Diagnostic Accuracy of the ADOS-2 in Children With Psychiatric Conditions

Précision diagnostique de l'ADOS-2 chez les enfants ayant des troubles psychiatriques

Angela Feehan
Shannon Napora
Karen Weis
Danielle Johnston
Keya Clegg Davis
Sharron McKinnon
Lonnie Zwaigenbaum

Angela Feehan[1], Shannon Napora[1], Karen Weis[1], Danielle Johnston[2], Keya Clegg Davis[1], Sharron McKinnon[1], and Lonnie Zwaigenbaum[3]

[1]Glenrose Rehabilitation Hospital, Edmonton, AB, CANADA

[2]Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, AB, CANADA

[3]Department of Pediatrics, University of Alberta, Edmonton, AB, CANADA

Abstract

The Autism Diagnostic Observation Schedule (2nd edition, ADOS-2) shows excellent diagnostic accuracy when used with children suspected of having either autism or language/intellectual delays; however, its accuracy has been lower in children with psychiatric conditions. The purpose of this study was to determine the diagnostic accuracy of the ADOS-2 in pediatric psychiatry patients and to explore factors related to misclassification. Retrospective chart reviews for 84 consecutive autism query referrals in a local child psychiatry program were completed. Patient charts were reviewed for demographic and diagnostic information as well as scores on the ADOS-2 and the Children's Communication Checklist-2. Forty-four of 84 children were ultimately diagnosed with autism. Sensitivity of the ADOS-2 was 93% and specificity was 58%. Positive and negative predictive values were 71% and 89%, respectively. Thus, a negative result on the ADOS-2 was more informative than a positive result. The positive likelihood ratio showed a small difference, and the negative likelihood ratio showed a large difference. Overall, the ADOS-2 produced high rates of false positives in this pediatric psychiatry population. False positives were not related to the total number of psychiatric diagnoses children had received, but children diagnosed with attention-deficit/hyperactivity disorder and anxiety disorders were more likely to receive a false positive result.

Abrégé

La deuxième édition de l'Échelle d'observation pour le diagnostic de l'autisme (*Autism Diagnostic Observation Schedule* [ADOS-2]) offre une excellente précision diagnostique lorsqu'elle est utilisée avec des enfants chez qui un trouble du spectre de l'autisme ou un retard langagier ou intellectuel est suspecté. Sa précision est toutefois moindre chez les enfants ayant des troubles psychiatriques. L'objectif de la présente étude était de déterminer la précision diagnostique de l'ADOS-2 chez des patients et patientes pédiatriques vus en psychiatrie et d'explorer les facteurs associés aux erreurs de classification. Les dossiers médicaux de 84 enfants référés consécutivement en pédopsychiatrie pour une suspicion de trouble du spectre de l'autisme ont été analysés rétrospectivement. Les données démographiques, les diagnostics, ainsi que les scores obtenus à l'ADOS-2 et au questionnaire *Children's Communication Checklist-2,* ont été extraits. Ultimement, 44 des 84 enfants ont reçu un diagnostic de trouble du spectre de l'autisme. La sensibilité de l'ADOS-2 était de 93 % et sa spécificité de 58 %. Les valeurs prédictives positives et négatives étaient respectivement de 71 % et 89 %. Par conséquent, un résultat négatif à l'ADOS-2 était plus informatif qu'un résultat positif. Le rapport de vraisemblance positif indiquait une petite différence, alors que le rapport de vraisemblance négatif indiquait une grande différence. De manière générale, l'ADOS-2 a produit un nombre élevé de faux positifs dans cette population pédopsychiatrique. Le nombre de faux positifs n'était pas associé au nombre total de diagnostics psychiatriques posés aux enfants. Cependant, les enfants ayant reçu un diagnostic de trouble du déficit de l'attention ou d'hyperactivité et de troubles anxieux étaient plus susceptibles de recevoir un résultat faussement positif.

Differential diagnosis of autism is not a straightforward process in individuals with psychiatric conditions. Behavioural features can overlap with autism in conditions such as psychosis, anxiety disorder, and depression. Accurate and timely diagnosis of autism is extremely important as it allows access to services and interventions.

The Autism Diagnostic Observation Schedule (2nd edition, ADOS-2; Lord et al., 2012) is an interactive assessment that uses standardized activities and "presses" to elicit communication, social interaction, and repetitive interests. It is widely considered a key component of a gold standard autism assessment (Kamp-Becker et al., 2018).

