



Published in final edited form as:

*Autism Res.* 2022 November ; 15(11): 2181–2191. doi:10.1002/aur.2801.

## Replication study for ADOS-2 cut-offs to assist evaluation of autism spectrum disorder

Ji Su Hong, MD<sup>1,3</sup>, Vini Singh, MPH<sup>1</sup>, Luke Kalb, PhD<sup>1,4</sup>, Rachel Reetzke, PhD<sup>1,3</sup>, Natasha N. Ludwig, PhD<sup>2,3</sup>, Danika Pfeiffer, PhD<sup>1,3</sup>, Calliope Hologue, PhD<sup>1,4</sup>, Deepa Menon, MBBS<sup>1,5</sup>, Qing Lu<sup>1,6</sup>, Ahlam Ashkar, BS<sup>1</sup>, Rebecca Landa, PhD<sup>1,3</sup>

<sup>1</sup>Center for Autism and Related Disorders, Kennedy Krieger Institute

<sup>2</sup>Department of Neuropsychology, Kennedy Krieger Institute

<sup>3</sup>Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine

<sup>4</sup>Department of Mental Health, Johns Hopkins University Bloomberg School of Public Health

<sup>5</sup>Department of Neurology, Johns Hopkins University School of Medicine

<sup>6</sup>Johns Hopkins University School of Arts and Sciences

### Abstract

The Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) has been widely used for ASD assessment. While prior studies investigated sensitivity and specificity of ADOS-2 Modules 1–3, there has been limited research addressing algorithm cut-off scores to optimize ADOS-2 classification. The goal of this study was to assess algorithm cut-off scores for diagnosing ASD with Modules 1–3, and to evaluate alignment of the ADOS-2 classification with best estimate clinical diagnosis. Participants included 3144 children aged 31 months or older who received ADOS-2 Modules 1–3, as well as a best estimate clinical diagnosis. Five classification statistics were reported for each module: sensitivity, specificity, positive predictive value, negative predictive value, and accuracy (i.e., Receiver Operator Classification Statistic), and these statistics were calculated for the optimal cut-off score. Frequency tables were used to compare ADOS-2 classification and best estimate clinical diagnosis. Half of the sample received Module 3, 21% received Module 2, and 29% received Module 1. The overall prevalence of ASD was 60%; the male to female ratio was 4:1, and half of the sample was non-White. Across all modules, the autism spectrum cut-off score from the ADOS-2 manual resulted in high sensitivity (95%+) and low specificity (63%–73%). The autism cut-off score resulted in better specificity (76–86%) with favorable sensitivity (81–94%). The optimal cut-off scores for all modules based on the current

**Corresponding Author:** Ji Su Hong MD, Center for Autism and Related Disorders, Kennedy Krieger Institute, 3901 Greenspring Ave, Baltimore, MD 21211, 443-923-7646, hongj@kennedykrieger.org; jhong57@jhmi.edu.

**Contributors Statement:** Dr. Hong conceptualized the study questions and design, interpreted the results, and wrote, revised, and finalized the manuscript. Ms. Singh conducted the statistical analyses, interpreted the results, wrote the results sections, and edited the manuscript. Dr. Kalb, Dr. Reetzke, Dr. Ludwig, Dr. Pfeiffer, Dr. Hologue, Dr. Menon, Ms. Lu, and Ms. Ashkar interpreted the results, and edited the manuscript. Dr. Landa assisted with conceptualizing the study questions and design, interpreted the results, and edited the manuscript. All authors approved the final manuscript as submitted and agreed to be accountable for all aspects of the work.

**Financial Disclosures:** The authors have no financial relationships relevant to this article to disclose.

**Conflict of Interest:** The authors have no conflicts of interest relevant to this article to disclose.

sample were within the autism spectrum classification range except Module 2 Algorithm 2. In the No ASD group, 29% had false positives (ADOS-2 autism spectrum classification or autism classification). The ADOS-2 autism spectrum classification did not indicate directionality for diagnostic outcome (ASD 56% vs. No ASD 44%). While cut-off scores of ADOS-2 Modules 1–3 in the manual yielded good clinical utility in ASD assessment, false positives and low predictability of the autism spectrum classification remain challenging for clinicians.

### Lay Abstract:

ADOS-2 Modules 1–3 have been widely used for ASD assessment, but there has been limited research on algorithm cut-off scores to optimize ADOS-2 clinical performance. Using a large independent sample, we examined alignment of the ADOS-2 classification with clinicians' best estimate clinical diagnosis, assessing algorithm cut-off scores. Cut-off scores of ADOS-2 Modules 1–3 in the manual yielded good clinical utility in ASD classification. The optimal cut-off scores based on the current sample were generally within the autism spectrum classification range.

### Keywords

autism spectrum disorder; diagnosis; classification; cut-off score

---

### Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by social impairment and repetitive behaviors. Since its first identification in medicine (Kanner, 1943), there have been remarkable advances in characterizing its phenotype and diagnostic criteria, as reflected in revisions of the Diagnostic and Statistical Manual of Mental Disorders (Harris, 2018). As the phenotype of ASD has been further clarified and diagnostic criteria have increasingly expanded, prevalence estimates have rapidly increased. Currently, ASD is estimated to occur 1 in 44 children (Maenner, 2021). As more children are identified with ASD, there has been an increasing need for assessment tools which yield diagnostic outcomes in a reliable and valid manner. Over the past two decades, many studies have investigated the use of ASD assessments tools, such as the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) (Lord et al., 2012), which is considered to be a gold standard ASD assessment tool (Falkmer et al., 2013; Harstad et al., 2015).

The ADOS-2 is a standardized, semi-structured observational measure of ASD symptoms, providing probes to evaluate communication, social interaction, play, restricted and repetitive behaviors, and other abnormal behaviors (hyperactivity, disruptive behaviors, anxiety). The ADOS-2 has 7 modules for different ages and expressive language levels (Hus Bal et al., 2020; Lord et al., 2012). Five modules are commercially available and the Adapted-ADOS is available for researchers. The initial psychometrics studies conducted by the developers indicated favorable validity statistics, with the exception of the Few to No words Algorithm of ADOS-2 Module 1, which led to the development of the ADOS-2 Toddler Module (Table 1) (Dorlack et al., 2018; Gotham et al., 2007, 2008; Hus Bal & Lord, 2014; Lebersfeld et al., 2021; Luyster et al., 2009).

