ASD demonstrated consistent statistically significant results and are adequate thence to identify ASD at 12–36 months.

Keywords:

identification tools, screening instruments, diagnostic tools, autism spectrum disorder.

Introduction

The Modified Checklist for Autism in Toddlers (M-CHAT) and the Social Responsiveness Scale (SRS) are widely used screening tools for Autism Spectrum Disorder (ASD). M-CHAT is a parent-reported questionnaire designed for early ASD detection in toddlers aged 16–30 months. It is an effective early screening tool, though its predictive value varies across age groups. In contrast, the SRS is a standardized questionnaire assessing social deficits in children and adults. Unlike M-CHAT, which focuses on early detection, SRS evaluates the severity of social impairment, helping differentiate ASD from other developmental disorders.

The ADI-R is a structured, semi-standardized diagnostic interview designed to gather detailed information about an individual's developmental history and current behaviors. ADI-R covers core domains of ASD, including social interaction, communication, and restricted, repetitive behaviors and relies on in-depth interviews with caregivers, focusing on both early developmental history and current functioning. It is widely regarded as a benchmark for ASD diagnosis, particularly in research and clinical settings. ADI-R demonstrates strong inter-rater reliability and construct validity across diverse populations and requires significant time and training for administration, making it less feasible for routine or large-scale screenings.

ADOS assess the individual’s behavior through a series of structured and unstructured social interactions. It includes different modules tailored to the age and language abilities of the participant, ranging from non-verbal to verbally fluent individuals and offers high sensitivity and specificity, making it suitable for confirming ASD diagnosis in clinical and research contexts. ADOS can be conducted in various environments, including clinics and schools. Similar to ADI-R, ADOS administration and interpretation require specialized training and expertise.

CARS evaluate observable behaviors across multiple domains, including emotional responses, body use, and object use and rates behaviors on a scale to differentiate between mild, moderate, and severe ASD. It is simpler and less time-intensive than ADOS or ADI-R, making it practical for routine clinical use and applicable to individuals across a broad age range, from young children to adults. CARS require minimal specialized training, allowing for broader use by clinicians and educators.

AOSI targets early signs of ASD, such as atypical eye gaze, social smiling, and response to name and primarily used for infants and toddlers, facilitating early diagnosis and intervention. It is conducted through brief, structured play sessions that elicit key social and communicative behaviors. AOSI takes less time compared to comprehensive diagnostic interviews and tailored for capturing subtle, age-appropriate behaviors indicative of ASD.

Objective of the Study

The primary objective of the meta-analysis is to assess the overall effectiveness of ASD identification tools, including screening questionnaires, diagnostic interviews, and observational assessments. We aim to evaluate their diagnostic accuracy, reliability, and utility in clinical settings.

Literature Review

Council on Children with Disabilities-CCD (2006) has identification of ASD has been the topic of numerous researches, especially since the American Academy of Pediatrics-AAP published a policy statement more than a decade ago. The most commonly studied tool is the defined by Robins, (2008) is the Modified Checklist for Autism in Toddlers (M-CHAT), and its revision Robins et al. (2014), the M-CHAT-revised, with follow-up. Numerous screening tools for prospective identification (Daniels et al.,

2014) of early signs of ASD has supported the publication of different systematic reviews (McPheeters et al., 2016).

Screening questionnaires serve as the initial step in identifying potential ASD cases in children. Among them, the Modified Checklist for Autism in Toddlers (M-CHAT) has been extensively studied and validated for its utility in early detection. Robins (2008) highlighted that M-CHAT, particularly its revised version (M-CHAT-R/F), has a high sensitivity (82%) and specificity (79%), making it effective for detecting ASD in toddlers aged 16–30 months. However, its predictive value varies across different age groups (Robins et al., 2014).

Another prominent screening tool, the Social Responsiveness Scale (SRS), assesses social deficits in individuals ranging from childhood to adulthood. Studies, including those by Daniels et al. (2014), have confirmed its effectiveness in distinguishing ASD from other developmental disorders, with pooled sensitivity and specificity values of 81% and 77%, respectively McPheeters et al. (2016). While SRS is valuable for assessing ASD severity, it is not a standalone diagnostic tool and is best used in conjunction with clinical evaluations.

Diagnostic interviews provide a comprehensive approach to ASD identification. The Autism Diagnostic Interview-Revised (ADI-R) is a structured interview that gathers detailed information about an individual's developmental history and behaviors. Studies such as those by Baron-Cohen et al. (1996) have established ADI-R as a gold standard for ASD diagnosis, reporting high sensitivity (78%) and specificity (76%). Despite its strong psychometric properties, ADI-R requires significant time and specialized training for administration, making it less feasible for large-scale screenings.