Like any clinical assessment, both false positives (FPs) and false negatives (FNs) can occur. The original ADOS (Lord et al., 1989) and ADOS-2 discriminate well between children with autism and those suspected of having language and/or intellectual delays (e.g., Corsello et al., 2013; Lord et al., 2012). A meta-analysis of diagnostic accuracy findings across 14 studies of children assessed with the ADOS-2 found that both sensitivity and specificity were above 80% (Lebersfeld et al., 2021). However, some studies have found high rates of misclassifications (particularly FP) in other types of clinical samples. For example, Molloy et al. (2011) found that the ADOS-2 produced low specificity when the children assessed presented with a broad range of developmental and behavioural disorders. Specifically, Molloy and colleagues measured sensitivity at 79% and specificity at 68% for the module 3 original algorithm when using the autism spectrum disorder (ASD) cutoff. Adults with psychiatric conditions have received high rates of misclassification on the ADOS-2 in some past research as

well (Bastiaansen et al., 2011; de Bildt et al., 2016; Maddox et al., 2017). **Table 1** summarizes these findings.

Some research with adults has investigated the specific characteristics of individuals who receive a FP. A history of psychosis and a diagnosis of schizophrenia have been common in FP cases (Bastiaansen et al., 2011; de Bildt et al., 2016; Maddox et al., 2017). Adamou et al. (2021) found that the Restricted Interests domain of the ADOS-2 module 4 was able to predict autism status. Compared to adults with schizophrenia, adults who received an autism diagnosis were found to have more stereotyped language, less reciprocal social interaction, poorer quality of social response, and poorer quality of rapport (Bastiaansen et al., 2011).

A few studies have shown that children with psychiatric conditions also receive high FP rates on the ADOS-2. Greene et al. (2022) examined diagnostic accuracy of ADOS-2 module 3 from a referred sample of children who presented with high rates of developmental, cognitive, and psychiatric concerns, reporting a sensitivity of 99%, specificity of 65%, and FP rate of 34%. Colombi et al. (2020) found that the ADOS-2 had low rates of both sensitivity and specificity in children and youth with psychiatric conditions. The sensitivity was reported at 58% for module 3 and 56% for module 4. Specificity was reported at 57% for module 3 and 60% for module 4. Forty percent of participants in the Colombi et al. study were misclassified.

Research investigating the specific characteristics of children receiving psychiatry services who are misclassified by the ADOS-2 is limited to date. In one study that piloted

Table 1				
ADOS Diagnostic Accuracy Studies in Adults With Psychiatric Conditions				
Authors (year)	ADOS version	Population	Findings	Data not reported
Bastiaansen et al. (2011)	ADOS	Males with autism, schizophrenia, psychopathy Controls	The ADOS correctly classified 74.2% of cases Mean scores for all ADOS domains were similar for participants with autism and participants with schizophrenia	Sensitivity Specificity Predictive values Likelihood ratios
de Bildt et al. (2016)	ADOS	Males with autism, schizophrenia, psychopathy Controls	Sensitivity for autism group = 55% Specificity for schizophrenia group = 67% Specificity for psychopathy group = 94% Specificity for control group = 95%	Predictive values Likelihood ratios
Maddox et al. (2017)	ADOS-2	Outpatients at community mental health centres	Sensitivity = 1.0 Specificity = .74 Positive predictive value = 25% Negative predictive value = 100%	Likelihood ratios

Note. ADOS data are reported for the autism spectrum disorder cutoff and the original algorithms. ADOS = Autism Diagnostic Observation Schedule.

the use of the ADOS (1st edition) in a psychiatric clinic, clinicians made note of anxiety disorders and attention-deficit/hyperactivity disorder (ADHD) in FP cases (Stadnick et al., 2015); however, only one study thus far has completed a statistical analysis of characteristics in FP cases: Greene et al. (2022). Greene and colleagues found higher rates of FPs in children and adolescents who were male, had low restricted and repetitive behaviour scores, had high anxiety levels during testing, and had trauma-based psychiatric concerns. The complexity of children assessed for autism (i.e., number of psychiatric conditions) has not been previously investigated as a factor in FP outcomes.

Social communication difficulties are a central component of autism. The speech-language pathologist assesses skills in this area of development during the autism differential diagnostic process. Social communication difficulties can also be present for children with other disorders, including ADHD, anxiety, and social (pragmatic) language disorder. The Children's Communication Checklist (2nd edition, CCC-2; Bishop, 2003) is a parent questionnaire that collects information about 10 domains of communication skills in children and has been used to assess social communication in many populations, including children with autism, ADHD, emotional-behavioural disorders, Williams syndrome, and anxiety (e.g., Bignell & Cain, 2007; Mackie & Law, 2010; Philofsky et al., 2007; van Steensel et al., 2013). The CCC-2 has been used in past research to identify communication concerns/disorders and to compare communication skills across diagnostic groups. The Social Interaction Difference Index (SIDI) of the CCC-2 is a measure of social communication skills. It is calculated by taking the difference between the scores of domains of basic communication (e.g., syntax and semantics) and scores of social domains (e.g., nonverbal communication). The SIDI represents an individual's social communication skills relative to their level of basic communication skills. The CCC-2 manual suggests a SIDI cutoff of −11 or below for identifying problems in social communication. In previous research, the SIDI of the CCC-2 has identified between 81% and 95% of children with autism as having social communication challenges (Philofsky et al., 2007; Volden & Phillips, 2010). The use of the CCC-2 in conjunction with the ADOS-2 for diagnosing autism in children with psychiatric conditions has not yet been studied.