Since its first psychometrics studies, there have been multiple replication studies to assess the clinical validity of ADOS-2 with research samples and clinical samples, in the US as well as other countries (see Table 1) (Bildt et al., 2009, 2016; Camodeca, 2018; Chojnicka & Pisula, 2017; Colombi et al., 2020; Dorlack et al., 2018; Gray et al., 2008; Hong et al., 2021; Hus Bal & Lord, 2015; Kamp-Becker et al., 2013; Langmann et al., 2017; Lee et al., 2019; Maddox et al., 2017; Medda et al., 2019; Molloy et al., 2011; Oosterling et al., 2010; Pugliese et al., 2015; Zander et al., 2015). These replication studies showed a wide range of variability in sensitivity and specificity based on study design, sample size, and characteristics of the non-ASD comparison group. Findings in some studies were greatly impacted by small sample size. While clinical validity has been studied for ADOS-2 Modules 1–3, there has been only one study that addressed algorithm cut-off scores to optimize ADOS-2 classification (Chojnicka & Pisula, 2017). Chojnicka and Pisula developed the ADOS-2 Polish version and determined cut-off scores of each module based on their Polish sample rather than following the developer’s initial studies and the ADOS-2 manual. Since the ADOS-2 is regarded as a gold standard in the autism field and widely used in clinical evaluations, there is a need for a replication study with a large independent sample to examine its clinical performance and determine the best algorithm cut-off scores to support the ASD diagnostic decision-making process.

To address this gap in the literature, the current study had two aims: (1) to assess Modules 1–3 algorithm cut-off scores for diagnosing ASD, and (2) to evaluate alignment of the ADOS-2 classification with best estimate clinical diagnosis.

## METHOD

### Participants

Data for this cross-sectional study were obtained from 3118 children who received 3144 ASD diagnostic evaluations at a Mid-Atlantic urban tertiary ASD-specialty clinic between October 2012 and June 2019. All children were referred for diagnostic clarification, ASD or not. The inclusion criteria were: 1) age 31 months or older, 2) completion of ADOS-2 Modules 1, 2, or 3 at the clinic, 3) completion of a comprehensive diagnostic evaluation by a physician, clinical psychologist, or neuropsychologist within 60 days of the ADOS-2 administration, 4) determination of best estimate clinical (BEC) diagnosis (ASD vs. No ASD), and 4) parental consent to join the local institution’s Institutional Review Board-approved clinical research registry, allowing their child’s de-identified information of medical record to be used for research purposes. The registry’s consent rate for this study was 59%; details on the registry were reported elsewhere (Kalb et al., 2019).

### Measures

**Sociodemographics**—Demographic information about the child was obtained from the medical record. This includes child age (at ADOS-2 administration), sex (male vs. female), race (classified as White, Black, Asian, Multi-Racial, Other/Unknown), ethnicity (Hispanic vs. Non-Hispanic), Insurance-Type (medical assistance vs. private insurance).

**ADOS-2 Modules 1–3**—For children aged 31 months and older, the ADOS-2 module selection is based on a child’s expressive language level and then age. Module 1 is for children aged 31 months and older, who are pre-verbal or have single words, but do not use phrases yet. Module 1 has two Algorithms; Algorithm 1 for children with less than 5 words (autism spectrum cut-off 11, autism cut-off 16), and Algorithm 2 for children with 5 words or more (autism spectrum cut-off 8, autism cut-off 12). Module 2 is for children who use phrases well but are not yet verbally fluent. Module 2 has two Algorithms; Algorithm 1 for children younger than 5 years (autism spectrum cut-off 7, autism cut-off 10), and Algorithm 2 for children aged 5 years or older (autism spectrum cut-off 8, autism cut-off 9). Module 3 is for children and young adolescents (under 16 years old) with fluent speech (expressive language level of 4 years old and above). Module 3 has a single Algorithm (autism spectrum cut-off 7, autism cut-off 9). ADOS-2 defines phrase speech as flexible and meaningful use of 3-word combinations which is non-echoed and sometimes involves a verb. Fluent speech is defined as having a range of flexible sentence types and grammatical forms that provide information about events out of immediate context, including some logical connections. ADOS-2 provides a classification system based on algorithm cut-off scores; non-spectrum classification, autism spectrum classification, and autism classification. The autism classification indicates that a child’s symptoms are within the range of a high proportion of patients with ASD who have similar expressive language level. The autism spectrum classification indicates significant symptoms similar to patients with ASD, but with less severity (Lord et al., 2012).

## Procedure

The majority of cases were evaluated by a paired team, including an ADOS-2 administrator (a speech-language pathologist or a clinical psychologist) and a diagnostic clinician (a physician, a clinical psychologist, or a neuropsychologist). In cases where ASD was considered extremely likely at triage, the child was evaluated by a clinical psychologist only for ADOS-2 administration and diagnostic determination. ADOS-2 was administered and coded by a speech-language pathologist (59%), or a clinical psychologist (41%), before the diagnostic decision was made. Clinicians administering ADOS-2 have maintained clinical reliability by training with a certified ADOS-2 trainer and having access to quarterly booster training and ad hoc consultations with a doctoral level ADOS-2 research reliable psychologist. Though some have established research reliability, the majority of the clinicians did not have ADOS-2 research reliability training. In our clinic ADOS-2 was used as a routine and standard clinical measure to inform diagnostic decision-making and ADOS-2 Modules 1–3 were conducted regardless of non-verbal mental age. After the ADOS-2 was administered and coded, diagnostic evaluation for ASD was conducted by physicians, clinical psychologists, or neuropsychologists who specialize in evaluating and treating neurodevelopmental disorders, with a subspecialty in autism spectrum disorder. When an ADOS-2 administrator and a diagnosing clinician were different, ADOS-2 results (algorithm score and classification) and the administrator’s observations were discussed with the diagnosing clinician. The diagnosing physician, clinical psychologist, or neuropsychologist used the ADOS-2 findings along with parent reported symptoms, medical and developmental history, as well as other standardized assessments (i.e., questionnaires completed by parents and/or teachers) to develop a BEC diagnosis (ASD vs. No ASD),

based on the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (American Psychiatric Association, 2013).

## Analysis

Statistical analyses were conducted separately for each module. Five classification statistics were reported: 1) sensitivity, 2) specificity, 3) positive predictive value (PPV), 4) negative predictive value (NPV), and 5) accuracy (i.e., Receiver Operator Classification Statistic; ROC). ROC values above 0.8 are considered optimal. These statistics were calculated for each ADOS-2 algorithm cut-off score. Next, frequency tables were used to compare ADOS-2 classification and BEC diagnosis.

## RESULTS

### Sample characteristics

Table 2 shows sample characteristics of 3118 children who received a total of 3144 ASD diagnostic evaluations (1865 with ASD, 1279 without ASD). There were 26 children who had two ASD diagnostic evaluations either more than 6 months apart or using different ADOS-2 modules. Therefore, each diagnostic evaluation was counted separately. Half of the sample received Module 3, 21% received Module 2, and 29% received Module 1. While the overall prevalence of ASD in our sample was 60%, it was not equal across modules. The highest rate of ASD diagnosis was with Module 1, while the lowest was with Module 3. The male to female ratio was 4:1 and half of the sample was non-White, with some variation across modules. There was no significant difference in age between Module 1 Algorithm 1, Module 1 Algorithm 2, and Module 2 Algorithm 1 (3.99, 3.89, 4.19, respectively), however those with Module 2 Algorithm 2 (6.63) and Module 3 (9.11) were significantly older. Information about parental education, marital status, and insurance type was available for 2144 (68.2%), 2161 (68.7%), 3133 (99.7%), respectively.

