Table 2. Past history of the ASD participants (N = 154)

Early developmental concerns	N (%)
Absent	29 (19.5)
Present	120 (80.5)
Age at first concern (median age)	30 months
Age at referral	48 months
Age at first diagnosis	123 months
Speech level at 6 years	
Words or two-word phrases	34 (24.8)
Sentences with more than three words	103 (75.2)
Early diagnoses before 4 years	
Diagnosed[a]	29 (18.8)
Service utilization[b]	
No utilization at any time	124 (80.5)
Continuous utilization through the entire life stages	22 (14.3)

[a]Includes diagnoses of ASD and other developmental disorders.
[b]Service refers to having professional advice regularly, participating in some therapeutic programs, receiving special educational aid.

Quality of life. Subjective QoL was measured using the WHOQOL-BREF, which is derived from the 100-item WHOQOL (The WHOQOL Group, 1995). The WHOQOL was developed to measure individuals' perceptions of their position in life in the context of the culture and value systems in which they live, and in relation to their goals, expectations, standards, and concerns, and is used widely and internationally. Its use has been validated for various populations, including psychiatric patients worldwide. Its shorter version, the WHOQOL-BREF, comprises 26 items on four domains of QoL: physical health, psychological health, social relationships, and environment. For the purpose of the present study, the six items of the 'psychological health' domain (bodily image and appearance, negative feelings, positive feelings, self-esteem, thinking, learning, and memory and concentration) and the three items of the 'social relationships' domain (personal relationships, social support, and sexual activity) of the Japanese version of the WHOQOL-BREF (WHOQOL 26) (Nakane et al., 1999; Tazaki and Nakane, 2007) were used. Each item is assessed by an individual diagnosed with ASD on a 5-point scale (1 = very poor/very dissatisfied/not at all; 2 = poor/dissatisfied/a little; 3 = neither poor nor good/a moderate amount; 4 = good/satisfied/very much; 5 = very good/very satisfied/extremely). The mean scores of these two domains were analysed.

Current family support. The family support situation was determined by responses to the question, 'Regarding the physical and psychological support provided by his/her family member, do you think it is actually helpful for him/her?' The facility staff who knew the person well answered using a 5-point rating scale (1 = very helpful; 2 = somewhat helpful; 3 = not helpful or unhelpful; 4 = not very helpful; 5 = not at all helpful) for the case of the father, mother, and sibling, respectively (Table 3).

Demographic characteristics. Demographic information was obtained through 17 items rated by parents, and included gender, age, residential status, marital status, education, employment, medical conditions, and challenging behaviors. In this study, we asked questions requiring yes or no answers regarding the presence or absence of self-injurious behaviors and aggressive behaviors. Self-injurious behaviors were defined as any kind of behaviors in which the ASD participants hurt themselves. Aggressive behaviors were defined as violent behaviors toward family members or

Table 3. Current family support to the ASD participants (N = 154)

	N (%)
Father	
Helpful	60 (39.0)
Not helpful	63 (40.9)
No father or unknown	31 (20.1)
Mother	
Helpful	119 (77.3)
Not helpful	15 (9.7)
No mother or unknown	20 (13.0)
Sibling	
Helpful	35 (22.7)
Not helpful	82 (53.3)
No sibling or unknown	37 (24.0)

other people, verbal aggression as statements such as 'Die' or 'I will kill you', and destructive behaviors as those causing serious material damage. The most important items are shown Table 1.

Past history. Developmental information was obtained through 19 items rated by the parents, and included age at parental concern, age at diagnosis, expressive language level at age 6, and service utilization. The most important items are shown in Table 2.

The self-rating part of the survey questionnaire was pilot tested in several clinical settings to confirm the ease of completion. It was confirmed that individuals with HFASD were able to understand and complete it satisfactorily, at levels similar to other psychiatric patients (Koyama et al., 2009).

Statistical analysis

First, to compare the QoL domain scores for our ASD participants with those for a healthy Japanese population obtained using stratified sampling methods (N = 828; 410 males; aged 20–49) (Nakane et al., 1999; Tazaki and Nakane, 2007), the raw domain scores were converted to z scores using the mean and standard deviation of the Japanese standardization sample by gender and by age group (20–29, 30–39, 40–49) (Tazaki and Nakane, 2007). To obtain the z scores of participants 18–19 years of age, we applied the mean and standard deviation for the age range of 20–29 years to their raw QoL domain scores. Second, Pearson correlations were calculated to assess associations between performance in the current environment and QoL in both the psychological health and social relationships domains. Third, using two sample *t*-tests, z scores of the QoL domain scores were compared between subgroups of the following demographic characteristics: gender (male vs. female), age (≤ 24, 25+), residential status (independent living vs. other), marital status (unmarried vs. other), education (≤high school vs. additional higher education), employment (unemployed vs. other), medical conditions and challenging behaviors (present vs. absent), sentence level at 6 years of age (present vs. absent), diagnosis before 4 years of age (present vs. absent), and service utilization (none vs. continuous). For current family support, the responses were classified into two categories of 'helpful' (1, 2) and 'not helpful' (3, 4, 5). Finally, a stepwise multiple regression analysis was used to identify the most important characteristics in predicting QoL domain scores. As independent