Most existing studies that have investigated ADOS-2 diagnostic accuracy in individuals with psychiatric concerns have focused on adult samples. The current study adds to previous literature by reporting diagnostic accuracy for a pediatric sample and providing information about

the characteristics of children who were misclassified by the ADOS-2. This study is the first to investigate social communication test scores with respect to autism diagnosis in this population. This study also adds further interpretability to diagnostic accuracy data by providing positive and negative likelihood ratio calculations, a measure that does not rely on the prevalence of autism in the sample for interpretation.

The overall goal of this study was to evaluate the diagnostic accuracy of the ADOS-2 in children and youth with psychiatric conditions. The first objective was to determine the sensitivity, specificity, negative predictive value (NPV), positive predictive value (PPV), positive likelihood ratios, and negative likelihood ratios of the ADOS-2 in children with psychiatric disorders. The second objective was to examine differences between FPs and TNs with regard to psychiatric diagnosis. Many psychiatric conditions involve a social communication component (e.g., social anxiety, ADHD) or repetitive motor and vocal behaviour (e.g., Tourette's syndrome), and thus, we may see higher rates of FPs in children with a greater number of these conditions. The third objective was to examine differences between FPs and TNs with regard to social language skills. The final objective was to describe subdomain scores on the ADOS-2 for FPs, TNs, FNs, and TPs. Characterizing the diagnostic accuracy of the ADOS-2 and factors that are more likely to lead to FPs can provide clinicians with much needed information about the uses and limitations of the ADOS-2 in autism diagnostic queries in psychiatric programs. We hypothesized that (a) children with a greater number of psychiatric conditions would be more likely to receive a FP outcome on the ADOS-2, (b) children with ADHD and anxiety disorders would be more likely to receive a FP, and (c) CCC-2 SIDI scores would be lower in children with autism compared to children without autism.

Methods

This research was approved by the University of Alberta Human Research Ethics Board (Pro00103319).

Chart Reviews

Retrospective chart reviews were conducted for 84 consecutive ADOS-2 referrals through child psychiatry at a Canadian pediatric rehabilitation hospital. All children were referred for the ADOS-2 in response to an autism query by the psychiatry team. Assessments were conducted between September 2018 and March 2022 and involved an ADOS-2 assessment with the occupational therapist, a psychiatric assessment, a developmental interview with the psychiatrist, a speech-language assessment, and a psychological assessment. The diagnostic team consisted

of three speech-language pathologists (with 8–11 years of experience), two occupational therapists (with 13–21 years of experience), and one psychologist (with 21 years of experience) who consulted across programs. Five different psychiatrists were working in distinct programs within psychiatry and collaborated with the team of consultants for autism queries.

A wide range of language assessment tools were used, including the Clinical Evaluation of Language Fundamentals (5th edition), the Oral and Written Language Scales (2nd edition), and the Comprehensive Assessment of Spoken Language (2nd edition). The CCC-2 was most consistently administered across participants. Psychology tools included the Weschler Intelligence Scale for Children (5th edition), Weschler Preschool and Primary Scale of Intelligence (4th edition), Weschler Adult Intelligence Scale (4th edition), and the Stanford-Binet (5th edition). There was some variability in clinical procedures across programs within the psychiatry department (e.g., some children did not receive psychological testing if it had been previously completed). Decisions regarding autism diagnosis were made by the clinical team in case conference meetings that the patient's psychiatric nurse also attended to provide observations. All information was considered together and diagnoses were based on the *Diagnostic Statistical Manual of Mental Disorders* criteria, (5th ed.; DSM-5; American Psychiatric Association, 2013).

Children seen through both inpatient and outpatient services were included in the chart review. The following information was extracted from each patient's medical chart by a member of the research team: age in years and months, psychiatric diagnosis, yes/no for formal autism diagnosis, yes/no for meeting ASD cutoff on the ADOS-2, ADOS-2 module 3 or 4 scores, CCC-2 scores, yes/no for a diagnosis of language disorder, and full-scale IQ. Information was recorded on an Excel spreadsheet by a member of the study team. Any information that could not be located on the electronic or paper chart was left blank. No identifiable information was recorded from patient charts.

Sensitivity, Specificity, Predictive Values, and Likelihood Ratios

The ADOS-2 consists of four modules for individuals of differing language levels, from minimal or no language (module 1) to fluent language (modules 3 and 4). ADOS-2 scores are categorized as either below or above the ASD cutoff. Scores below the ASD cutoff indicate that a diagnosis of autism is not supported whereas scores above the ASD cutoff indicate that a diagnosis of autism is supported.