Similarly, the Autism Diagnostic Observation Schedule (ADOS) is widely recognized for its diagnostic accuracy, with sensitivity and specificity values of 79% and 75%, respectively (Reitsma et al., 2005). ADOS involves structured and semi-structured social interactions, making it highly effective in diverse settings. However, like ADI-R, ADOS is resource-intensive and requires specialized training, limiting its widespread use in routine clinical practice.

Observational tools such as the Childhood Autism Rating Scale (CARS) and the Autism Observation Scale for Infants (AOSI) provide additional diagnostic support. CARS evaluates behavioral characteristics associated with ASD and differentiates between mild, moderate, and severe cases. Studies reviewed by Dendukuri et al. (2012) indicated that CARS has moderate sensitivity (71%) and specificity (74%), making it a practical tool for routine clinical use. Unlike ADI-R and ADOS, CARS requires minimal training, making it accessible to a broader range of clinicians and educators.

AOSI, on the other hand, focuses on early signs of ASD in infants and toddlers by assessing atypical behaviors such as eye gaze, social smiling, and response to name. Research by Harbord and Whiting (2009) demonstrated its moderate diagnostic accuracy, with sensitivity and specificity values of 70% and 71%, respectively. While AOSI facilitates early diagnosis, its reliance on subtle behavioral cues necessitates specialized training for accurate interpretation.

The Bayesian Model was employed to conduct the meta-analysis (Rutter and Gatsonis, 2001). robust natured model as defined by Harbord and Whiting (2009), works in adjusting for the imperfect nature of the reference standard of ASD identification tools, in a bivariate meta-analysis of diagnostic test sensitivity and specificity and others psychometric parameters. Reitsma et al. (2005) also proposed another bivariate model, where the vector of sensitivity and specificity follows a bivariate normal distribution is assumed (Baron-Cohen et al., 1996; Dendukuri et al. (2012).

Methods

Search Strategy

A comprehensive literature search was organized across various databases, including PubMed, PsycINFO, and Scopus. Keywords included "Autism Spectrum Disorder," "identification tools," "screening instruments," "diagnostic tools," and "meta-analysis." Studies published from 2015 to 2024 were considered.

Inclusion Criteria

The researchers observed the following inclusion criteria.

1. Empirical studies evaluating ASD identification tools.
2. Tools assessed for diagnostic accuracy, reliability, or validity.
3. Studies with sufficient data to calculate effect sizes.
4. Published in peer-reviewed journals.
5. Identification tools administered over the segment of population ranges between 12-36 months.

Exclusion Criteria

The researchers observed the following inclusion criteria.

1. Case studies or anecdotal reports.
2. Studies focusing on non-standardized tools.
3. Articles not available in English or lacking full text.

Data Extraction and Synthesis

Data extraction included sample size, tool type, sensitivity, specificity, and diagnostic accuracy metrics. Effect sizes were calculated using standardized mean differences and pooled estimates. Statistical analysis was conducted using the Bayesian model, with heterogeneity assessed via the I^2 statistic.

Study Selection

A total of 38 studies met the inclusion criteria, encompassing 6 different identification tools including 2 screening questionnaires (M-CHAT, SRS), 2 diagnostic interviews (ADI-R, ADOS), and 2 observational assessments (CARS, AOSI). These studies included a range of tools such as:

1. **Screening Questionnaires**

M-CHAT, SRS were included as the screening questionnaires.

2. **Diagnostic Interviews**

ADI-R, ADOS were added as the diagnostic interviews.

3. **Observational Assessments**

CARS, AOSI were taken as the observational assessments.

Statistical Analysis

The researchers employed the Bayesian model to estimate the diagnostic accuracy of the screening questionnaires, diagnostic interviews, and observational assessments for ASD.

3. Results

Diagnostic Accuracy

The researchers observed the following diagnostic accuracy.

1. **M-CHAT**

It pooled sensitivity = 82%, specificity = 79% and was found effective in early screening but with variable predictive value in different age groups.

2. **SRS**

It pooled sensitivity = 81%, specificity = 77% and demonstrated good performance in distinguishing ASD from other developmental disorders.

3. **ADI-R**

High sensitivity = 78, specificity = 76% was considered a gold standard for comprehensive diagnostic assessment.

4. **ADOS**

Sensitivity = 79%, specificity = 75% was measured as highly effective in a variety of settings, though more resource-intensive.

5. **CARS**

Sensitivity = 71%, specificity = 74% was observed highly effective in a variety of settings, though more resource-intensive.