### Accuracy and Classification statistics

Table 3 shows the classification statistics most proximal to algorithm cut-off scores of the ADOS-2 manual (See supplemental tables 1–5 for the full range of scores). Across all modules, the autism spectrum cut-off from the ADOS-2 manual resulted in high levels of sensitivity (95%+), however specificity was low (63%–73%). The autism cut-off led to substantial improvement in specificity (76–86%) with modest decline in sensitivity (81–94%). Generally, all metrics for the autism cutoff were strong (all >70%, except NPV for Module 1 Algorithm 1). ROC values for all modules were very high as well (0.89–0.95).

The optimal cut-offs for all modules from the current sample's data were within the autism spectrum classification range except Module 2 Algorithm 2, for which it was in the autism classification range. Using these optimal cut-offs led to favorable accuracy (83%–91%) and excellent sensitivity (91%–95%), however specificity (73%–82%) was relatively lower.

Figure 1 represents the frequency of algorithm scores by BEC diagnosis. While ADOS-2 algorithm scores were well separated between ASD and No ASD groups for each module, there was an overlap area of scores which did not indicate a direction of diagnostic status.

Table 4 summarizes the comparison of existing ADOS-2 classification vs. BEC diagnosis. In the ASD group, the proportion of participants having ADOS-2 autism spectrum classification or autism classification (true positives) was high across all modules (97%, ranging 95%–99%). In the No ASD group, 377 of 1279 children (29%, ranging 27%–38%) had false positives with ADOS-2 autism spectrum classification or autism classification. When considering children who received ADOS-2 autism spectrum classification (358 of 3144, 11%, ranging 5%–15%), they were almost equally likely to have ASD (56%, ranging 21%–79%) versus not (44%, ranging 21%–79%) and it indicated that ADOS-2 autism spectrum classification did not have a directionality for the clinical diagnosis of ASD or not. More than half of our sample (58%) fell into ADOS-2 autism classification, and the true rate of ASD in this group was high (88%, ranging 82%–98%).

## DISCUSSION

The current study represents the largest independent evaluation of the ADOS-2 algorithm cut-off scores to date. In our sample, the established autism spectrum cut-off performed very well in classifying the ASD group as ADOS-2 positive, given the very high sensitivity. However, the specificity was moderate, resulting in elevated false positives. The autism cut-off improved specificity and false positive while lowering sensitivity. When using both the autism spectrum cut-off and the autism cut-off, with clinical interpretation, ADOS-2 can yield a clinically useful guide in detecting ASD with favorable sensitivity and specificity (Table 3).

Compared to the developers' studies (Gotham et al., 2007, 2008), the current study indicated higher sensitivity and lower specificity across all modules. While replication studies showed a wide range of variability in sensitivity (61–100%) and specificity (20–96%), the current study showed consistently high sensitivity (95 %+) and moderately favorable specificity (62–73%) in all modules.

A second goal of this study was to examine optimal algorithm cut-off scores for all modules in this clinical sample. The optimal cut-offs were located in the autism spectrum classification range, except Module 2 Algorithm 2. Given very narrow classification ranges in Module 2 Algorithm 2 (autism spectrum cut-off: 8, autism cut-off: 9), the chosen cut-off of 10 was very close to the cut-offs in the ADOS-2 manual (Table 3).

As shown in Table 1, since the initial ADOS-2 studies conducted by developers in 2007, there have been 8 replication studies with Module 1 Algorithms 1 and 2 (Chojnicka & Pisula, 2017; Gotham et al., 2007, 2008; Gray et al., 2008; Medda et al., 2019; Molloy et al., 2011; Oosterling et al., 2010; Zander et al., 2015). Among those, six studies included toddlers aged 30 months and below, into Module 1. Though Chojnicka et al. and Medda et al. correctly excluded those toddlers from Module 1, their sample size was small (37 and 81) and it is difficult to generalize the findings. The current study correctly excluded toddlers appropriate for the Toddler Module, designed for ages 30 months and below, and showed that both Module 1 Algorithms 1 and 2 yielded good sensitivity and specificity with the autism spectrum cut-off and the autism cut-off, comparable to other modules. This finding

provides important confidence and reassurance for using existing Module 1 cut-offs in the ADOS-2 manual for clinical evaluation.

While ADOS-2 Modules 1–3 provided good validity statistics, the false positives remains a significant challenge for clinicians. Nearly one third of the No ASD group were misclassified by ADOS-2 manual cut-offs. In the group with ADOS-2 autism spectrum classification, 44% did not have ASD, and in the group with ADOS-2 autism classification, 12% did not have ASD. This indicates concerns about positive predictability of existing ADOS-2 manual cut-offs, specifically for ADOS-2 autism spectrum classification. This finding is consistent with prior studies that showed clinically referred samples with a wide range of developmental and mental disorders presenting as false positive with ADOS/ ADOS-2 algorithm cut-off scores (Colombi et al., 2020; Conner et al., 2019; Maddox et al., 2017; Medda et al., 2019; Molloy et al., 2011; Sedgewick et al., 2019; Sikora et al., 2008; Stadnick et al., 2015), and Havdahl et al. (Havdahl et al., 2016) indicated that mental health symptoms-adjusted cut-offs could improve specificity. While this is not entirely surprising, given the clinical complexity of referred samples, clinicians in ASD-specialty clinic settings should interpret ADOS-2 findings with other clinical information during diagnostic decision-making. As recommended by the ADOS-2 developers (Lord et al., 2012), clinicians should use a combination of all clinical data to determine diagnosis, including parent reported symptoms, medical and developmental history, standardized assessments of cognitive and language function, along with direct child observation. Our findings further reinforce language from the ADOS-2 manual that states “The ADOS-2 classification should never be used in isolation to determine an individual’s clinical diagnosis or eligibility for services.” (Lord et al., 2012, p.187).

The current findings should be interpreted in terms of the study’s strengths and limitations. For strengths, the ASD diagnosis was provided by clinicians with ASD expertise, the sample size was very large, and the sample was clinically as well as socio-demographically heterogeneous. For limitations, as per standards of best clinical practice, diagnoses were not defined independently of ADOS-2 results and this could have resulted in diagnostic biases. This study also lacked access to important developmental data (e.g., intelligence, language test results, co-occurring diagnoses) that would have permitted better sample characterization. The lack of non-verbal mental age in Module 1 Algorithm 1 may have affected the frequency of false positives, given that ADOS-2 manual indicated that inclusion of children with non-verbal mental age  $\leq$  15 months increased false positive in Module 1 Algorithm 1 (Lord et al., 2012, p.226–228 and 243). The majority of ADOS-2 administrators were not research reliable, although they were monitored for ADOS-2 scoring calibration with research-reliable clinicians. Though the sample size is very large, the subsamples of No ASD group in Module 1 Algorithm 1, Module 1 Algorithm 2, and Module 2 Algorithm 2 were relatively small and it may have affected false positives. Lastly, the current sample came from a single ASD-specialty center and the presence of selection and referral bias is possible. Despite these limitations, this study provides important evidence for clinical validity of the ADOS-2 as well as meaningful insights for interpreting ADOS-2 results.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGEMENT

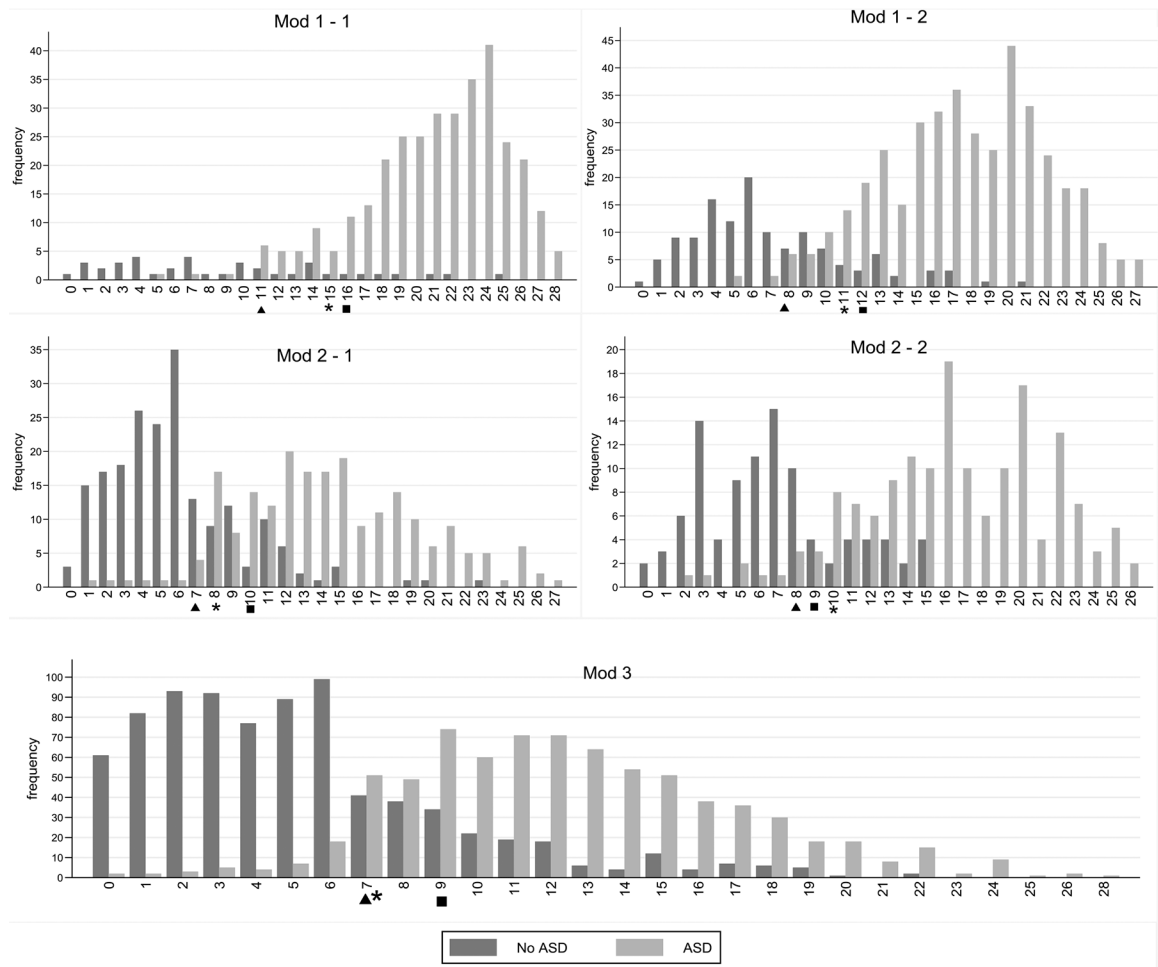
The authors have no financial relationships or conflict of interest relevant to this article to disclose. This work was supported by P50 HD103538.

## REFERENCE

- de Bildt A, Sytema S, van Lang ND, Minderaa RB, van Engeland H, & de Jonge MV. (2009). Evaluation of the ADOS revised algorithm: The applicability in 558 Dutch children and adolescents. *Journal of Autism and Developmental Disorders*, 39(9), 1350–1358. [PubMed: 19452268]
- de Bildt A, Sytema S, Meffert H, & Bastiaansen JA (2016). The Autism Diagnostic Observation Schedule, Module 4: Application of the revised algorithms in an independent, well-defined, Dutch sample (n= 93). *Journal of Autism and Developmental Disorders*, 46(1), 21–30. [PubMed: 26319249]
- Camodeca A (2018). Utility of three N-Item scales of the child behavior checklist 6–18 in autism diagnosis. *Research in Autism Spectrum Disorders*, 51, 75–85. 10.1016/j.rasd.2018.04.004
- Chojnicka I, & Pisula E (2017). Adaptation and Validation of the ADOS-2, Polish Version. *Frontiers in Psychology*, 8. 10.3389/fpsyg.2017.01916
- Colombi C, Fish A, & Ghaziuddin M (2020). Utility of the ADOS-2 in children with psychiatric disorders. *European Child & Adolescent Psychiatry*, 29(7), 989–992. 10.1007/s00787-019-01411-8 [PubMed: 31587085]
- Conner CM, Cramer RD, & McGonigle JJ (2019). Examining the Diagnostic Validity of Autism Measures Among Adults in an Outpatient Clinic Sample. *Autism in Adulthood*, 1(1), 6–68. 10.1089/aut.2018.0023
- Dorlack T, Myers O, & Kodituwakku P (2018). A Comparative Analysis of the ADOS-G and ADOS-2 Algorithms: Preliminary Findings. *Journal of Autism and Developmental Disorders*, 48(6), 2078–2089. 10.1007/s10803-018-3475-3 [PubMed: 29380271]
- Falkmer T, Anderson K, Falkmer M, & Horlin C (2013). Diagnostic procedures in autism spectrum disorders: A systematic literature review. *European Child & Adolescent Psychiatry*, 22(6), 329–340. 10.1007/s00787-013-0375-0 [PubMed: 23322184]
- Gotham K, Risi S, Dawson G, Tager-Flusberg H, Joseph R, Carter A, Hepburn S, McMahon W, Rodier P, Hyman SL, Sigman M, Rogers S, Landa R, Spence MA, Osann K, Flodman P, Volkmar F, Hollander E, Buxbaum J, ... Lord C (2008). A Replication of the Autism Diagnostic Observation Schedule (ADOS) Revised Algorithms. *Journal of the American Academy of Child and Adolescent Psychiatry*, 47(6), 642–651. 10.1097/CHI.0b013e31816bffb7 [PubMed: 18434924]
- Gotham K, Risi S, Pickles A, & Lord C (2007). The Autism Diagnostic Observation Schedule: Revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders*, 37(4), 613. [PubMed: 17180459]
- Gray KM, Tonge BJ, & Sweeney DJ (2008). Using the Autism Diagnostic Interview-Revised and the Autism Diagnostic Observation Schedule with young children with developmental delay: Evaluating diagnostic validity. *Journal of Autism and Developmental Disorders*, 38(4), 657–667. 10.1007/s10803-007-0432-y [PubMed: 17690967]
- Harris J (2018). Leo Kanner and autism: A 75-year perspective. *International Review of Psychiatry (Abingdon, England)*, 30(1), 3–17. 10.1080/09540261.2018.1455646 [PubMed: 29667863]
- Harstad EB, Fogler J, Sideridis G, Weas S, Mauras C, & Barbaresi WJ (2015). Comparing diagnostic outcomes of autism spectrum disorder using DSM-IV-TR and DSM-5 criteria. *Journal of Autism and Developmental Disorders*, 45(5), 1437–1450. [PubMed: 25384720]
- Havdahl KA, Hus Bal V, Huerta M, Pickles A, Øyen A-S, Stoltenberg C, Lord C, & Bishop SL (2016). Multidimensional Influences on Autism Symptom Measures: Implications for Use in Etiological