Sensitivity, specificity, NPV, and PPV are commonly reported in research as measures of diagnostic accuracy. Sensitivity refers to a test's ability to correctly identify when the target condition is truly present (true positive TP), and specificity refers to a test's ability to correctly identify when the condition is truly absent (true negative TN). PPVs estimate how likely it is that a person who tests positive truly has the condition, and NPVs estimate how likely it is that a person who tests negative truly does not have the condition. Higher sensitivity, specificity, and predictive values indicate a better performing test, with the maximum value being 100% and the lowest being 0%.

Likelihood ratios indicate how probable the test result is for clients with and without the condition. They combine information from sensitivity and specificity into one value and do not rely on clinical prevalence. Negative likelihood ratios look at how having a negative test result changes the chances of an individual having a condition. Negative likelihood ratios range from 1 to 0, with .5 indicating a *small difference*, .2 indicating a *medium difference*, and .1 indicating a *large difference*. For the negative likelihood ratio, values closer to 0 indicate higher importance of the test (Raslich et al., 2007). Positive likelihood ratios look at how having a positive test changes the chances of an individual having the condition. Positive likelihood ratios range from 1 to 10, with 2 indicating *a small difference*, 5 indicating a *moderate difference*, and 10 indicating a *large difference*. Positive likelihood ratios closer to 10 indicate higher importance of the test (Raslich et al., 2007).

Data Analysis

Descriptive statistics (means and standard deviations) were calculated to summarize the sample characteristics. FP, TN, FN, and TP rates were calculated using the ADOS-2 as the index test, and the formal autism diagnosis as the reference standard. Sensitivity, specificity, positive and negative predictive values, and positive and negative likelihood ratios were calculated to examine the performance of the ADOS-2 with this population. Mean ADOS-2 domain scores, mean number of psychiatric diagnoses, and mean scores on the CCC-2 SIDI were calculated for all four groups (FP, TN, FN, TP). Normality of the data was assessed using the Shapiro-Wilk test and the TN group was found to violate the assumptions of normality when comparing mean number of psychiatric diagnoses. Therefore, the nonparametric Mann-Whitney U test was used to examine the difference between the mean number of diagnoses in the FP and TN groups. As a post hoc exploratory analysis, chi-square tests were used to analyze differences in rates of ADHD and anxiety for FP and TN groups. Due to the exploratory nature of this analysis, *p* values were not adjusted (.05 was used).

Because the assumptions of normality were met, independent t tests were used to examine differences between the children without autism (FP and TN groups) and the children with autism (FN and TP groups) for mean CCC-2 SIDI scores and ADOS-2 Social Affect scores. Cohen's *d* was used as a measure of effect size. For ADOS-2 Restricted and Repetitive Behaviour scores, the assumptions of normality were not met, therefore, a Mann-Whitney U was conducted to compare autism and no-autism groups. Eta-squared statistics were calculated to describe the magnitude of difference for Restricted and Repetitive Behaviour. Module 4 test scores were not compared statistically because very few of our participants were given this module. All data analyses were conducted using Statistical Package for Social Sciences (SPSS) software, version 25 (IBM SPSSv25).

Results

Descriptive statistics are presented in **Table 2**. Participants' ages ranged from 6 years, 7 months to 17 years, 11 months, and full-scale IQs ranged from 71 to 132. The groups with and without autism were similar with respect to male-to-female ratio, mean age, and mean IQ. Both groups

also had similar ratios of children who were assessed using ADOS-2 module 3 compared to module 4. Eight psychiatric conditions were listed from the patient charts: the most common co-occurring condition across groups was ADHD and the second most common was anxiety.

Twenty-six children received scores below the ASD cutoff on the ADOS-2; however, three of them still received a diagnosis of autism because the psychiatric team's conclusion, when considering all assessment information together, differed from the ADOS-2 classification (representing FNs). Likewise, 58 children received scores above the ASD cutoff; however, 17 of them were not diagnosed with autism by the team (representing FPs). The number of children in each of the FP, TN, FN, and TP groups is presented in **Table 3** along with diagnostic accuracy of the ADOS-2.

The positive likelihood ratio was 2.19 and showed a small difference. The negative likelihood ratio was 0.12 and showed a large difference. The number of psychiatric diagnoses for children ranged from 1 to 5 (*M* = 2.7) in the FP group, 0 to 3 (*M* = 2.0) in the TN group, 2 to 6 (*M* = 3.7)

