Validating the Autism Diagnostic Interview-Revised in the Korean Population

Miae Oh[1]*, Da-Yea Song[2]*, Guiyoung Bong[2], Nan-He Yoon[3], So Yoon Kim[4],
Joo-Hyun Kim[2], Jongmyeong Kim[2], and Hee Jeong Yoo[2,5] ✉

[1]Department of Psychiatry, Kyung Hee University Hospital, Seoul, Republic of Korea
[2]Department of Psychiatry, Seoul National University Bundang Hospital, Seongnam, Republic of Korea
[3]Department of Health Administration, Hanyang Cyber University, Seoul, Republic of Korea
[4]Teacher Education, Duksung Women's University, Seoul, Republic of Korea
[5]Seoul National University College of Medicine, Seoul, Republic of Korea

Objective   This study aimed to examine the validity of the Korean version of the Autism Diagnostic Interview-Revised (K-ADI-R) and determine its efficacy in identifying individuals with autism spectrum disorder (ASD).

Methods   Data were pooled from several past and ongoing studies as well as clinical records acquired at Seoul National University Bundang Hospital from 2008 to 2017. The K-ADI-R were administered and scored by trained research reliable examiners. Measurements to investigate the validity of the K-ADI-R was through sensitivity, specificity, positive predictive values (PPV), negative predictive values (NPV), and Cohen's kappa.

Results   A total of 1,271 (age 88.9±62.42 months, male=927) participants were included. The K-ADI-R yielded strong psychometric properties with high sensitivity (86.06–99.27%), specificity (84.75–99.55%), PPV (92.33–99.72%), and NPV (79.43–98.64%). There were significant differences in item scores across the K-ADI-R diagnostic algorithm regardless of age and sex (p<0.001). Agreement between the K-ADI-R and other ASD related measurements ranged between levels of good to excellent.

Conclusion   Despite language or cultural boundaries, the K-ADI-R demonstrated high levels of sensitivity, specificity, PPV, and NPV within a wide range of participants; hence, suggesting promising usage as a valuable diagnostic instrument for individuals with ASD.
Psychiatry Investig 2021;18(3):196-204

Key Words   Autism spectrum disorder, Autism Diagnostic Interview-Revised, Validation, K-ADI-R.

INTRODUCTION

Autism Spectrum Disorder (ASD) is defined by persistent deficits in social communication, social interaction, and restricted, repetitive patterns of behavior, interests, or activities.[1] The prevalence of ASD has risen significantly in recent years and is considered one of the fastest-growing developmental disabilities that affect individuals worldwide.[2,3] As ASD influences multiple domains of an individual throughout their lifespan, the timing of detection is crucial to ensure early linkage to care and optimize long-term outcomes.[4-6] However, manifesting a heterogeneous phenotype with continuous variation, diagnosing ASD can be complex based on the need for extensive information ranging from early childhood development to school life and social relationships with peers.[7]

While initially developed for research purposes, the Autism Diagnostic Interview-Revised (ADI-R) and the Autism Diagnostic Observation Schedule (ADOS) are considered the gold standard diagnostic instruments for ASD.[8-11] Both closely in parallel with the Diagnostic and Statistical Manual of Mental Disorders-IV (DSM-IV-TR) dimensions and the ICD-10 criteria, the ADI-R and ADOS are often administered together to improve consensus with clinical judgments by using multiple sources of information acquired on an individual's past and current behavior.[12-14]

The ADI-R is a semi-structured interview administered by trained examiners to caregivers of individuals with suspected

ASD. The assessment solicits information from caregivers that focuses on the domains of: 1) social interactions, 2) communication, 3) restricted, repetitive and stereotyped behaviors (RRB), and 4) the period in which concerns or ASD related symptoms became apparent. Current and past behaviors of each item are scored by the examiner based on the level of severity, with higher scores indicating abnormality, based on the evidence provided during the interview. Depending on the individual's age, a diagnostic algorithm systematically combines a subset of the questions that result in the classification of ASD or non-ASD based on whether the domain scores exceed the cut-off points.

Having demonstrated good to excellent levels of sensitivity and specificity of diagnostic validity across varying samples, the ADI-R has gained attention internationally to examine the equivalence of its psychometric properties across diverse cultures.[15,16] Previous research has shown that adaptation of instruments is a complicated and challenging process that involves linguistic translation and adjusting to cultural values or customs while considering its feasibility.[17,18] Despite increased awareness, understanding, and recognition of ASD, different cultures vary in the levels of over or under-reporting concerns due to different standards to which behaviors are considered acceptable and expected.[19,20] Hence, there is an increasing need to adequately assess the compatibility of the ADI-R in different population groups to ensure the instrument's appropriateness.

To date, the ADI-R has been translated into 17 languages, with validation studies conducted in Germany, Greece, Finland, Brazil, Japan, the Netherlands, China, Sweden, the Latino population in the United States, South Korea, and Poland.[18,21-30] While all of the studies have reported promising results for the use of the ADI-R in clinical and research settings, most have been restricted to small sample sizes or based on participants with relatively narrow age ranges, which further limits the generalizability of the findings. For instance, Kim et al.[29] demonstrated moderate-to-high classifications of individuals with ASD using the ADI-R in South Korea. However, this study only included school-age (7−14-year-old) children with a lack of evidence for its use on younger children, adolescents, and adults. Additionally, with concerns being raised regarding individual factors (i.e., age or intelligent quotient) as well as administrative factors influencing the ADI-R results, additional research on applicability is required.[31,32] Therefore, the purpose of this study is to further expand on previous studies and investigate the diagnostic validity of the Korean translated version of ADI-R (K-ADI-R) in a wider age range and examine whether its ability to detect individuals with ASD is comparable to the original validation conducted by Lord et al.[8]

METHODS

Participants and data collection

Data was compiled from past and ongoing projects such as a genetic study to identify ASD related biomarkers, a social skills training intervention for individuals with ASD, the development of an early ASD screening instrument, and by obtaining the outpatient clinical records at Seoul National University Bundang Hospital (SNUBH) from 2008 to 2017. Participants were recruited through the combination of diverse routes (i.e., the psychiatric clinic and Pediatric and Child Rehabilitation clinic at SNUBH, local primary clinics, community health and mental health centers, daycare centers, advertisements on both online and offline bulletin boards at public institutions, and with the aid of parents' self-help communities). While specific projects (i.e., the genetics study) only included participants whose biological parents are Korean, other studies did not limit participation based on ethnicity but required caregivers to comprehend interview questions and were comfortable being evaluated by the K-ADI-R. Each participant provided informed consent to the corresponding study they were enrolled in, and the retrospective analysis of the collected data was approved by the Institutional Review Board (IRB no. B-1711-435-106) at SNUBH to address the aim of the present study.

As seen in Table 1, a total of 1,271 children, adolescents, and adults with ages ranging from 24 months to 34 years old, comprised of 825 participants with ASD (males; 84.6%), 446 non-ASD participants (males; 51.1%), were included for analyses. Participants included individuals with ASD, unaffected siblings, and typically developing individuals. A subgroup of the participants in the non-ASD group who scored lower than 80 points on either the full-scale intelligence quotient (FSIQ), the Korean version of the Vineland Adaptive Behavior Scale, second edition (K-VABS), or the Korean Vineland Social Maturity Scale (K-SMS) while not meeting the ASD diagnostic criteria were separately categorized as other developmental disorders (OD). In addition to a battery of parent-report questionnaires, described in further detail below, trained research reliable examiners administered the Korean translated versions of ADOS and ADI-R during the participants one time visit on-site. Sessions of the diagnostic evaluations were video-recorded and checked to ensure adequate levels of inter-rater reliability. 43.4% of the videos were watched by two independent raters, while another 10% of the cases were viewed together during weekly research meetings. Upon reviewing the comprehensive information gathered, two board-certified psychiatrists made the best clinical diagnosis by following the DSM-IV-TR and DSM-5 criteria.