- Research. *Journal of the American Academy of Child & Adolescent Psychiatry*, 55(12), 1054–1063.e3. 10.1016/j.jaac.2016.09.490 [PubMed: 27871640]
- Hong JS, Singh V, Kalb L, Ashkar A, & Landa R (2021). Replication study of ADOS-2 Toddler Module cut-off scores for autism spectrum disorder classification. *Autism Research*, e02496.
- Hus Bal V, & Lord C (2014). The Autism Diagnostic Observation Schedule, Module 4: Revised Algorithm and Standardized Severity Scores. *Journal of Autism and Developmental Disorders*, 44(8), 1996–2012. 10.1007/s10803-014-2080-3 [PubMed: 24590409]
- Hus Bal V, & Lord C (2015). Replication of Standardized ADOS Domain Scores in the Simons Simplex Collection. *Autism Research*, 8(5), 583–592. 10.1002/aur.1474 [PubMed: 25712123]
- Hus Bal V, Maye M, Salzman E, Huerta M, Pepa L, Risi S, & Lord C (2020). The adapted ADOS: A new module set for the assessment of minimally verbal adolescents and adults. *Journal of Autism and Developmental Disorders*, 50(3), 719–729. [PubMed: 31736004]
- Kalb L, Jacobson L, Zisman C, Mahone E, Landa R, Azad G, Menon D, Singh V, Zabel A, & Pritchard A (2019). Interest in Research Participation Among Caregivers of Children with Neurodevelopmental Disorders. *Journal of Autism and Developmental Disorders*, 49(9), 3786–3797. [PubMed: 31172337]
- Kamp-Becker I, Ghahreman M, Heinzl-Gutenbrunner M, Peters M, Remschmidt H, & Becker K (2013). Evaluation of the revised algorithm of Autism Diagnostic Observation Schedule (ADOS) in the diagnostic investigation of high-functioning children and adolescents with autism spectrum disorders. *Autism*, 17(1), 87–102. [PubMed: 21610187]
- Kanner L (1943). Autistic disturbances of affective contact. *Nervous Child*, 2(3), 217–250.
- Langmann A, Becker J, Poustka L, Becker K, & Kamp-Becker I (2017). Diagnostic utility of the autism diagnostic observation schedule in a clinical sample of adolescents and adults. *Research in Autism Spectrum Disorders*, 34, 34–43.
- Lebersfeld JB, Swanson M, Clesi CD, & O’Kelley SE (2021). Systematic Review and Meta-Analysis of the Clinical Utility of the ADOS-2 and the ADI-R in Diagnosing Autism Spectrum Disorders in Children. *Journal of Autism and Developmental Disorders*. 10.1007/s10803-020-04839-z
- Lee KS, Chung SJ, Thomas HR, Park J, & Kim SH (2019). Exploring diagnostic validity of the autism diagnostic observation schedule-2 in South Korean toddlers and preschoolers. *Autism Research*, 12(9), 1356–1366. 10.1002/aur.2125 [PubMed: 31112376]
- Lord C, Rutter M, DiLavore P, Risi S, Gotham K, & Bishop S (2012). *Autism diagnostic observation schedule—2nd edition (ADOS-2)*. Los Angeles, CA. Western Psychological Corporation.
- Luyster R, Gotham K, Guthrie W, Coffing M, Petrak R, Pierce K, Bishop S, Esler A, Hus V, Oti R, Richler J, Risi S, & Lord C (2009). The Autism Diagnostic Observation Schedule—Toddler Module: A New Module of a Standardized Diagnostic Measure for Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 39(9), 1305–1320. 10.1007/s10803-009-0746-z [PubMed: 19415479]
- Maddox B, Brodtkin E, Calkins M, Shea K, Mullan K, Hostager J, Mandell D, & Miller J (2017). The Accuracy of the ADOS-2 in Identifying Autism among Adults with Complex Psychiatric Conditions. *Journal of Autism and Developmental Disorders*, 47(9), 2703–2709. 10.1007/s10803-017-3188-z [PubMed: 28589494]
- Maenner MJ (2021). Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years—Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2018. *MMWR. Surveillance Summaries*, 70. 10.15585/mmwr.ss7011a1
- Medda JE, Cholemkery H, & Freitag CM (2019). Sensitivity and specificity of the ADOS-2 algorithm in a large german sample. *Journal of Autism and Developmental Disorders*, 49(2), 750–761. [PubMed: 30238180]
- Molloy CA, Murray DS, Akers R, Mitchell T, & Manning-Courtney P (2011). Use of the Autism Diagnostic Observation Schedule (ADOS) in a clinical setting. *Autism*, 15(2), 143–162. 10.1177/1362361310379241 [PubMed: 21339248]
- Oosterling I, Roos S, de Bildt A, Rommelse N, de Jonge M, Visser J, Lappenschaar M, Swinkels S, van der Gaag RJ, & Buitelaar J (2010). Improved Diagnostic Validity of the ADOS Revised Algorithms: A Replication Study in an Independent Sample. *Journal of Autism and Developmental Disorders*, 40(6), 689–703. 10.1007/s10803-009-0915-0 [PubMed: 20148299]