Table 2			
Sample Characteristics			
Characteristic	Full sample (autism and no autism diagnosis) *N* = 84*	Autism diagnosis *n* = 44*	No autism diagnosis *n* = 40*
Sex, *n* (%) Female Male	20 (24) 64 (76)	10 (23) 34 (77)	10 (25) 30 (75)
Mean age (*SD*)	12 years, 9 months (3.23)	12 years, 10 months (3.24)	12 years, 8 months (3.26)
Mean IQ (*SD*)	96.90 (15.37) [a]	95.08 (15.03) [b]	98.89 (15.72) [c]
Module 3, *n* (%)	61 (75) [d]	31 (72) [e]	30 (79) [f]
Module 4, *n* (%)	20 (25) [d]	12 (28) [e]	8 (21) [f]
Psychiatric Diagnoses, *n* (%) Attention-deficit/hyperactivity disorder Anxiety disorders Tourette syndrome Obsessive compulsive disorder Learning disability Depression/mood disorder Parent-child relational problem [j] Disruptive behaviour disorder	54 (64) 41 (49) 25 (30) 20 (24) 20 (24) 14 (17) 11 (13) 6 (7)	27 (61) 20 (45) 14 (32) 7 (16) 13 (30) 7 (16) 5 (11) 2 (5)	27 (68) 21 (53) 11 (28) 13 (33) 7 (18) 7 (18) 6 (15) 4 (10)
Developmental language disorder, *n* (%) [k]	22 (30) [g]	14 (35) [h]	8 (24) [i]

Note. Due to missing data, numbers vary in some categories as listed in specific notes below.
[a] *n* = 73. [b] *n* = 38. [c] *n* = 35 [d] *n* = 81. [e] *n* = 43. [f] *n* = 38. [g] *n* = 73. [h] *n* = 40. [i] *n* = 33. [j] This is listed as a "condition" (rather than a diagnosis) in the *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; American Psychiatric Association, 2013). [k] For children diagnosed with autism, the diagnostic label is "Language disorder in the context of autism spectrum disorder."

Table 3

Sensitivity, Specificity, and Predictive Power of the ADOS-2

ADOS-2 results	Diagnosed with autism n (%)	Not diagnosed with autism n (%)	Predictive value %
Autism	41 (49) True positives	17 (20) False positives	PPV = 70.7
No autism	3 (4) False negatives	23 (27) True negatives	NPV = 88.5
	Sensitivity = 93.2%	Specificity = 57.5%	

Note. ADOS-2 = Autism Diagnostic Observation Schedule (2nd edition); PPV = positive predictive value; NPV = negative predictive value.

in the FN group, and 0 to 4 (M = 2.0) in the TP group (see **Table 4**). The difference between the mean number of psychiatric diagnoses in the FP and TN groups was not significant (U = 139.000, N_1 = 23, N_2 = 17, p = .126). Differences in diagnosis rates for FP and TN groups were significant for ADHD, X^2 (1, N = 40) = 5.80, p = .016, and anxiety, X^2 (1, N = 40) = 3.88, p = .049.

Table 5 presents the social interaction and ADOS-2 domain scores by classification. The range of SIDI scores was 6 to −24 (n = 13) for the FP group, 14 to −20 (n = 12) for the TN group, −11 to −28 (n = 3) for the FN group, and −2 to −37 (n = 29) for the TP group. Of children who received an autism diagnosis, 69% had SIDI scores at or below the −11 cutoff. Of children who did not receive a diagnosis of autism, 56% had SIDI scores at or below the −11 cutoff. Children with an autism diagnosis had lower scores on the CCC-2 SIDI (M = −14.00) than did children without an autism diagnosis (M = −10.24). This difference was significant, and the size

Table 4

Psychiatric and Language Diagnosis by ADOS-2 Classification

Diagnoses	No autism group		Autism group	
	False positive diagnosis (n = 17)	True negative diagnosis (n = 23)	False negative diagnosis (n = 3)	True positive diagnosis (n = 41)
Total number of psychiatric diagnoses, M (SD)	2.71 (1.11)	2.00 (1.09)	3.67 (2.08)	1.95 (0.95)
	3.00 (1–5) [a]	2.00 (0–3) [a]		
Psychiatric diagnoses, n (%)				
Attention-deficit/hyperactivity disorder	15[*] (88)	12[*] (52)	1 (33)	26 (63)
Anxiety disorders	12[*] (71)	9[*] (39)	0 (0)	20 (49)
Tourette syndrome	6 (35)	5 (22)	1 (33)	13 (32)
Obsessive compulsive disorder	4 (24)	9 (39)	0 (0)	7 (17)
Learning disability	4 (24)	3 (13)	3 (100)	10 (24)
Depression/mood disorder	3 (18)	4 (17)	2 (67)	5 (12)
Parent-child relational problem	2 (12)	4 (17)	3 (100)	2 (5)
Disruptive behaviour disorder	1 (6)	3 (13)	1 (33)	1 (2)
Developmental language disorder, n (%) [b]	5 (31) [c]	3 (18) [d]	1 (100) [e]	13 (33) [f]

[a] Median (range) are reported because assumptions of normality were violated. Results were not significant. [b] Numbers vary due to missing data. [c] n = 16. [d] n = 17. [e] n = 1. [f] n = 39.
[*] p < .05

Table 5

Social Interaction and ADOS-2 Domain Scores by False Positive, True Negative, False Negative, and True Positive Classification