Table 1. Participant characteristics between the ASD and non-ASD groups

Characteristics	N	ASD (N=825; 65%) Mean (SD)	Non-ASD (N=446; 35%) Mean (SD)	Total (N=1,271) Mean (SD)	t or χ²
Male (%)†	1,271	84.6	51.1	72.9	164.14***
Age (months)‡	1,271	92.2 (64.1)	82.9 (58.6)	88.9 (62.4)	2.53**
FSIQ‡	647	82.5 (26.6)	105.3 (18.9)	90.5 (26.5)	-12.64***
ADI-R score‡					
Algorithm A	1,271	19.9 (6.4)	2.7 (3.1)	13.9 (9.9)	64.34***
Algorithm B verbal	864	15.9 (4.8)	1.7 (2.3)	10.0 (8.0)	57.20***
Algorithm B nonverbal	407	10.8 (2.9)	2.3 (2.1)	8.9 (4.5)	30.41***
Algorithm B nonverbal only	1,271	10.0 (3.3)	1.4 (1.8)	7.0 (5.0)	60.34***
Algorithm C	1,271	5.3 (2.6)	0.7 (1.1)	3.7 (3.1)	45.23***
Algorithm D	1,271	3.6 (1.2)	0.4 (0.9)	2.5 (1.9)	54.91***
Level of functioning (%)†					
Language delay	1,271	48.7	6.5	33.3	230.31***
Regression of language	1,271	5.2	0.9	3.7	15.14***
Loss of skills	1,271	3.6	0.0	2.4	16.61***

**p<0.01, ***p<0.001, †chi-square test, ‡t-test. ASD: autism spectrum disorder, FSIQ: full scale intelligence quotient, ADI-R: Autism Diagnostic Interview-Revised, Algorithm A: Social Interaction, Algorithm B: Communication, Algorithm C: Restricted, Repetitive, and Stereotyped Patterns of Behavior, Algorithm D: Abnormality of development before 36 months

Measurements

Autism Diagnostic Interview-Revised (ADI-R)[9]

The ADI-R is a semi-structured interview that allows parents or primary caregivers to report on a child's current behavior as well as reflect on the early developmental history, provided that he or she has a mental age above two years. Consisting of 93 questions, the ADI-R uses the information passed on by the caregiver to rate each item on a scale of 0 (socially appropriate) to 3 (evidence of severe abnormality). Depending on the individual's age, scores are converted, and a diagnostic algorithm made up by a subset of the question items are summed to four domains of 1) social interactions; 2) communication (verbal and non-verbal); 3) RRB; and 4) whether developmental concerns were present before three years of age. All four domains need to reach the cut-off limits to meet the ASD diagnostic criteria. The K-ADI-R used in the present study was translated and back-translated by Yoo et al.[33] and approved through Western Psychological Services.

Autism Diagnostic Observation Schedule (ADOS) and Autism Diagnostic Observation Schedule-second edition (ADOS-2)[10,11]

The ADOS and ADOS-2 are an observation instrument using play-based methods to assess communication and social behaviors in a series of standardized contexts. Participants' behavior or interactions during the assessment are recorded by trained examiners and used to determine item scores in verbal and non-verbal communication, social interaction, imagination/creativity, and RRBs. Taking into account the developmental trajectory of an individual by age and expressive verbal abilities, the ADOS-2 consists of five modules, each with slightly different tasks and diagnostic algorithms. Summation of the converted scores for a batch of items results in the domains of Social Affect and RRBs. Although each module has distinct combinations of items and cut-offs points, the total score derived from the addition of the two domains can classify individuals of autism, ASD, and non-spectrum combined. While the Western Psychological Services approved both versions of the Korean translated ADOS[34] and ADOS-2,[35] the latter was administered for toddlers and any participant data collected following the publication of the second edition.

Korean version of the Childhood Autism Rating Scale (K-CARS)[36]

The K-CARS is a rating scale used to identify the presence of ASD related behaviors and assess the severity of symptoms in children over 24 months. Fifteen items with scores ranging from 1 (appropriate behavior for age level) to 4 (severe deviance compared to age level) are rated by evaluators using information gathered from clinical observations, caregiver interviews, and other questionnaires. Domain scores are added together to derive a total, which then classifies an individual as non-ASD or indicates mild, moderate, or severe ASD symp-

toms. While studies have well documented the psychometrics of CARS, the cut-off scores used in this study were set to 28 grounded on previous validations in Korea.[37]

Korean Vineland Social Maturity Scale (K-SMS)[38]

The K-SMS, based on Doll's Vineland Social Maturity Scale (SMS),[39] is used as an interview and behavior-observation scale to evaluate social competence and adaptive functioning. Conducted with a caregiver who is familiar with the person being assessed, it can be administered to individuals from birth up to 30 years of age. Six domains, each organized into year levels, were built upon the standardization using a representative sample of Korean participants.[40] Results of the K-SMS can be used to determine the social maturity and social quotient of an individual.

Social Reciprocity Scale (SRS)[41]

The SRS is a 65-item caregiver-report questionnaire that is often used as a screening instrument to recognize ASD-related behaviors and capture its severity. Having high internal consistency and good discriminant validity, the SRS has been useful in distinguishing individuals with and without ASD.[42,43] Items are rated on a 4-point Likert scale with higher T-scores suggesting more significant impairments. The Korean translated version of the SRS was approved for usage in each of the past and ongoing studies by the Western Psychological Services.

Social Communication Questionnaire (SCQ)[44]

The SCQ is a 40-item questionnaire based on the ADI-R. Items are rated as "yes" or "no" by caregivers of individuals aged 24 months or above. There are two versions of the SCQ: the current form that focuses on the child's behavior in the last three months and the lifetime form that asks about the developmental history in the past 12 months. Behaviors are assessed on the domains of social interaction, language and communication, and RRB. Initially intended as a screening instrument for children who are four years of age, the SCQ threshold has been adjusted by researchers when including younger children.[45] In Korea, Kim et al.[46] reported that cut off scores of 10 points for children under 47 months and 12 points for children over 48 months were most effective at maximizing the sensitivity and specificity of individuals with ASD.

Korean version of the Vineland Adaptive Behavior Scale, second edition (K-VABS)[47,48]

The K-VABS questionnaire was used to measure an individual's adaptive functioning from birth through 90 years of age. Items are arranged in order of developmental sequence across nine subdomains. Caregivers rate how often each item is performed on a 3-point scale. Scores are standardized with the mean score of 100 and standard deviations of 15 points into domains of Communication, Socialization, Daily Living Skills, and Motor Skills. An overall rating derived by taking into account all four domains were used to determine whether an individual has OD.

Statistical analyses

Group differences of participant characteristics for categorical variables were analyzed using a chi-square test, while continuous variables such as the K-ADI-R domain and subdomain scores were examined using independent sample t-tests. Further investigations were performed by independent sample t-tests to see whether there were effects of sex on the K-ADI-R scores and was followed by an additional exploration looking into age differences (i.e., children, adolescents, adults). Additionally, following the ADI-R diagnostic algorithm, age groups were divided into verbal and non-verbal individuals below 47 months and those who were 48 months and older.