- Pugliese C, Kenworthy L, Bal V, Wallace G, Yerys B, Maddox B, White S, Popal H, Armour A, Miller J, Herrington J, Schultz R, Martin A, & Anthony L (2015). Replication and Comparison of the Newly Proposed ADOS-2, Module 4 Algorithm in ASD Without ID: A Multi-site Study. *Journal of Autism and Developmental Disorders*, 45(12), 3919–3931. 10.1007/s10803-015-2586-3 [PubMed: 26385796]
- Sedgewick F, Kerr-Gaffney J, Leppanen J, & Tchanturia K (2019). Anorexia Nervosa, Autism, and the ADOS: How Appropriate Is the New Algorithm in Identifying Cases? *Frontiers in Psychiatry*, 10, 507. 10.3389/fpsyt.2019.00507 [PubMed: 31379625]
- Sikora DM, Hartley SL, McCoy R, Gerrard-Morris AE, & Dill K (2008). The performance of children with mental health disorders on the ADOS-G: A question of diagnostic utility. *Research in Autism Spectrum Disorders*, 2(1), 188–197. 10.1016/j.rasd.2007.05.003
- Stadnick N, Brookman-Frazee L, Williams KN, Cerda G, & Akshoomoff N (2015). A pilot study examining the use of the autism diagnostic observation schedule in community-based mental health clinics. *Research in Autism Spectrum Disorders*, 20, 39–46. [PubMed: 26379765]
- Zander E, Sturm H, & Bölte S (2015). The added value of the combined use of the Autism Diagnostic Interview–Revised and the Autism Diagnostic Observation Schedule: Diagnostic validity in a clinical Swedish sample of toddlers and young preschoolers. *Autism*, 19(2), 187–199. 10.1177/1362361313516199 [PubMed: 24413849]



**Figure 1.**  
Frequency of algorithm scores between ASD group vs. No ASD group  
Horizontal axis: ADOS-2 algorithm score  
▲ autism spectrum cut-off in ADOS-2 manual ■ autism cut-off in ADOS-2 manual \*  
optimal cut-off in the current sample

**Table 1.**

ADOS-2 psychometrics studies

Study (country)	Toddler Algorithm 1 Sens/Spec (ASD/NS)	Toddler Algorithm 2 Sens/Spec (ASD/NS)	Module 1 Algorithm 1 Sens/Spec (ASD/NS)	Module 1 Algorithm 2 Sens/Spec (ASD/NS)	Module 2 Algorithm 1 Sens/Spec (ASD/NS)	Module 2 Algorithm 2 Sens/Spec (ASD/NS)	Module 3 Sens/Spec (ASD/NS)	Module 4 Sens/Spec (ASD/NS)
Gotham *2007 (US)			0.92/0.49 (446/49)	0.87/0.87 (276/76)	0.91/0.87 (107/30)	0.91/0.87 (162/30)	0.82/0.80 (315/83)	
Gotham *2009 (US)			0.86/0.80 (203/46)	0.92/0.83 (175/46)	0.80/1.00 (69/18)		0.71/0.90 (384/73)	
Luyster *2009 (US)	0.87/0.86 (87/64)	0.81/0.83 (59/24)						
Hus *2014 (US)								0.91/0.82 (347/90)
Gray 2008 (Australia)			0.92/0.86 (50/7)	0.78/0.92 (89/49)				
Bildt 2009 (Dutch)				0.89/0.67 (75/24)		0.70/0.70 (74/50)	0.77/0.69 (264/70)	
Oosterling 2010 (Dutch)			0.93/0.70 (69/10)	0.72/0.86 (98/80)	0.61/0.83 (75/30)	0.64/0.85 (58/40)		
Kamp-Becker 2011 (Germany)							0.95/0.69 (126/126)	
Molloy 2011 (US)			0.96/0.29 (70/17)	0.98/0.45 (58/32)	0.76/0.60 (54/53)	0.88/0.60 (51/40)	0.87/0.34 (96/113)	
Zander 2015 (Sweden)			0.99/0.52 (72/21)	0.94/0.76 (64/29)	1.00/0.65 (32/40)			
Pugliese 2015 (US)								0.85/0.72 (253/68)
Langmann 2017 (Germany)								0.86/0.80 (165/91)
Maddox 2017 (US)								1.00/0.70 (6/69)
Chojnicka 2017 (Poland)	0.96/0.91 (34/42)	0.71/0.88 (41/40)	0.92/0.82 (41/40)	0.93/0.83 (41/40)	0.91/0.96 (40/41)	0.84/0.75 (40/41)	0.90/0.81 (39/43)	0.92/0.74 (39/42)
Medda 2018 (Germany)			1.00/0.20 (32/5)	0.99/0.18 (102/17)	1.00/1.00 (12/4)	0.91/0.66 (106/29)	0.91/0.66 (290/82)	0.93/0.29 (129/18)
Camodeca 2018 (US)							0.92/0.96 (78/160)	0.94/0.94 (36/32)
Lee 2019 (South Korea)	0.94/0.82 (18/11)	1.00/0.86 (6/7)		1.00/0.88 (19/16)	1.00/1.00 (7/5)			
Colombi 2019 (US)							0.58/0.57 (n=35)	0.56/0.60 (n=23)
Hong 2021 (US)	0.98/0.75 (236/61)	0.96/0.85 (69/46)						

Toddler Algorithm 1: 12–20month, and 21–30 months with < 5 words.

Toddler Algorithm 2: 21–30 months with 5 words.

Module 1 Algorithm 1: 31 months and older with < 5 words.

Module 1 Algorithm 2: 31 months and older with 5 words.

Module 2 Algorithm 1: < 5years.

Module 2 Algorithm 2: 5 years

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

ASD: autism spectrum disorder. NS: No autism spectrum disorder

\* indicates studies conducted by ADOS-2 developers.

Gotham et al. (2007) included children aged 14–30 months. Gotham et al. (2009) included children aged 18–30 months, and children whose nonverbal mental age was 15 months or below were excluded. Gray et al. (2008) included children aged 20–30 months. Oosterling et al. (2010) included children aged 18–30 months. Molly et al. (2011) included children aged 26–30 months. Zander et al. (2015) included children aged 20–30 months.

Chojnicka et al. (2017) developed their own cut-off scores and did not use developers' cut-off scores.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2.**