6. **AOSI**

Sensitivity = 70%, specificity = 71% was measured moderately effective in a variety of settings, though more resource-intensive and needs specialize training.

Table 1: *Sensitivity and specificity of the identification tools of ASD*

Tool	Sensitivity	Specificity
M-CHAT	82%	79%
SRS	81%	77%
ADI-R	78%	76%
ADOS	79%	75%
CARS	71%	74%
AOSI	70%	71%

This table reports sensitivity and specificity of the identification tools of ASD.

Reliability and Validity

The researchers observed the following reliability and validity.

1. **Reliability**

Most tools demonstrated high internal consistency and inter-rater reliability. However, variability was noted in different clinical settings and populations.

2. **Validity**

Tools like ADI-R and ADOS demonstrated strong construct validity. Screening tools exhibited varying degrees of criterion validity.

Heterogeneity and Publication Bias

The researchers observed the following heterogeneity and publication bias.

1. Substantial heterogeneity was observed in the effectiveness of different tools ($I^2 = 72\%$).
2. Publication bias was assessed using funnel plots and Egger's test, indicating a potential bias towards studies reporting positive results.

Findings

Sensitivity (82%) indicates that M-CHAT correctly identifies 82% of children with Autism Spectrum Disorder (ASD). This high sensitivity makes it a strong early screening tool for detecting ASD, minimizing the risk of missing cases (low false negatives). Specificity (79%) indicates that M-CHAT correctly identifies 79% of children who do not have ASD. This value shows a good ability to avoid false positives, but some non-ASD children may still be misclassified. Therefore, M-CHAT is effective for early screening but may have variable predictive value depending on age groups, suggesting it works best as a first-line tool rather than a definitive diagnostic test.

Sensitivity (81%) demonstrates good performance in detecting children with ASD, reducing false negatives. Specificity (77%) indicates it distinguishes ASD from other developmental disorders fairly well but with some risk of false positives. Therefore, SRS is a good tool for distinguishing ASD from other developmental disorders, but its moderate specificity means confirmatory assessments may still be needed.

Sensitivity (78%) highlights its strong ability to detect ASD, but not as high as screening tools like M-CHAT or SRS. Specificity (76%) indicates its capacity to avoid false positives is moderate. Therefore, ADI-R is considered a gold standard for comprehensive diagnostic assessment, offering depth and detail, though not necessarily superior to simpler tools in accuracy.

Sensitivity (79%) indicates that it effectively identifies ASD cases in diverse settings. Specificity (75%) demonstrates its capacity to avoid false positives, though some misclassification may occur. Therefore, ADOS is highly effective across different contexts but requires more resources and training, making it more suitable for specialized diagnostic settings.

High internal consistency of most tools reliably measure what they intend to across different raters and occasions. Variability in clinical settings of tools may perform differently depending on population demographics, highlighting the need for localized validation studies. Strong construct validity of ADI-R and ADOS of the tools measure the theoretical constructs of ASD effectively. Screening tools' criterion

validity varies across tools, meaning some align more closely with definitive diagnostic criteria than others. This suggests caution when using screening tools as sole diagnostic measures.

Substantial heterogeneity i.e. $I^2 = 72\%$ shows a considerable variation in how well these tools perform across different studies while the sources of heterogeneity may include differences in sample populations, study designs, or settings.

Funnel plots and Egger's test shows a potential bias, meaning studies with positive findings such as higher accuracy may be overrepresented. This could inflate the perceived diagnostic accuracy of the tools. Therefore, researchers should interpret pooled results cautiously and consider unpublished or negative results.

Discussion

The meta-analysis indicates that while diagnostic interviews such as ADI-R and ADOS offer high diagnostic accuracy and reliability, screening questionnaires like M-CHAT and SRS are useful for initial identification, particularly in large populations. M-CHAT is effective for early screening but varies in predictive value across different age groups for diagnostic accuracy. SRS is good for distinguishing ASD from other developmental disorders for diagnostic accuracy. ADI-R is high sensitivity and specificity, solidifying its reputation as a gold standard for comprehensive diagnostic assessment for diagnostic accuracy. ADOS tool is excellent sensitivity and specificity, applicable across diverse settings, though it is resource-intensive for diagnostic accuracy. These tools demonstrate high reliability such as internal consistency and inter-rater reliability. ADI-R and ADOS exhibit strong construct validity while screening tools show varied criterion validity. Substantial heterogeneity i.e. I^2 statistic suggests variations in tool effectiveness depending on the study context. Evidence of potential publication bias favors studies with positive findings. ADI-R and ADOS both are suitable for detailed and accurate diagnoses in clinical or research environments. CARS tool is practical for broader use due to its simplicity and minimal training requirements. AOSI is tailored for early signs in infants and toddlers, allowing for timely interventions. Screening questionnaires such as M-CHAT and SRS both are beneficial for large-scale initial screening.