Measure	No autism group		Autism group	
	False positive diagnosis	**True negative diagnosis**	**False negative diagnosis**	**True positive diagnosis**
Children's Communication Checklist (2nd edition): Social Interaction Difference Index, M (SD)	−10.23 (8.57) $n = 13$	−10.25 (9.28) $n = 12$	−16.67 (9.82) $n = 3$	−13.72 (7.00) $n = 29$
	−10.24[*] (8.72) $n = 25$		−14.00[*] (7.16) $n = 32$	
ADOS-2 domain scores, M (SD) Social affect, module 3	8.07 (2.15) $n = 15$	4.33 (1.68) $n = 15$	3.50 (3.54) $n = 2$	9.55 (2.41) $n = 29$
	6.2[*] (2.68) $n = 30$		9.16[*] (2.85) $n = 31$	
Restricted and repetitive behavior, module 3	1.5 (1.41) $n = 15$	0.47 (0.92) $n = 15$	0.00 (0.00) $n = 2$	1.59 (1.02) $n = 29$
	1.0 (1.29) 0.5 [a*] (0–4) $n = 30$		1.48 (1.06) 1.0 [a*] (0–4) $n = 31$	
Social communication total, module 4	13.00 (0.00) $n = 2$	6.00 (2.10) $n = 6$	3.00 (0.00) $n = 1$	10.36 (3.50) $n = 11$
	7.75 (3.69) $n = 8$		9.75 (3.96) $n = 12$	
Stereotyped behaviours and restricted interests total, module 4	1.00 (0.00) $n = 2$	0.83 (0.75) $n = 6$	0.00 (0.00) $n = 1$	1.73 (1.19) $n = 11$
	0.88 (0.64) $n = 8$		1.58 (1.24) $n = 12$	

[a] Median (range) are reported because assumptions of normality were violated.
[*] $p < .05$

of the effect was moderate ($t = −1.788$, $df = 55$, $p = .0395$, one tailed, $d = 0.47$). Children with an autism diagnosis had higher scores on the Social Affect domain ($M = 9.16$) than did children without an autism diagnosis ($M = 6.2$). This difference was significant, and the size of the effect was large (t = 4.173, $df = 59$, $p = .00$, one tailed, $d = 1.069$). Children with an autism diagnosis had higher scores on the Restricted and Repetitive Behaviour domain ($M = 1.48$) than children without an autism diagnosis ($M = 1.00$), the difference was significant (U = 330.000, $N_1 = 30$, $N_2 = 31$, $p = .043$), and the size of the effect was large ($\eta^2 = .132$).

Discussion

In our study, sensitivity and specificity values, predictive values, and likelihood ratios all showed that a negative result on the ADOS-2 is more informative than a positive result when children have co-occurring psychiatric conditions. Sensitivity and NPV were higher than specificity and PPV. Like in our sample, higher sensitivity relative to specificity has frequently been found in adult psychiatric samples referred for autism evaluation (i.e., Bastiaansen et al., 2011; de Bildt et al., 2016; Maddox et al., 2017). Our sensitivity and specificity values are comparable to those reported by Greene et al. (2022) in their sample with high rates of developmental, cognitive, and psychiatric concerns. This means that, although the ADOS-2 ASD cutoff is effective at capturing individuals with autism, many individuals without autism also receive elevated scores on this assessment tool. High FP rates mean that clinicians using the ADOS-2 with psychiatric populations must be wary of overdiagnosing autism. Our findings indicate that around 40% of children meeting the ASD cutoff on the ADOS-2 did not meet DSM-5

criteria for a formal autism diagnosis. A comprehensive assessment that involves collecting a detailed developmental history, history of presenting symptoms including cognitive ability, language development/skills, and behavioural symptoms is the hallmark of autism evaluation and must be used in conjunction with the ADOS-2.

As a child's psychiatric presentation becomes more complex, determining if an additional diagnosis of autism is appropriate can become more challenging. In the present study, complexity was defined as the number of additional diagnoses the child presented with. Contrary to what we expected, the mean number of diagnoses did not differ significantly across FP and TN groups; however, the range of additional conditions across the two groups did differ descriptively, with those who had a higher number of conditions (4–5) all falling into the FP group and those who had zero additional conditions all falling into the TN group. The ADOS-2 scoring may not be sensitive to the number of conditions a child has per se, but to the additive effect of social communication and behavioural differences associated with ADHD and/or anxiety. Like our findings, previous research has noted high rates of ADHD and anxiety disorders in FP cases (Stadnick et al., 2015). Scores above cutoff on the ADOS-2 for children with these conditions may be more common as the symptoms could affect observable behaviours during the assessment. Unfortunately, information about anxiety levels during the testing session was not available on the patient charts, so we could not corroborate the findings of Greene et al. (2022) that test levels of anxiety were high in FP cases.