Sample characteristics

	Module 1 Algorithm 1 <sup>a</sup> N=364		Module 1 Algorithm 2 <sup>b</sup> N=534		Module 2 Algorithm 1 <sup>c</sup> N=413		Module 2 Algorithm 2 <sup>d</sup> N=257		Module 3 <sup>e</sup> N=1576		Total <sup>f</sup> N=3144	
	No ASD	ASD	No ASD	ASD	No ASD	ASD	No ASD	ASD	No ASD	ASD	No ASD	ASD
Age <sup>b,e,f</sup>	N=40	N=324	N=129	N=405	N=200	N=213	N=98	N=159	N=812	N=764	N=1279	N=1865
	3.99 (1.79)	3.98 (1.74)	3.89 (1.49)	3.89 (1.49)	4.19 (0.67)	4.19 (0.67)	6.63 (1.79)	6.63 (1.79)	9.11 (2.71)	9.11 (2.71)	6.78 (3.21)	6.41 (3.25)
	4.07 (2.20)	3.98 (1.74)	3.55 (1.43)	4.00 (1.49)	4.17 (0.69)	4.20 (0.65)	6.47 (1.80)	6.73 (1.78)	8.97 (2.66)	9.25 (2.75)	7.33 (3.21)	6.41 (3.25)
Sex: <sup>d</sup>												
Female	11 (27.5%)	68 (21.0%)	18 (14.0%)	75 (18.5%)	47 (23.5%)	53 (24.9%)	24 (24.5%)	19 (11.9%)	165 (20.3%)	144 (18.8%)	265 (20.7%)	359 (19.2%)
Male	29 (72.5%)	256 (79.0%)	111 (86.0%)	330 (81.5%)	153 (76.5%)	160 (75.1%)	74 (75.5%)	140 (88.1%)	647 (79.7%)	620 (81.2%)	1014 (79.3%)	1506 (80.8%)
Race/Ethnicity: <sup>a,c,f</sup>												
White	12 (30.0%)	122 (37.7%)	63 (48.8%)	164 (40.5%)	111 (55.5%)	119 (55.9%)	37 (37.8%)	57 (35.8%)	512 (63.1%)	461 (60.3%)	735 (57.5%)	923 (49.5%)
Black	10 (25.0%)	122 (37.7%)	26 (20.2%)	107 (26.4%)	44 (22.0%)	39 (18.3%)	37 (37.8%)	55 (34.6%)	153 (18.8%)	160 (20.9%)	270 (21.1%)	483 (25.9%)
Asian	5 (12.5%)	34 (10.5%)	14 (10.9%)	50 (12.3%)	12 (6.00%)	18 (8.45%)	6 (6.12%)	12 (7.55%)	20 (2.46%)	46 (6.02%)	57 (4.46%)	160 (8.58%)
Multi-Racial	6 (15.0%)	10 (3.09%)	13 (10.1%)	25 (6.17%)	16 (8.00%)	14 (6.57%)	6 (6.12%)	10 (6.29%)	83 (10.2%)	42 (5.50%)	124 (9.70%)	101 (5.42%)
Hispanic/Spanish	2 (5.00%)	13 (4.01%)	7 (5.43%)	13 (3.21%)	9 (4.50%)	5 (2.35%)	6 (6.12%)	6 (3.77%)	19 (2.34%)	23 (3.01%)	43 (3.36%)	60 (3.22%)
Other	5 (12.5%)	23 (7.10%)	6 (4.65%)	46 (11.4%)	8 (4.00%)	18 (8.45%)	6 (6.12%)	19 (11.9%)	25 (3.08%)	32 (4.19%)	50 (3.91%)	138 (7.40%)
Education: <sup>b,c,f</sup>												
Less than High School	1 (5.26%)	10 (5.10%)	6 (8.00%)	7 (2.62%)	2 (1.54%)	3 (2.10%)	4 (6.45%)	6 (5.50%)	12 (2.05%)	11 (1.97%)	25 (2.87%)	37 (2.90%)
High School	7 (36.8%)	64 (32.7%)	16 (21.3%)	63 (23.6%)	35 (26.9%)	20 (14.0%)	24 (38.7%)	30 (27.5%)	169 (28.9%)	133 (23.8%)	251 (28.9%)	310 (24.3%)
Trade School/ Associates	8 (42.1%)	44 (22.4%)	26 (34.7%)	49 (18.4%)	20 (15.4%)	16 (11.2%)	9 (14.5%)	18 (16.5%)	134 (22.9%)	113 (20.2%)	197 (22.6%)	240 (18.8%)
Bachelors	2 (10.5%)	38 (19.4%)	13 (17.3%)	77 (28.8%)	42 (32.3%)	51 (35.7%)	16 (25.8%)	26 (23.9%)	141 (24.1%)	153 (27.4%)	214 (24.6%)	345 (27.1%)
Graduate	1 (5.26%)	40 (20.4%)	14 (18.7%)	71 (26.6%)	31 (23.8%)	53 (37.1%)	9 (14.5%)	29 (26.6%)	128 (21.9%)	149 (26.7%)	183 (21.0%)	342 (26.8%)

	Module 1 Algorithm 1 <sup>a</sup> N=364		Module 1 Algorithm 2 <sup>b</sup> N=534		Module 2 Algorithm 1 <sup>c</sup> N=413		Module 2 Algorithm 2 <sup>d</sup> N=257		Module 3 <sup>e</sup> N=1576		Total <sup>f</sup> N=3144	
	No ASD	ASD	No ASD	ASD	No ASD	ASD	No ASD	ASD	No ASD	ASD	No ASD	ASD
Marital: <sup>c,e,f</sup>												
Married/Living Together	13 (65.0%)	132 (66.3%)	51 (68.9%)	193 (71.5%)	126 (87.5%)	126 (87.5%)	40 (63.5%)	70 (63.1%)	334 (57.2%)	358 (63.5%)	531 (60.8%)	879 (68.2%)
Divorced/Separated	2 (10.0%)	14 (7.04%)	6 (8.11%)	19 (7.04%)	7 (4.86%)	7 (4.86%)	7 (11.1%)	10 (9.01%)	94 (16.1%)	97 (17.2%)	123 (14.1%)	147 (11.4%)
Never Married	4 (20.0%)	47 (23.6%)	14 (18.9%)	51 (18.9%)	11 (7.64%)	11 (7.64%)	12 (19.0%)	29 (26.1%)	134 (22.9%)	96 (17.0%)	186 (21.3%)	234 (18.2%)
Widowed	0 (0.00%)	1 (0.50%)	0 (0.00%)	2 (0.74%)	0 (0.00%)	0 (0.00%)	0 (0%)	0 (0%)	5 (0.86%)	1 (0.18%)	6 (0.69%)	4 (0.31%)
Other	1 (5.00%)	5 (2.51%)	3 (4.05%)	5 (1.85%)	2 (1.52%)	2 (1.52%)	4 (6.35%)	2 (1.80%)	17 (2.91%)	12 (2.13%)	27 (3.09%)	24 (1.86%)
Insurance: <sup>c</sup>												
Medical Assistance	25 (62.5%)	155 (48.1%)	56 (43.4%)	147 (36.3%)	47 (22.2%)	47 (22.2%)	41 (42.3%)	56 (35.4%)	299 (37.0%)	276 (36.2%)	502 (39.4%)	681 (36.6%)
Private	15 (37.5%)	167 (51.9%)	73 (56.6%)	258 (63.7%)	165 (77.8%)	165 (77.8%)	56 (57.7%)	102 (64.6%)	509 (63.0%)	486 (63.8%)	772 (60.6%)	1178 (63.4%)
ADOS-2												
CSS <sup>*,f</sup>	3.12 (2.23)	7.44 (1.80)	3.12 (1.89)	7.55 (1.69)	2.98 (1.94)	7.08 (2.00)	3.28 (1.89)	7.24 (1.70)	3.04 (2.30)	7.12 (2.06)	3.06 (2.17)	7.27 (1.91)
SA CSS <sup>*,f</sup>	3.10 (2.13)	6.86 (1.83)	2.99 (1.74)	7.00 (1.96)	3.35 (1.95)	7.09 (2.02)	3.74 (2.00)	7.26 (1.98)	3.62 (2.46)	7.25 (2.07)	3.51 (2.29)	7.11 (1.99)
RRB CSS <sup>*,f</sup>	5.00 (2.85)	8.38 (1.65)	5.57 (2.77)	8.29 (1.77)	4.26 (2.46)	7.04 (2.19)	4.13 (2.40)	7.30 (2.35)	3.15 (2.60)	6.59 (2.59)	3.70 (2.71)	7.38 (2.35)