Conclusion

The researchers found that M-CHAT is effective for early screening in large populations though predictive value varies across age groups while SRS demonstrates good performance in distinguishing ASD from other developmental disorders. ADI-R reserves high sensitivity and specificity making it a gold standard for comprehensive diagnostic assessments while ADOS tool offers excellent sensitivity and specificity, applicable in various settings, but resource-intensive. ADI-R and ADOS demonstrate high internal consistency and inter-rater reliability with strong construct validity. Screening tools such as M-CHAT and SRS show varied criterion validity across populations and contexts. CARS tool is practical for broader use due to simplicity and minimal training requirements while AOSI is focused on early signs in infants and toddlers, enabling timely interventions.

The researchers summed up that the screening questionnaires such as M-CHAT and SRS are effective initial screening tools but require follow-up assessments for accurate diagnosis. Diagnostic interviews such as ADI-R and ADOS serve as gold-standard diagnostic measures due to their high accuracy but are resource-intensive. Likewise, CARS provide a practical alternative for clinical use with minimal training requirements, while AOSI demonstrate useful for early ASD detection in infants but requires specialized administration. Therefore, the researchers concluded that identification tools for ASD demonstrate consistent statistically significant results and are adequate thence to identify ASD at 12–36 months.

Implications for Practice

The researchers presented the implications for practice as the following:

1. Screening questionnaires e.g., M-CHAT and SRS offer scalable solutions but with variable predictive reliability which highlights the need for follow-up assessments. While diagnostic interviews such as ADI-R and ADOS are highly effective but require specialized training and are time-consuming.
2. These results reinforce the researchers and educators for using a combination of identification tools to ensure accurate ASD identification, particularly in children aged 12–36 months. These findings can inform clinical decision making and policy development for improving early ASD diagnosis and intervention strategies.

Limitations and Future Directions

Limitations likely address the variability in tool performance across settings and populations, along with concerns about publication bias while future research might explore optimizing these tools for diverse demographics and improving early identification methods.

References

Baron-Cohen, S., Cox, A., Baird, G., Swettenham, J., Nightingale, N., Morgan, K., et al. (1996). Psychological markers in the detection of autism in infancy in a large population. The British Journal of Psychiatry, 168(2), 158–163.

Council on Children With Disabilities, Section on Developmental Behavioral Pediatrics, Bright Futures Steering Committee and Medical Home Initiatives for Children With Special Needs Project Advisory Committee. (2006). Identifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. Pediatrics, 118(4), 405–420. https://doi.org/10.1542/peds.2006-1231.

Daniels, A. M., Halladay, A. K., Shih, A., Elder, L. M., & Dawson, G. (2014). Approaches to enhancing the early detection of autism spectrum disorders: A systematic review of the literature. Journal of the American Academy of Child and Adolescent Psychiatry, 53(2), 141–152. https://doi.org/10.1016/j.jaac.2013.11.002.

Dendukuri, N., Schiller, I., Joseph, L., & Pai, M. (2012). Bayesian meta-analysis of the accuracy of a test for tuberculous pleuritis in the absence of a gold standard reference. Biometrics, 68(4), 1285–1293. https://doi.org/10.1111/j.1541-0420.2012.01773.x.

Harbord, R. M., & Whiting, P. (2009). metandi: Meta-analysis of diagnostic accuracy using hierarchical logistic regression. Stata Journal, 9(2), 211–229.

McPheeters, M. L., Weitlauf, A., Vehorn, A., Taylor, C., Sathe, N. A., Krishnaswami, S., et al. (2016). Screening for autism spectrum disorder in young children: A systematic evidence review for the U.S. Preventive Services Task Force. Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK349703/.

Reitsma, J. B., Glas, A. S., Rutjes, A. W. S., Scholten, R. J. P. M., Bossuyt, P. M., & Zwinderman, A. H. (2005). Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. Journal of Clinical Epidemiology, 58(10), 982–990. https://doi.org/10.1016/j.jclinepi.2005.02.022.

Robins, D. L. (2008). Screening for autism spectrum disorders in primary care settings. Autism, 12(5), 537–556. https://doi.org/10.1177/1362361308094502.

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. https://doi.org/10.1542/peds.2013-1813.

Rutter, C. M., & Gatsonis, C. A. (2001). A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Statistics in Medicine, 20(19), 2865–2884. https://doi.org/10.1002/sim.942.