The possibility of using the CCC-2 SIDI as a measure of social communication to complement the ADOS-2 was also investigated in this research. The recommended SIDI cutoff score of –11 surprisingly only identified 69% of children with autism as having social communication challenges. This is lower than numbers published in previous research (81%–95%; Philofsky et al., 2007; Volden & Phillips, 2010). Because social communication differences are a core feature of autism, we would have expected a higher percentage of children to be identified with social communication challenges. The CCC-2 SIDI is a measure of the discrepancy between basic communication skills and social communication skills and thus may miss out on identifying children with social communication challenges who do not have this discrepancy (who also have low basic language skills). The CCC-2 SIDI also identified many children without autism as having social communication challenges (56%). Not surprisingly, this indicates a high prevalence of social communication concerns in children referred for autism assessment through child psychiatry. Overall, our results

indicate that we can generally expect lower SIDI scores for children with autism and higher scores for children without autism; however, there were children with autism who received scores above the cutoff (in the –2 to –11 range) and children without autism who received very low SIDI scores (as low as –24). All children with scores below –25 received an autism diagnosis. Interestingly, all FNs received a SIDI at or below the –11 cutoff; however, the number of FNs was too small in our sample to make any conclusions regarding the utility of the SIDI for identifying FNs.

Significant differences between module 3 Social Affect scores for the autism and no-autism groups are not surprising given the importance of affective differences in the autism population. We also expected to see a difference between ADOS-2 scores in the autism versus no-autism groups for Restricted and Repetitive Behaviours, because this domain of the test captures the second core feature of autism (characteristics such as focused interests, sensory processing differences, difficulty with change, and stereotyped motor movements like hand flapping, toe walking, rocking, etc.).

ADOS-2 FPs in adult populations have been linked to a diagnosis of psychosis and schizophrenia in past research (Bastiaansen et al., 2011; de Bildt et al., 2016; Maddox et al., 2017). These diagnoses are rarely given during childhood and adolescence and no cases were observed in our sample. Adult literature also cites high restricted interests, greater stereotyped language, poorer reciprocal social interaction, poorer quality of social response, and poorer quality of rapport on the ADOS-2 as factors that can differentiate FPs from individuals with autism (Adamou et al., 2021; Bastiaansen et al., 2011). We have presented domain scores for module 4 descriptively. Greene et al. (2022) found that children who received a FP on the ADOS-2 were more likely to be male and often had a positive history of trauma. We did not consider these variables in our analysis as none of our participants presented with trauma-related diagnoses and a high number of males were seen across all groups due to the composition of our sample.

Future Directions

Future research should consider how to measure the complexity of psychiatric cases referred for autism assessment and how complexity may relate to a FP outcome on the ADOS-2. Following these cohorts over time could help to characterize those who do and do not meet criteria for an autism diagnosis and to understand their performance on the ADOS-2, particularly with respect to diagnoses that are given later in adulthood, such as psychosis and schizophrenia.

Small numbers of participants with specific conditions (e.g., OCD) made certain comparisons difficult in our sample. Future research with larger sample sizes is needed to investigate ADOS-2 classification accuracy for children with specific conditions. Future research can also focus on characterizing those who receive a FN outcome. Finally, there is a need for ADOS-2 diagnostic accuracy data in different age groups including older youth and young adults.

Limitations

One limitation of this research was that some data could not be located on patient charts, which led to missing data points in some cases and a differing *n* across variables. This is a limitation of using retrospective chart analysis as a data collection method.

Our study was adequately powered to address the primary research questions; however, statistical analysis of some additional variables was not possible due to the sample size. For example, it was difficult to characterize the children who received FNs in our sample, as our numbers were very low in this group. Colombi et al. (2020) found that their sample of children referred for an autism assessment through psychiatry had high FNs. Because there is a possibility of high FNs in some samples, information about the characteristics of these children can be helpful in predicting and identifying FNs. In this study, we used the team's diagnostic decision at a single point in time and did not monitor the stability of the diagnostic decision over time; therefore, changes to autism diagnosis could not be accounted for in our results.

Conclusions

The ADOS-2 produced a high number of FP outcomes in this sample of children with psychiatric conditions evaluated for autism. Caution is needed in interpreting positive ADOS-2 results when individuals referred for autism evaluation present with psychiatric concerns. Any comprehensive assessment for autism will combine a number of tools and should utilize a team approach to collecting information. With low numbers of FNs, the ADOS-2 continues to be a useful clinical tool in child psychiatry, as it is sensitive for identifying individuals who may be autistic.