Consensus regarding participant categorization by the K-ADI-R algorithm standards and the final clinical diagnosis was evaluated by sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and Cohen's kappa (k). To view whether there was a difference in diagnostic algorithm standards, the analysis was conducted in two approaches. While the diagnostic algorithm requires individuals to equal or exceed the thresholds across all four domains to be considered in the ASD group, additional analysis with participants who met at least one of the cut-off score domains were included. Agreement between the K-ADI-R results with other instruments measuring ASD related traits or symptoms was weighed by k coefficients and interpreted based on the division proposed by Landis and Koch (Slight: 0−0.20; Fair: 0.21−0.40; Moderate: 0.41−0.60; Substantial: 0.61−0.80; Almost Perfect: 0.81−1).

To examine whether the K-ADI-R could differentiate individuals with ASD from individuals with OD, sensitivity, specificity, PPV, NPV, and Cohen's kappa were calculated between the two groups. As each research study followed different protocols and acquired various measurements, only those participants who had available data to categorize groups with diagnostic certainty were included for further analysis. All statistical analyses were performed using SPSS Statistics 26.0 (IBM Corp., Armonk, NY, USA).

RESULTS

Of the total 1,271 participants, 72.9% (n=927) were male, and there was a statistically significant difference (p<0.001) in sex between the ASD (84.6% male) and non-ASD (51.1%) group. Based on the DSM-IV criteria, out of the 825 individ-

Table 2. Mean scores of the K-ADI-R diagnostic algorithm per domain by age group

	ASD	Non-ASD	Total	t†
	Mean (SD)	Mean (SD)	Mean (SD)	
Children	(N=662; 63%)	(N=386; 37%)	(N=1,048)	
Algorithm A				
A1	3.72 (1.59)	0.33 (0.65)	2.47 (2.10)	48.56***
A2	5.40 (2.47)	0.96 (1.25)	3.76 (3.00)	38.50***
A3	4.44 (1.55)	0.78 (1.09)	3.09 (2.25)	44.63***
A3	6.28 (2.24)	0.85 (1.16)	4.28 (3.24)	51.60***
Algorithm B				
B1	5.42 (2.38)	0.50 (1.05)	3.61 (3.10)	45.98***
B2	1.79 (1.84)	0.26 (0.72)	1.23 (1.70)	18.92***
B3	1.63 (1.98)	0.25 (0.62)	1.12 (1.75)	16.60***
B4	4.67 (1.43)	0.94 (1.18)	3.29 (2.25)	45.52***
Algorithm C				
C1	1.85 (1.28)	0.24 (0.60)	1.26 (1.33)	27.50***
C2	1.05 (1.19)	0.15 (0.46)	0.72 (1.08)	17.33***
C3	0.88 (0.91)	0.05 (0.25)	0.57 (0.84)	21.96***
C4	1.52 (0.68)	0.27 (0.53)	1.06 (0.87)	33.52***
Algorithm D				
D	3.74 (1.16)	0.41 (0.93)	2.51 (1.94)	50.72***
Adolescents	(N=124; 72%)	(N=48; 28%)	(N=172)	
Algorithm A				
A1	3.98 (1.71)	0.17 (0.48)	2.91 (2.26)	22.61***
A2	6.45 (1.60)	0.73 (1.30)	4.85 (2.99)	24.21***
A3	4.60 (1.63)	0.58 (1.27)	3.48 (2.37)	17.17***
A4	6.04 (2.37)	0.48 (0.83)	4.49 (3.24)	22.83***
Algorithm B				
B1	4.49 (2.70)	0.13 (0.64)	3.27 (3.04)	16.83***
B2	3.33 (1.12)	0.27 (0.68)	2.48 (1.71)	21.87***
B3	2.49 (1.67)	0.06 (0.25)	1.81 (1.79)	15.81***
B4	5.06 (1.37)	0.69 (1.17)	3.84 (2.36)	19.51***
Algorithm C				
C1	2.31 (1.26)	0.19 (0.61)	1.72 (1.47)	14.83***
C2	1.19 (1.23)	0.04 (0.20)	0.87 (1.17)	10.00***
C3	0.45 (0.80)	0.00 (0.00)	0.33 (0.71)	6.29***
C4	1.24 (0.81)	0.08 (0.28)	0.92 (0.88)	13.93***
Algorithm D				
D	3.15 (1.24)	0.13 (0.39)	2.30 (1.73)	24.16***
Adults	(N=39; 77%)	(N=12; 24%)	(N=51)	
Algorithm A				
A1	3.97 (2.07)	0.08 (0.29)	3.06 (2.46)	11.38***
A2	6.51 (1.45)	0.50 (1.24)	5.10 (2.93)	12.96***
A3	5.05 (1.41)	0.33 (0.65)	3.94 (2.39)	16.04***
A4	6.08 (2.25)	0.42 (0.79)	4.75 (3.14)	13.25***

Table 2. Mean scores of the K-ADI-R diagnostic algorithm per domain by age group (continued)

	ASD	Non-ASD	Total	t†
	Mean (SD)	Mean (SD)	Mean (SD)	
Algorithm B				
B1	5.23 (2.55)	0.08 (0.29)	4.02 (3.13)	12.36***
B2	3.56 (0.64)	0.33 (0.65)	2.80 (1.52)	15.22***
B3	2.67 (1.87)	0.00 (0.00)	2.04 (1.99)	8.91***
B4	5.15 (0.96)	0.58 (0.90)	4.08 (2.17)	14.61***
Algorithm C				
C1	2.26 (0.94)	0.33 (0.65)	1.80 (1.20)	6.61***
C2	1.82 (1.49)	0.00 (0.00)	1.39 (1.51)	7.66***
C3	0.59 (0.82)	0.00 (0.00)	0.45 (0.76)	4.50***
C4	1.21 (0.92)	0.25 (0.45)	0.98 (0.93)	4.84***
Algorithm D				
D	3.62 (1.50)	0.25 (0.62)	2.82 (1.97)	11.24***

***p<0.001, †t-test. ASD: autism spectrum disorder, Algorithm A: Social Interaction, A1: Failure to use nonverbal behaviors to regulate social interaction, A2: Failure to develop peer relationships, A3: Lack of shared enjoyment, A4: Lack of socioemotional reciprocity, Algorithm B: Communication, B1: Lack of or delay in spoken language and failure to compensate through gesture, B2: Relative failure to initiate or sustain conversational interchange, B3: Stereotyped, repetitive, or idiosyncratic speech, B4: Lack of varied spontaneous make-believe or social imitative play, Algorithm C: Restricted, Repetitive, and Stereotyped Patterns of Behavior, C1: Encompassing preoccupation or circumscribed pattern of interest, C2: Apparently compulsive adherence to nonfunctional routines or rituals, C3: Stereotyped and repetitive motor mannerism, C4: Preoccupation with parts of objects or nonfunctional elements of material, Algorithm D: Abnormality of development before 36 months

uals diagnosed with ASD, 448 were classified as autistic disorder, 291 as Asperger, and 86 as pervasive developmental disorder not otherwise specified. The participants' average age was 88.9 months (SD=62.4), with the ASD group being around 92.2 months (SD=64.1). The FSIQ of the ASD group was 82.5 (SD=26.6), which was significantly lower than that of the non-ASD group (105.3, SD=18.9) (p<0.001). The mean K-ADI-R algorithm score, as indicated in Table 1, shows statistically significant group differences between the ASD and non-ASD groups across all domains (p<0.001). The t-test results show that the percentage of individuals with language delay, language regression, or loss of skills, measured based on the K-ADI-R items, were significantly higher in the ASD group (p<0.001).