There were missing data; Education N=2144, Marital N=2161, Insurance N=3133, CSS N=3140, SA CSS N=3142, RRB CSS N=3143

CSS: Calibrated Severity Score, SA: Social Affect, RRB: Restricted repetitive behavior

p<0.05 :

<sup>a</sup>- Module 1 Algorithm 1;

<sup>b</sup>- Module 1 Algorithm 2;

<sup>c</sup>- Module 2 Algorithm 1;

<sup>d</sup>- Module 2 Algorithm 2;

<sup>e</sup>- Module 3;

<sup>f</sup>- All modules

Total  
7

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 3.**

Classification statistics with various algorithm cut-off scores in the current study sample

Module	Algorithm cut-off score	Sensitivity	Specificity	False negative	False positive	Positive predictive value	Negative predictive value	Accuracy	Area under ROC curve
Module 1 Algorithm 1	10	99	55	1	45	95	88	94	0.92
	11▲	99	63	1	38	96	89	95	
	12	97	68	3	33	96	75	94	
	13	96	70	4	30	96	67	93	
	14	94	73	6	28	97	60	92	
	15*	91	80	9	20	97	53	90	
	16■	90	83	10	18	98	50	89	
	17	87	85	14	15	98	44	86	
18	83	88	17	13	98	38	83		
Module 1 Algorithm 2	7	100	56	0	44	88	97	89	0.95
	8▲	99	64	1	36	90	95	90	
	9	98	69	2	31	91	90	91	
	10	96	77	4	23	93	86	91	
	11*	94	82	6	18	94	80	91	
	12■	90	85	10	15	95	73	89	
	13	86	88	15	12	96	66	86	
	14	79	92	21	8	97	59	82	
15	76	94	24	6	97	55	80		
Module 2 Algorithm 1	6	98	52	2	48	68	95	75	0.92
	7▲	97	69	3	31	77	96	84	
	8*	95	76	5	24	81	94	86	
	9	87	80	13	20	82	86	84	
	10■	84	86	16	14	86	83	85	
	11	77	88	23	12	87	78	82	
12	71	93	29	7	91	75	82		
Module 2 Algorithm 2	7	97	50	3	50	76	91	79	0.93
	8▲	96	65	4	35	82	90	84	
	9■	94	76	6	24	86	88	87	
	10*	92	80	8	20	88	86	87	
	11	87	82	13	18	89	79	85	
	12	83	86	17	14	90	75	84	
	13	79	90	21	10	93	72	83	
14	73	94	27	6	95	68	81		
Module 3	6	97	61	3	39	70	96	78	0.89
	7▲*	95	73	5	27	77	94	83	

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Module	Algorithm cut-off score	Sensitivity	Specificity	False negative	False positive	Positive predictive value	Negative predictive value	Accuracy	Area under ROC curve
	8	88	78	12	22	79	87	83	
	9 <sup>▲</sup>	81	83	19	17	82	83	82	
	10	72	87	28	13	84	77	80	

▲ autism spectrum cut-off in ADOS-2 manual

■ autism cut-off in ADOS-2 manual

\* optimal cut-off in the current sample

non-spectrum classification	autism spectrum classification	autism classification
-----------------------------	--------------------------------	-----------------------

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 4.**

Summary using algorithm cut-off scores of ADOS-2 manual

<b>Module 1 Algorithm 1</b>	<b>No ASD (N=40)</b>	<b>ASD (N=324)</b>	<b>All (N=364)</b>	<b>Correctly predicting BEC diagnosis</b>
non-spectrum (0–10)	25 (62.5%)	3 (0.9%)	28 (7.7%)	89%
autism spectrum (11–15)	8 (20%)	30 (9.3%)	38 (10.4%)	79%
autism (16–28)	7 (17.5%)	291 (89.8%)	298 (81.9%)	98%
	TN 62%, FP 38%	TP 99%, FN 1%		
<b>Module 1 Algorithm 2</b>	<b>No ASD (N=129)</b>	<b>ASD (N=405)</b>	<b>All (N=534)</b>	<b>Correctly predicting BEC diagnosis</b>
non-spectrum (0–7)	82 (63.6%)	4 (1%)	86 (16.1%)	95%
autism spectrum (8–11)	28 (21.7%)	36 (8.9%)	64 (12%)	56%
autism (12–28)	19 (14.7%)	365 (90.1%)	384 (71.9%)	95%
	TN 64%, FP 36%	TP 99%, FN 1%		
<b>Module 2 Algorithm 1</b>	<b>No ASD (N=200)</b>	<b>ASD (N=213)</b>	<b>All (N=413)</b>	<b>Correctly predicting BEC diagnosis</b>
non-spectrum (0–6)	138 (69%)	6 (2.8%)	144 (34.9%)	96%
autism spectrum (7–9)	34 (17%)	29 (13.6%)	63 (15.3%)	46%
autism (10–28)	28 (14%)	178 (83.6%)	206 (49.9%)	86%
	TN 69%, FP 31%	TP 97%, FN 3%		
<b>Module 2 Algorithm 2</b>	<b>No ASD (N=98)</b>	<b>ASD (N=159)</b>	<b>All (N=257)</b>	<b>Correctly predicting BEC diagnosis</b>
non-spectrum (0–7)	64 (65.3%)	6 (3.8%)	70 (27.2%)	91%
autism spectrum (8)	11 (11.2%)	3 (1.9%)	14 (5.4%)	21%
autism (9–28)	23 (23.5%)	150 (94.3%)	173 (67.3%)	86%
	TN 65%, FP 35%	TP 96%, FN 4%		
<b>Module 3</b>	<b>No ASD (N=812)</b>	<b>ASD (N=764)</b>	<b>All (N=1576)</b>	<b>Correctly predicting BEC diagnosis</b>
non-spectrum (0–6)	593 (73%)	41 (5.4%)	634 (40.2%)	94%
autism spectrum (7–8)	79 (9.7%)	100 (13.1%)	179 (11.4%)	56%
autism (9–28)	140 (17.2%)	623 (81.5%)	763 (48.4%)	82%
	TN 73%, FP 27%	TP 95%, FN 5%		
<b>Total sample</b>	<b>No ASD (N=1279)</b>	<b>ASD (N=1865)</b>	<b>All (N=3144)</b>	<b>Correctly predicting BEC diagnosis</b>
non-spectrum	902 (70.5%)	60 (3.3%)	962 (30.7%)	94%
autism spectrum	160 (12.5%)	198 (10.6%)	358 (11.3%)	56%
autism	217 (17%)	1607 (86.2%)	1824 (58%)	88%
	TN 71%, FP 29%	TP 97%, FN 3%		

TN: true negative, FP: false positive, TP: true positive, FN: false negative