References

Adamou, M., Jones., S. L., & Wetherhill, S. (2021). Predicting diagnostic outcome in adult autism spectrum disorder using the autism diagnostic observation schedule, second edition. *BMC Psychiatry, 21*, Article 24. https://doi.org/10.1186/s12888-020-03028-7

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Bastiaansen, J. A., Meffert, H., Hein, S., Huizinga, P., Ketelaars, C., Pijnenborg, M., Bartels, A., Minderaa, R., Keysers, C., & de Bildt, A. (2011). Diagnosing autism spectrum disorders in adults: The use of Autism Diagnostic Observation Schedule (ADOS) module 4. *Journal of Autism and Developmental Disorders, 41*, 1256–1266. https://doi.org/10.1007/s10803-010-1157-x

Bignell, S., & Cain, K. (2007). Pragmatic aspects of communication and language comprehension in groups of children differentiated by teacher ratings of inattention and hyperactivity. *British Journal of Developmental Psychology, 25*(4), 499–512. https://doi.org/10.1348/026151006X171343

Bishop, D. (2003). *Children's Communication Checklist* (2nd ed.). Pearson.

Colombi, C., Fish., A., & Ghaziuddin, M. (2020). Utility of the ADOS-2 in children with psychiatric disorders. *European Child and Adolescent Psychiatry, 29*(7), 989–992. https://doi.org/10.1007/s00787-019-01411-8

Corsello, C. M., Akshoomoff, N., & Stahmer, A. C. (2013). Diagnosis of autism spectrum disorders in 2-year-olds: A study of community practice. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 54*(2), 178–185. https://doi.org/10.1111/j.1469-7610.2012.02607.x

de Bildt, A., Sytema, S., Meffert, H., & Bastiaansen, J. A. (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*, 21–30. https://doi.org/10.1007/s10803-015-2532-4

Greene, R. K., Vasile, I., Bradbury, K. R., Olsen, A., & Duvall, S. W. (2022). Autism Diagnostic Observation Schedule (ADOS-2) elevations in a clinical sample of children and adolescents who do not have autism: Phenotypic profiles of false positives. *Clinical Neuropsychologist, 36*(5), 943–959. https://doi.org/10.1080/13854046.2021.1942220

Kamp-Becker, I., Albertowski, K., Becker, J., Ghahreman, M., Langmann, A., Mingebach, T., Poustka, L., Weber, L., Schmidt, H., Smidt, J., Stehr, T., Roessner, V., Kucharczyk, K., Wolff, N., & Stroh, S. (2018). Diagnostic accuracy of the ADOS and ADOS-2 in clinical practice. *European Child and Adolescent Psychiatry, 27*, 1193–1207. https://doi.org/10.1007/s00787-018-1143-y

Lebersfeld, J. B., Swanson, M., Clesi, C. D., & O'Kelley, S. E. (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, 51*(11), 4101–4114. https://doi.org/10.1007/s10803-020-04839-z

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). *Autism Diagnostic Observation Schedule* (2nd ed.). Western Psychological Services.

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., & Schopler, E. (1989). Autism Diagnostic Observation Schedule: A standardized observation of communicative and social behavior. *Journal of Autism and Developmental Disorders, 19*(2), 185–212. https://doi.org/10.1007/BF02211841

Mackie, L., & Law, J. (2010). Pragmatic language and the child with emotional/behavioural difficulties (EBD): A pilot study exploring the interaction between behaviour and communication disability. *International Journal of Language and Communication Disorders, 45*(4), 397–410. https://doi.org/10.3109/13682820903105137

Maddox, B. B., Brodkin, E. S., Calkins, M. E., Shea, K., Mullan, K., Hostager, J., Mandell, D. S., & 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. https://doi.org/10.1007/s10803-017-3188-z

Molloy, C. A., Murray, D. S., 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. https://doi.org/10.1177/1362361310379241

Philofsky, A., Fidler, D. J., & Hepburn, S. (2007). Pragmatic language profiles of school-age children with autism spectrum disorders and Williams syndrome. *American Journal of Speech Language Pathology, 16*(4), 368–380. https://doi.org/10.1044/1058-0360(2007/040)

Raslich, M. A., Markert, R. J., & Stutes, S. A. (2007). Selecting and interpreting diagnostic tests. *Biochemia Medica, 17*(2),151–161. https://doi.org/10.11613/bm.2007.014

Stadnick, N., Brookman-Frazee, L., Nguyen Williams, K., 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. https://doi.org/10.1016/j.rasd.2015.08.007

van Steensel, F. J. A., Bögels, S. M., & Wood, J. J. (2013). Autism spectrum traits in children with anxiety disorders. *Journal of Autism and Developmental Disorders, 43*(2), 361–370. https://doi.org/10.1007/s10803-012-1575-z

Volden, J., & Phillips, L. (2010). Measuring pragmatic language in speakers with autism spectrum disorders: Comparing the Children's Communication Checklist-2 and the Test of Pragmatic Language. *American Journal of Speech Language Pathology, 19*(3), 204–212. https://doi.org/10.1044/1058-0360(2010/09-0011)

Authors' Note

Correspondence concerning the article should be addressed to Angela Feehan, Alberta Health Services, Glenrose Rehabilitation Hospital, 10230 111 Ave. – Rm. 337, Edmonton, AB, CANADA, T5G 0B7.
Email: **afeehan@ualberta.ca**

Disclosures

No conflicts of interest, financial or otherwise, are declared by the authors.