Analysis of the data resulted in significantly higher mean subdomain scores in the K-ADI-R for the ASD group to the non-ASD group (Table 2) (p<0.001). The stratified analysis to further investigate the potential effects of sex differences by dividing males and females showed that the ASD group, regardless of sex, were significantly higher than those of the non-

Table 3. Validity between the K-ADI-R and the clinical best estimate diagnosis

	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	Cohen's Kappa (p-value)
ASD (N=825) vs. non-ASD (N=446)					
ADI-R†	99.27 (98.42−99.73)	84.75 (81.08−87.96)	92.33 (90.63−93.75)	98.44 (96.59−99.29)	0.868 (<0.001)
ADI-R‡	86.06 (83.51−88.35)	99.55 (98.39−99.95)	99.72 (98.89−99.93)	79.43 (76.52−82.06)	0.809 (<0.001)
ASD (N=825) vs. OD (N=54)					
ADI-R†	99.27 (98.42−99.73)	59.26 (45.03−72.43)	97.38 (96.43−98.09)	84.21 (69.99−92.42)	0.679 (<0.001)
ADI-R‡	86.06 (83.51−88.35)	98.15 (90.11−99.95)	99.86 (99.03−99.98)	31.55 (27.93−35.41)	0.424 (<0.001)

†K-ADI-R ASD group defined as participants who equaled or exceeded the cut-off score in at least one domain, ‡K-ADI-R ASD group defined as participants who equaled or exceeded the cut-off score in all four domains. ADI-R: Autism Diagnostic Interview-Revised, PPV: positive predictive value, NPV: negative predictive value

ASD group (Supplementary Table 1 in the online-only Data Supplement). Similarly, despite being broken down by age and verbal ability, the results showed significant differences with consistently higher scores in the ASD group (Supplementary Tables 2 and 3 in the online-only Data Supplement).

When the ASD group was defined as participants who equaled or exceeded the cut-off score in at least one domain of the K-ADI-R, it resulted in high sensitivity (99.27%), specificity (84.75%), PPV (92.33%) and, NPV (98.64%) across all age groups regardless of expressive verbal abilities (Table 3). Limiting the ASD group to those who met all four domains of the K-ADI-R also demonstrated high sensitivity (86.06%), specificity (99.55%), PPV (99.72%), and NPV (79.43%). Similar patterns were seen when analyzed by age groups.

Based on the pre-defined standards, 54 participants were classified into the OD group (Supplementary Table 4 in the online-only Data Supplement). Attempts to explore the validity in differentiating ASD and OD by using the K-ADI-R resulted in ranges of 86.06−99.27% sensitivity, 59.26−98.15% specificity, 97.38−99.86% PPV, and 31.55−84.21% NPV, depending on the number of domains used to classify individuals into the ASD group (Table 3). The small number of participants with OD restricted further analysis by age groups.

As seen in Table 4, Cohen's kappa for the agreement between the K-ADI-R and other ASD related measurements such as the K-ADOS, K-CARS, SCQ, and SRS demonstrated moderate to excellent levels. Grouping ASD participants as those who met at least one domain of the K-ADI-R had k values ranging from 0.429 to 0.947, whereas participants clustered as ASD when satisfying all four domains had k values ranging from 0.481 to 1.00.

DISCUSSION

This study examines the diagnostic validity of the K-ADI-R, which differentiated ASD and non-ASD participants by showing significant differences in the algorithm's scores between

Table 4. Agreement with existing instruments by age group

	N	Cohen's Kappa (p-value)	
		ADI-R† (lower cut-off)	ADI-R‡ (higher cut-off)
Children			
ADOS (lower cut-off)	1,047	0.833 (<0.001)	0.773 (<0.001)
ADOS (higher cut-off)	1,047	0.652 (<0.001)	0.718 (<0.001)
CARS	413	0.429 (<0.001)	0.578 (<0.001)
SCQ	839	0.578 (<0.001)	0.634 (<0.001)
SRS	710	0.605 (<0.001)	0.683 (<0.001)
Adolescents			
ADOS (lower cut-off)	172	0.866 (<0.001)	0.784 (<0.001)
ADOS (higher cut-off)	172	0.633 (<0.001)	0.682 (<0.001)
CARS	19	N/a	1.000 (<0.001)
SCQ	161	0.588 (<0.001)	0.613 (<0.001)
SRS	161	0.675 (<0.001)	0.700 (<0.001)
Adults			
ADOS (lower cut-off)	51	0.947 (<0.001)	0.760 (<0.001)
ADOS (higher cut-off)	51	0.683 (<0.001)	0.523 (<0.001)
CARS	4	N/a	N/a
SCQ	48	0.670 (<0.001)	0.697 (<0.001)
SRS	49	0.557 (<0.001)	0.481 (<0.001)

†K-ADI-R ASD group defined as participants who equaled or exceeded the cut-off score in at least one domain, ‡K-ADI-R ASD group defined as participants who equaled or exceeded the cut-off score in all four domains. ADOS: Autism Diagnostic Observation Schedule, CARS: Childhood Autism Rating Scale, SCQ: Social Communication Questionnaire, SRS: Social Responsiveness Scale

the two groups. This study is particularly meaningful because the analyses were based on a large sample size with wide age ranges that covered individuals from 24 months to 34-year-olds. Additionally, the analyses compared participants who met the K-ADI-R diagnostic algorithm thresholds on at least one domain versus all four domains. Results reported high sensitivity, specificity, PPV, and NPV in all age groups and dem-

onstrated good to excellent instrument agreement levels with ADOS, CARS, SCQ, and SRS. Incorporating the results of the K-ADI-R in making the best estimate clinical diagnosis could have impacted our findings. To complement this issue and by following previous studies demonstrating increased diagnostic accuracy when combining multiple sources of information,[49] our best estimate clinical diagnosis was based upon the combination of direct observations, caregiver questionnaires and interviews, as well as other psychological assessments.

This study found indications of excellent sensitivity (99.27%) and specificity (84.75%) of the K-ADI-R. The overall PPV and NPV were 92.33 and 98.44%, respectively, indicating excellent clinical utility. When ASD groups were limited to those who satisfied all four domains of the K-ADI-R, they also showed high sensitivity (86.06%), specificity (99.55%), PPV (99.72%), and NPV (79.43%). In addition to sensitivities, specificity, PPV, and NPV, this study also compared algorithm scores for each domain and showed significant differences in all categories, making this study more meaningful. Overall, these results suggest that the K-ADI-R is an effective tool with good to excellent validity in diagnosing ASD. Even when limited to comparisons with the OD group, the K-ADI-R demonstrated high validity in detecting individuals with ASD.

Previous studies investigating the diagnostic validity of ADI-R among various samples with different languages and cultures have reported good to excellent sensitivity for the diagnosis of ASD.[18,22,24-26,50] Lord et al.[8] conducted a validation study on the original English version of ADI-R and found high sensitivity (96%) and specificity (92%). Compared to the initial validation of the ADI-R, the current study showed higher sensitivity but lower specificity. Findings have been inconsistent with studies using other languages reporting different results with varying ranges. For example, while the study with Brazilian subjects showed perfect sensitivity and specificity,[22,24] the study with a Japanese sample of 317 participants aged 2 to 19 years, found the sensitivity and specificity to be 92 and 89%, and PPV and NPV to be 89 and 90%, respectively.[25] The ADI-R sensitivity and specificity in Greek samples were somewhat lower than other studies (85 and 75%, respectively). The Greek and Brazilian studies had small sample sizes (77 and 40 individuals, respectively), and the use of a homogeneous control group could have accounted for the differences in validity. The participants' chronological and mental age could have also affected the results with studies consisting of younger subjects reporting lower specificity and sensitivity.[22,51] One study looking at participants who ranged from 16 to 31 months showed poor sensitivity (56%) and moderate specificity (67%).[51] However, the study in Brazil, which showed 100% sensitivity and specificity, was conducted on participants aged 7 to 18 years old.[24] The systematic review supports the findings that ADI-R

sensitivity for children under 3 years of age (82%) is much lower than that for children over 3 years of age (91%).[32] The differences in validity measures can be attributed to varying sample sizes, the severity of symptoms, characteristics of comparison groups, as well as linguistic and cultural factors.[18]

Cultural factors may shape the perception of ASD symptoms due to differences in determining risk signs, which may stem from the varying values and standards in what is deemed appropriate behaviors.[52] For instance, Bong et al.[53] found potential influences of cultural differences to lower sensitivity in social smiles among Korean parents. Question items may also need to be worded in a certain way, with examples applicable to the communities being assessed.[54] For example, white parents tended to report higher levels of communication symptoms than Latina parents.[55] In other studies looking at cross cultural comparisons, a study showed that Latina parents underreported social interactions, another study found that Latina parents reported lower levels of restricted, repetitive and stereotyped behaviors than white parents.[56,57] As our study was primarily focused on the validity of the K-ADI-R in detecting individuals with ASD, we did not focus on specific question items but urge future studies to investigate whether patterns of cultural differences affect the validity of the assessment when used in non-Western populations.

Several interviews, questionnaires, and observational instruments have been developed and used to support clinicians in assessing the diagnosis of specific behaviors found in individuals with ASD. ADI-R, ADOS, CARS, SCQ, and SRS are typical examples, and past studies have analyzed the diagnostic agreement between these assessments. According to a study conducted with 119 children aged 9 to 13 years, the SCQ and SRS were highly correlated with ADI-R, while other studies showed that agreement between ADI-R and CARS reached 85.7%.[58,59] According to Cohen's kappa (0.432−0.842), this study also showed moderate to excellent diagnostic consistency between the ADI-R and other diagnostic tools. The CARS was an exception, showing the lowest level of agreement with ADI-R (k=0.432). This finding might be due to the psychometric characteristics of the K-CARS. The cut-off scores of the K-CARS used in the study were 28, based upon standardization and validation study results for the Korean version of CARS conducted in 1998.[37] As there have been high levels of false negatives (26.2%), especially among high-functioning subjects, and with the emphasis on sensory sensitivity, the cut-off scores were recently re-adjusted.[60] The SRS and the SCQ are questionnaires reported by parents, caregivers, and teachers, and how they interpret the questions may differ from the ADI-R, which is scored by trained clinicians.[7] Depending on caregivers and teachers, they may tend to reduce or enlarge a child's behaviors, or they may not even be aware of the symptoms or risk

signs of ASD. Also, the answer to the question items may vary depending on which aspect they focus on. Therefore, it may be beneficial to reach a diagnosis using multiple tools rather than relying on one source.

Although the current study has many positive findings, this study has several limitations. First, this study has not demonstrated how ASD severity and functioning levels are reflected in the ADI-R. Second, we had a relatively small sample of adult participants which restricted further analysis by age groups. Third, in addition to validating the ADI-R, the study tried to identify the difference between individuals with TD, OD, and ASD. However, missing part of the information used to classify individuals into the OD group, such as the FSIQ, VABS, and SMS, limited the classification of non-ASD groups. Further studies are encouraged to confirm differences between TD, OD, and ASD.

To our knowledge, this study used the largest sample size in Korea to examine the validity of the K-ADI-R and showed high sensitivity, specificity, PPV, and NPV within a wide range of age and demonstrated excellent agreement with other ASD diagnostic instruments. This study suggests that ADI-R might be a valuable diagnostic instrument for individuals with ASD across countries with different languages, cultural backgrounds, and levels of awareness for ASD.

Supplementary Materials

The online-only Data Supplement is available with this article at https://doi.org/10.30773/pi.2020.0337.

Acknowledgments

Conflicts of Interest

The authors have no potential conflicts of interest to disclose.

Author Contributions

Conceptualization: Miae Oh, Guiyoung Bong, Hee Jeong Yoo. Data curation: Guiyoung Bong, Joo-Hyun Kim. Formal analysis: Miae Oh, Da-Yea Song, Nan-He Yoon. Investigation: Da-Yea Song. Methodology: Guiyoung Bong. Project administration: Miae Oh, Guiyoung Bong. Resources: Guiyoung Bong, Jongmyeong Kim. Supervision: Hee Jeong Yoo. Validation: Miae Oh, Da-Yea Song, So Yoon Kim. Visualization: Da-Yea Song. Writing—original draft: Miae Oh, Da-Yea Song, Hee Jeong Yoo. Writing—review & editing: Miae Oh, Da-Yea Song, Hee Jeong Yoo.

ORCID iDs

Miae Oh   https://orcid.org/0000-0002-1927-3637
Da-Yea Song   https://orcid.org/0000-0002-7144-4739
Guiyoung Bong   https://orcid.org/0000-0001-8630-9399
Nan-He Yoon   https://orcid.org/0000-0003-1268-0149
So Yoon Kim   https://orcid.org/0000-0003-1349-8031
Joo-Hyun Kim   https://orcid.org/0000-0002-9000-0092
Jongmyeong Kim   https://orcid.org/0000-0002-6917-9358
Hee Jeong Yoo   https://orcid.org/0000-0003-0521-2718

REFERENCES

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). Washington DC: American Psychiatric Pub; 2013.
2. Elsabbagh M, Divan G, Koh YJ, Kim YS, Kauchali S, Marcin C, et al. Global prevalence of autism and other pervasive developmental disorders. Autism Res 2012;5:160-179.
3. Ardhanareeswaran K, Volkmar F. Introduction. Focus: autism spectrum disorders. Yale J Biol Med 2015;88:3-4.
4. Dawson G. Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. Dev Psychopathol 2008;20:775-803.
5. Reichow B. Overview of meta-analyses on early intensive behavioral intervention for young children with autism spectrum disorders. J Autism Dev Disord 2012;42:512-520.
6. Fein D, Barton M, Eigsti IM, Kelley E, Naigles L, Schultz RT, et al. Optimal outcome in individuals with a history of autism. J Child Psychol Psychiatry 2013;54:195-205.
7. Volkmar FR, Lord C, Bailey A, Schultz RT, Klin A. Autism and pervasive developmental disorders. J Child Psychol Psychiatry 2004;45:135-170.
8. Lord C, Rutter M, Le Couteur A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. J Autism Dev Disord 1994;24:659-685.
9. Rutter M, Le Couteur A, Lord C. Autism Diagnostic Interview-Revised. Los Angeles, CA: Western Psychological Services; 2003.
10. Lord C, Rutter M, DiLavore PC, Risi S, Gotham K, Bishop SL. Autism Diagnostic Observation Schedule. Manual. Los Angeles: Western Psychological Services; 1999.
11. Lord C, Rutter M, DiLavore P, Risi S, Gotham K, Bishop S. Autism Diagnostic Observation Schedule, (ADOS-2) Modules 1-4. Los Angeles, California: Western Psychological Services; 2012.
12. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. Washington DC: American Psychiatric Publishing; 2013.
13. Zivetz L. The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines. Geneva: World Health Organization; 1992
14. Le Couteur A, Haden G, Hammal D, McConachie H. Diagnosing autism spectrum disorders in pre-school children using two standardised assessment instruments: the ADI-R and the ADOS. J Autism Dev Disord 2008;38:362-372.
15. De Bildt A, Sytema S, Ketelaars C, Kraijer D, Mulder E, Volkmar F, et al. Interrelationship between autism diagnostic observation schedule-generic (ADOS-G), autism diagnostic interview-revised (ADI-R), and the diagnostic and statistical manual of mental disorders (DSM-IV-TR) classification in children and adolescents with mental retardation. J Autism Dev Disord 2004;34:129-137.
16. Gray KM, Tonge BJ, Sweeney DJ. Using the Autism Diagnostic Interview-Revised and the Autism Diagnostic Observation Schedule with young children with developmental delay: evaluating diagnostic validity. J Autism Dev Disord 2008;38:657-667.
17. Al Maskari TS, Melville CA, Willis DS. Systematic review: cultural adaptation and feasibility of screening for autism in non-English speaking countries. Int J Ment Health Syst 2018;12:22.
18. Vanegas SB, Magana S, Morales M, McNamara E. Clinical Validity of the ADI-R in a US-Based Latino Population. J Autism Dev Disord 2016; 46:1623-1635.
19. Matson JL, Worley JA, Fodstad JC, Chung KM, Suh D, Jhin HK, et al. A multinational study examining the cross cultural differences in reported symptoms of autism spectrum disorders: Israel, South Korea, the United Kingdom, and the United States of America. Res Autism Spectr Disord 2011;5:1598-1604.
20. Matson JL, Matheis M, Burns CO, Esposito G, Venuti P, Pisula E, et al. Examining cross-cultural differences in autism spectrum disorder: a multinational comparison from Greece, Italy, Japan, Poland, and the

United States. Eur Psychiatry 2017;42:70-76.

21. Poustka F, Lisch S, Ruhl D, Sacher A, Schmotzer G, Werner K. The standardized diagnosis of autism, Autism Diagnostic Interview-Revised: interrater reliability of the German form of the interview. Psychopathology 1996;29:145-153.

22. Papanikolaou K, Paliokosta E, Houliaras G, Vgenopoulou S, Giouroukou E, Pehlivanidis A, et al. Using the Autism Diagnostic Interview-Revised and the Autism Diagnostic Observation Schedule-Generic for the diagnosis of autism spectrum disorders in a Greek sample with a wide range of intellectual abilities. J Autism Dev Disord 2009;39:414-420.

23. Lampi KM, Sourander A, Gissler M, Niemela S, Rehnstrom K, Pulkkinen E, et al. Brief report: validity of Finnish registry-based diagnoses of autism with the ADI-R. Acta Paediatr 2010;99:1425-1428.

24. Becker MM, Wagner MB, Bosa CA, Schmidt C, Longo D, Papaleo C, et al. Translation and validation of Autism Diagnostic Interview-Revised (ADI-R) for autism diagnosis in Brazil. Arq Neuropsiquiatr 2012;70:185-190.

25. Tsuchiya KJ, Matsumoto K, Yagi A, Inada N, Kuroda M, Inokuchi E, et al. Reliability and validity of autism diagnostic interview-revised, Japanese version. J Autism Dev Disord 2013;43:643-662.

26. de Bildt A, Oosterling IJ, van Lang ND, Kuijper S, Dekker V, Sytema S, et al. How to use the ADI-R for classifying autism spectrum disorders? Psychometric properties of criteria from the literature in 1,204 Dutch children. J Autism Dev Disord 2013;43:2280-2294.

27. Sun X, Allison C, Auyeung B, Zhang ZX, Matthews FE, Baron-Cohen S, et al. Validation of existing diagnosis of autism in mainland China using standardised diagnostic instruments. Autism 2015;19:1010-1017.

28. Zander E, Sturm H, Bölte S. 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 2015;19:187-199.

29. Kim SH, Kim YS, Koh YJ, Lim EC, Kim SJ, Leventhal BL. Often asked but rarely answered: can Asians meet DSM-5/ICD-10 autism spectrum disorder criteria? J Child Adolesc Psychopharmacol 2016;26: 835-842.

30. Chojnicka I, Pisula E. Cross-cultural validation of the Polish version of the ADI-R, including new algorithms for toddlers and young preschoolers. Child Psychiatry Hum Dev 2019;50:591-604.

31. Jones RM, Risi S, Wexler D, Anderson D, Corsello C, Pickles A, et al. How interview questions are placed in time influences caregiver description of social communication symptoms on the ADI-R. J Child Psychol Psychiatry 2015;56:577-585.

32. Falkmer T, Anderson K, Falkmer M, Horlin C. Diagnostic procedures in autism spectrum disorders: a systematic literature review. Eur Child Adolesc Psychiatry 2013;22:329-340.

33. Yoo H, Park K, Kwak Y, Cho S, Ban G, Kim B. Korean Version of Autism Diagnostic Interview-Revised (ADI-R). Seoul: Inpsyt; 2007.

34. Yoo H, Kwak Y. Korean Version of Autism Diagnostic Observation Schedule (ADOS). Seoul: Hakjisa; 2007.

35. Yoo HJ, Bong GY, Kwak YS, Lee MS, Cho SH, Kim BN, et al. Korean Autism Diagnostic Observation Schedule-2 (K-ADOS-2). Seoul: Hakjisa; 2018.

36. Kim T, Park R. Korean Version of Childhood Autism Rating Scale. Seoul: Special Education; 1995.

37. Shin M, Kim Y. Standardization study for the Korean Version of Childhood Autism Rating Scale: reliability, validity and cut-off score. Korean J Clin Psychol 1998;17:1-15.

38. Kim S, Kim O. Korean Vineland Social Maturity Scale. Seoul: Chung-Ang Jeokseong Press; 2002.

39. Doll EA. The Measurement of Social Competence: A Manual for the Vineland Social Maturity Scale. Minneapolis: Educational Test Bureau, Educational Publishers; 1953.

40. Kim S, Kim Y. Korean Version of the Vineland Social Maturity Scale. Seoul: Chung Ang Jeok Seong Press; 1985.

41. Constantino JN, Gruber CP. Social Responsiveness Scale: SRS-2. Torrance, CA: Western Psychological Services; 2012.

42. Cholemkery H, Kitzerow J, Rohrmann S, Freitag CM. Validity of the social responsiveness scale to differentiate between autism spectrum disorders and disruptive behaviour disorders. Eur Child Adolesc Psychiatry 2014;23:81-93.

43. Wang J, Lee LC, Chen YS, Hsu JW. Assessing autistic traits in a Taiwan preschool population: cross-cultural validation of the Social Responsiveness Scale (SRS). J Autism Dev Disord 2012;42:2450-2459.

44. Rutter M, Bailey A, Lord C. The Social Communication Questionnaire: Manual. Los Angeles, CA: Western Psychological Services; 2003.

45. Wiggins LD, Reynolds A, Rice CE, Moody EJ, Bernal P, Blaskey L, et al. Using standardized diagnostic instruments to classify children with autism in the study to explore early development. J Autism Dev Disord 2015;45:1271-1280.

46. Kim JH, Sunwoo HJ, Park SB, Noh DH, Jung YK, Cho IH, et al. A validation study of the Korean version of Social Communication Questionnaire. J Korean Acad Child Adolesc Psychiatry 2015;26:197-208.

47. Volkmar FR. Vineland Adaptive Behavior Scales, Second Edition, Encyclopedia of Autism Spectrum Disorders. New York: Springer; 2013.

48. Hwang ST, Kim JH, Hong SH. Korean Vineland Adaptive Behavior Scales-II. Daegu: Korea Psychology Co.; 2014.

49. Risi S, Lord C, Gotham K, Corsello C, Chrysler C, Szatmari P et al. Combining information from multiple sources in the diagnosis of autism spectrum disorders. Am Acad Child Adolesc Psychiatry 2006;45: 1094-1103.

50. de Bildt A, Sytema S, Zander E, Bolte S, Sturm H, Yirmiya N, et al. Autism iagnostic Interview-Revised (ADI-R) algorithms for toddlers and young preschoolers: application in a non-US sample of 1,104 children. J Autism Dev Disord 2015;45:2076-2091.

51. Ventola PE, Kleinman J, Pandey J, Barton M, Allen S, Green J, et al. Agreement among four diagnostic instruments for autism spectrum disorders in toddlers. J Autism Dev Disord 2006;36:839-847.

52. Matson JL, Matheis M, Burns CO, Esposito G, Venuti P, Pisula E, et al. Examining cross-cultural differences in autism spectrum disorder: a multinational comparison from Greece, Italy, Japan, Poland, and the United States. Eur Psychiatry, 2017;42:70-76.

53. Bong G, Kim JH, Hong Y, Yoon NH, Sunwoo H, Jang JY, et al. The feasibility and validity of autism spectrum disorder screening instrument: Behavior Development Screening for Toddlers (BeDevel)—a pilot study. Autism Res 2019;12:1112-1128.

54. Lee KS, Jung SJ, Cho SW, Shin YJ, Jeon MK, Ahn SI. Application of SACS-K to Korean infants and toddlers with autism spectrum disorders (ASD): comparison of markers between ASD, developmental delay and typical children. J Korean Assoc Pers Autism 2016;16:1-26.

55. Blacher J, Cohen SR, Azad G. In the eye of the beholder: reports of autism symptoms by Anglo and Latino mothers. Res Autism Spectr Disord 2014;8:1648-1656.

56. Overton T, Fielding C, de Alba RG. Differential diagnosis of Hispanic children referred for autism spectrum disorders: Complex issues. J Autism Dev Disord 2007;37:1996-2007.

57. Magaña S, Smith MJ. Psychological distress and well–being of Latina and non–Latina white mothers of youth and adults with an autism spectrum disorder: Cultural attitudes towards coresidence status. Am J Orthopsychiatry 2006;76:346-357.

58. Charman T, Baird G, Simonoff E, Loucas T, Chandler S, Meldrum D, et al. Efficacy of three screening instruments in the identification of autistic-spectrum disorders. Br J Psychiatry 2007;191:554-559.

59. Pilowsky T, Yirmiya N, Shulman C, Dover R. The Autism Diagnostic Interview-Revised and the Childhood Autism Rating Scale: differences between diagnostic systems and comparison between genders. J Autism Dev Disord 1998;28:143-151.

60. Mayes SD, Calhoun SL, Murray MJ, Morrow JD, Yurich KK, Mahr F, et al. Comparison of scores on the Checklist for Autism Spectrum Disorder, Childhood Autism Rating Scale, and Gilliam Asperger's Disorder Scale for children with low functioning autism, high functioning autism, Asperger's disorder, ADHD, and typical development. J Autism Dev Disord 2009;39:1682-1693.

Supplementary Table 1. Domain scores of the K-ADI-R diagnostic algorithm by sex

	ASD	Non-ASD	Total	$t^\dagger$
	Mean (SD)	Mean (SD)	Mean (SD)	
Male	(N=698; 75%)	(N=228; 25%)	(N=926)	
Algorithm A				
A1	3.74 (1.63)	0.32 (0.64)	2.90 (2.07)	45.67***
A2	5.68 (2.36)	1.04 (1.35)	4.53 (2.94)	36.68***
A3	4.52 (1.56)	1.01 (1.16)	3.66 (2.11)	36.29***
A4	6.25 (2.25)	0.94 (1.18)	4.94 (3.06)	46.12***
Algorithm B				
B1	5.27 (2.48)	0.54 (1.00)	4.10 (3.01)	41.23***
B2	2.17 (1.83)	0.34 (0.82)	1.72 (1.82)	20.75***
B3	1.86 (1.97)	0.29 (0.65)	1.48 (1.87)	18.24***
B4	4.79 (1.39)	1.14 (1.17)	3.89 (2.07)	39.03***
Algorithm C				
C1	2.04 (1.26)	0.32 (0.66)	1.62 (1.36)	26.65***
C2	1.11 (1.22)	0.18 (0.50)	0.88 (1.16)	16.53***
C3	0.80 (0.91)	0.04 (0.23)	0.61 (0.86)	20.43***
C4	1.50 (0.70)	0.34 (0.57)	1.22 (0.84)	25.37***
Algorithm D				
D	3.68 (1.17)	0.51 (1.03)	2.90 (1.78)	39.01***
Female	(N=127; 37%)	(N=218; 63%)	(N=345)	
Algorithm A				
A1	3.91 (1.64)	0.28 (0.61)	1.62 (2.07)	24.02***
A2	5.24 (2.32)	0.80 (1.14)	2.43 (2.72)	20.15***
A3	4.32 (1.57)	0.48 (0.97)	1.90 (2.22)	24.98***
A4	6.15 (2.34)	0.66 (1.06)	2.68 (3.12)	24.99***
Algorithm B				
B1	5.27 (2.34)	0.36 (1.02)	2.17 (2.88)	22.41***
B2	1.76 (1.81)	0.19 (0.57)	0.77 (1.41)	9.50***
B3	1.51 (1.87)	0.16 (0.50)	0.66 (1.37)	8.01***
B4	4.50 (1.48)	0.66 (1.13)	2.07 (2.25)	25.30***
Algorithm C				
C1	1.39 (1.24)	0.16 (0.52)	0.61 (1.04)	10.69***
C2	1.06 (1.26)	0.09 (0.36)	0.45 (0.94)	8.54***
C3	0.79 (0.90)	0.05 (0.24)	0.32 (0.68)	9.08***
C4	1.26 (0.78)	0.16 (0.41)	0.56 (0.78)	14.82***
Algorithm D				
D	3.46 (1.38)	0.23 (0.69)	1.42 (1.85)	24.52***

***p<0.001, †t-test. ASD: Autism spectrum disorder, Algorithm A: Social Interaction, A1: Failure to use nonverbal behaviors to regulate social interaction, A2: Failure to develop peer relationships, A3: Lack of shared enjoyment, A4: Lack of socioemotional reciprocity, Algorithm B: Communication, B1: Lack of or delay in spoken language and failure to compensate through gesture, B2: Relative failure to initiate or sustain conversational interchange, B3: Stereotyped, repetitive, or idiosyncratic speech, B4: Lack of varied spontaneous make-believe or social imitative play, Algorithm C: Restricted, Repetitive, and Stereotyped Patterns of Behavior, C1: Encompassing preoccupation or circumscribed pattern of interest, C2: Apparently compulsive adherence to nonfunctional routines or rituals, C3: Stereotyped and repetitive motor mannerism, C4: Preoccupation with parts of objects or nonfunctional elements of material, Algorithm D: Abnormality of development before 36 months

Supplementary Table 2. Domain scores of the K-ADI-R verbal algorithm by age

	ASD	Non-ASD	Total	t[†]
	Mean (SD)	Mean (SD)	Mean (SD)	
Below 48 months (verbal)	(N=56; 43%)	(N=75; 57%)	(N=131)	
Algorithm A	13.00 (4.13)	3.29 (3.21)	7.48 (6.03)	15.20***
A1	2.61 (1.24)	0.47 (0.83)	1.39 (1.48)	11.33***
A2	2.63 (1.21)	0.95 (1.06)	1.67 (1.40)	8.51***
A3	3.26 (1.32)	0.84 (0.95)	1.89 (1.64)	11.78***
A4	4.49 (1.82)	1.04 (1.27)	2.53 (2.30)	12.22***
Algorithm B	12.35 (4.49)	2.19 (2.44)	6.58 (6.13)	15.45***
B1	4.00 (2.25)	0.51 (1.03)	2.02 (2.41)	10.88***
B2	2.39 (1.33)	0.36 (0.78)	1.23 (1.46)	10.22***
B3	2.37 (1.50)	0.43 (0.81)	1.27 (1.50)	8.87***
B4	3.60 (1.33)	0.89 (1.02)	2.06 (1.78)	13.19***
Algorithm C	4.96 (2.80)	0.80 (1.08)	2.60 (2.88)	10.64***
C1	1.84 (1.28)	0.23 (0.56)	0.92 (1.23)	8.91***
C2	1.23 (1.28)	0.28 (0.63)	0.69 (1.07)	5.14***
C3	0.63 (0.79)	0.07 (0.30)	0.31 (0.63)	5.10***
C4	1.26 (0.79)	0.23 (0.45)	0.67 (0.81)	8.85***
Algorithm D	2.86 (1.08)	0.41 (0.88)	1.47 (1.55)	14.43***
48 months and over (verbal)	(N=450; 61%)	(N=283; 39%)	(N=733)	
Algorithm A	21.28 (5.88)	2.27 (3.04)	13.94 (10.52)	57.50***
A1	4.00 (1.67)	0.17 (0.48)	2.52 (2.30)	45.72***
A2	6.51 (1.68)	0.93 (1.36)	4.35 (3.14)	49.23***
A3	4.64 (1.58)	0.64 (1.09)	3.09 (2.41)	40.52***
A4	6.14 (2.21)	0.53 (0.92)	3.97 (3.29)	47.68***
Algorithm B	16.35 (4.71)	1.62 (2.28)	10.66 (8.19)	56.55***
B1	4.90 (2.58)	0.23 (0.72)	3.10 (3.08)	36.27***
B2	3.54 (0.85)	0.33 (0.79)	2.30 (1.77)	52.23***
B3	3.01 (1.73)	0.23 (0.59)	1.94 (1.95)	31.32***
B4	4.90 (1.39)	0.83 (1.16)	3.33 (2.37)	42.64***
Algorithm C	5.80 (2.72)	0.55 (1.05)	3.77 (3.39)	36.87***
C1	2.35 (1.21)	0.24 (0.62)	1.53 (1.45)	30.99***
C2	1.38 (1.35)	0.08 (0.35)	0.88 (1.25)	19.41***
C3	0.68 (0.88)	0.01 (0.08)	0.42 (0.77)	16.22***
C4	1.39 (0.77)	0.23 (0.49)	0.94 (0.88)	24.99***
Algorithm D	3.50 (1.24)	0.23 (0.73)	2.23 (1.92)	45.06***

***p<0.001, [†]t-test. ASD: Autism spectrum disorder, Algorithm A: Social Interaction, A1: Failure to use nonverbal behaviors to regulate social interaction, A2: Failure to develop peer relationships, A3: Lack of shared enjoyment, A4: Lack of socioemotional reciprocity, Algorithm B: Communication, B1: Lack of or delay in spoken language and failure to compensate through gesture, B2: Relative failure to initiate or sustain conversational interchange, B3: Stereotyped, repetitive, or idiosyncratic speech, B4: Lack of varied spontaneous make-believe or social imitative play, Algorithm C: Restricted, Repetitive, and Stereotyped Patterns of Behavior, C1: Encompassing preoccupation or circumscribed pattern of interest, C2: Apparently compulsive adherence to nonfunctional routines or rituals, C3: Stereotyped and repetitive motor mannerism, C4: Preoccupation with parts of objects or nonfunctional elements of material, Algorithm D: Abnormality of development before 36 months

Supplementary Table 3. Domain scores of the K-ADI-R non-verbal algorithm by age

	ASD	Non-ASD	Total	t†
	Mean (SD)	Mean (SD)	Mean (SD)	
Below 48 months (non-verbal)	(N=173; 66%)	(N=88; 34%)	(N=261)	
Algorithm A	15.55 (4.46)	3.99 (3.03)	11.64 (6.80)	24.65***
A1	3.15 (1.41)	0.58 (0.72)	2.28 (1.72)	19.37***
A2	2.78 (1.16)	0.89 (1.06)	2.14 (1.44)	12.84***
A3	3.91 (1.37)	1.03 (1.21)	2.93 (1.90)	16.61***
A4	5.72 (1.97)	1.49 (1.27)	4.29 (2.67)	20.93***
Algorithm B	9.69 (2.75)	2.27 (2.14)	7.18 (4.35)	23.93***
B1	5.38 (2.07)	1.11 (1.43)	3.93 (2.76)	19.40***
B4	4.27 (1.33)	1.14 (1.32)	3.21 (1.99)	18.05***
Algorithm C	4.09 (2.04)	0.91 (1.07)	3.02 (2.33)	16.50***
C1	1.26 (1.08)	0.26 (0.56)	0.92 (1.05)	9.81***
C2	0.62 (0.85)	0.17 (0.46)	0.47 (0.77)	5.54***
C3	0.72 (0.88)	0.14 (0.41)	0.52 (0.80)	7.27***
C4	1.50 (0.64)	0.34 (0.59)	1.11 (0.83)	14.60***
Algorithm D	3.78 (1.14)	0.81 (1.19)	2.77 (1.82)	19.60***

Non-verbal individuals over 48 months were not included as there were no non-ASD participants who met the criteria. ***p<0.001, †t-test. ASD: Autism spectrum disorder, Algorithm A: Social Interaction, A1: Failure to use nonverbal behaviors to regulate social interaction, A2: Failure to develop peer relationships, A3: Lack of shared enjoyment, A4: Lack of socioemotional reciprocity, Algorithm B: Communication, B1: Lack of or delay in spoken language and failure to compensate through gesture, B4: Lack of varied spontaneous make-believe or social imitative play, Algorithm C: Restricted, Repetitive, and Stereotyped Patterns of Behavior, C1: Encompassing preoccupation or circumscribed pattern of interest, C2: Apparently compulsive adherence to nonfunctional routines or rituals, C3: Stereotyped and repetitive motor manner-ism, C4: Preoccupation with parts of objects or nonfunctional elements of material, Algorithm D: Abnormality of development before 36 months

Supplementary Table 4. Participant characteristics between the ASD and OD groups

Characteristics	N	ASD (N=825; 94%)	OD (N=54; 6%)	Total (N=879)	t or χ^2
		Mean (SD)	Mean (SD)	Mean (SD)	
Male (%)[†]	879	84.6	64.8	83.4	14.33***
Age (months)[‡]	879	92.2 (64.1)	68.9 (45.8)	90.7 (63.3)	3.52***
FSIQ[‡]	454	82.5 (26.6)	83.9 (26.9)	82.6 (26.6)	-0.30
ADI-R score[‡]					
Algorithm A	879	19.9 (6.4)	4.7 (3.9)	19.0 (7.3)	26.58***
Algorithm B verbal	544	15.9 (4.8)	3.0 (3.1)	15.0 (5.8)	23.60***
Algorithm B nonverbal	335	10.8 (2.9)	3.3 (2.0)	10.4 (3.3)	14.43***
Algorithm B nonverbal only	879	10.0 (3.3)	2.1 (2.0)	9.5 (3.8)	26.56***
Algorithm C	879	5.3 (2.6)	1.1 (1.3)	5.1 (2.7)	21.42***
Algorithm D	879	3.6 (1.2)	1.1 (1.4)	3.5 (1.4)	14.72***
Level of functioning (%)[†]					
Language delay	879	48.7	11.1	46.4	28.83***
Regression of language	879	5.2	1.9	5.0	1.20
Loss of skills	879	3.6	0.0	3.4	2.03

***p<0.001, [†]chi-square test, [‡]t-test. ASD: Autism spectrum disorder, OD: other developmental disorder, FSIQ: full scale intelligence quotient, ADI-R: Autism Diagnostic Interview-Revised, Algorithm A: Social Interaction, Algorithm B: Communication, Algorithm C: Restricted, Repetitive, and Stereotyped Patterns of Behavior, Algorithm D: Abnormality of development before 36 months