Table 4. Means and SD for raw scores on psychological and social domains of WHOQOL 26 rated by ASD participants themselves (N = 154), and z scores converted from the raw scores of the ASD participants (N = 154)

QOL domain[a]	Mean (SD)	*t*	*p*	95% confidential interval
Psychological health				
Raw score	2.78 (0.74)			
Z score	−.80 (1.24)	−8.0	.0001	−1.0 to −.60
Social relationships				
Raw score	2.71 (0.82)			
Z score	−.63 (1.25)	−6.2	.0001	−.83 to −.43

[a]Psychological domain contains six items (1–5) and social domain contains three items (1–5). The mean raw domain scores of the ASD participants were converted to z scores using the mean and standard deviation of the Japanese standardization sample by gender and by age range (20–29, 30–39, 40–49) (Tazaki and Nakane, 2007).

variables, characteristics that were found to be significant on *t*-tests and all the demographic characteristics were used. A p-value < .05 was considered to indicate statistical significance. Statistical analysis was performed using SPSS version 18.0 (SPSS Inc., Chicago, USA).

Results

Psychological and social aspects of QoL in the ASD participants

The raw score means of the 'psychological health' and 'social relationships' domains of the WHOQOL 26 as rated by the ASD participants were 2.78 and 2.71, respectively (Table 4), whereas those of the Japanese standardization sample aged 20–49 ranged from 3.26 to 3.32 for the psychological health domain and from 3.19 to 3.25 for the social relationships domain, varying slightly by gender and age (Tazaki and Nakane, 2007). The differences in mean z scores of the ASD participants from those of the Japanese standardization sample were –0.80 for the psychological domain and -0.63 for the social domain, indicating that psychological and social aspects of QoL of the ASD participants were significantly lower (worse) than those of the healthy Japanese population (p < .000 for both).

Associations between psychological and social QoL and performance in the current environment in the ASD participants

Pearson correlations revealed that the QoL scores in both psychological and social domains for the ASD participants were not significantly correlated (r = 0.06 and r = 0.01, respectively, n.s.) with everyday performance in the current environment. This may be interpreted to suggest that the psychological and social domains of QoL capture more of the subjective aspects of QoL, and not objective function or capacity.

Factors related to psychological and social QoL in the ASD participants

As shown in Table 5, *t*-tests revealed that higher QoL was significantly associated with being male (p < .05 for psychological domain, p < .01 for social domain), having received a diagnosis before 4 years of age (p < .05 for psychological domain), and mother's support being helpful (p < .001 for both psychological and social domains). Lower QoL was significantly associated with suffering

Table 5. Comparison between two ASD subgroups by demographic characteristics, medical conditions past history, and current family support (N = 154)

	Psychological health domain QoL			Social relationships domain QoL		
	Mean	score	*t*	Mean	score	*t*
Demographic characteristics						
Gender (male/female)	2.83	2.56	2.35*	2.73	2.60	2.68**
Age (18–24/25+)	2.76	2.79	0.34	2.80	2.64	0.77
Residential status (living with family/independent living)	2.78	2.73	0.49	2.74	2.27	1.73
Marital status (unmarried/married or partnered)	2.78	2.74	0.37	2.72	2.46	1.37
Education (≤ high school/further higher education)	2.69	2.90	−1.51	2.64	2.80	−1.14
Employment (unemployed/employed)	2.76	2.85	−0.48	2.69	2.80	−0.40
Medical conditions						
Comorbid psychiatric conditions (absent/present)	2.91	2.58	2.76**	2.85	2.43	3.14**
Comorbid physical conditions (absent/present)	2.76	2.89	−0.38	2.71	2.60	0.42
Self-injurious behaviors (absent/present)	2.81	2.50	1.68	2.74	2.52	1.30
Aggressive behaviors (absent/present)	2.85	2.44	2.58*	2.80	2.36	2.38*
Past history						
Sentence level at 6 years (word or two-word phrases/ sentence with more than three words)	3.05	2.71	2.32*	2.93	2.65	1.64
Early diagnoses before 4 years (not diagnosed/diagnosed)	2.71	3.04	−2.02*	2.69	2.86	−1.18
Service utilization (no use at any time/continuous use through all life stages)	2.77	3.04	−1.63	2.71	2.73	−0.13
Current family support						
Father (not helpful/helpful)	2.73	2.81	−0.73	2.66	2.85	−1.25
Mother (not helpful/helpful)	2.14	2.84	−4.01***	1.91	2.84	−4.50***
Sibling (not helpful/helpful)	2.71	2.72	0.00	2.61	2.83	−0.92

**p<.05, **p<.01, ***p<.001.

Table 6. Summary of a stepwise multiple regression analysis investigating the predictive variables of demographic characteristics, medical conditions, past history, and current family support on QoL 'psychological health' domain scores of the ASD participants (N = 154)

Variables entered	Standardized coefficients (β)	*t*	*p*-value
Early diagnosis before 4 years	0.22	2.22	.05
Mother's support being helpful	0.32	3.24	.01

Adjusted R^2=0.16. Excluded variables by a stepwise procedure were gender, age, residential status, marital status, education, employment, comorbid psychiatric conditions, aggressive behaviors, speech level at 6 years.

from comorbid psychiatric conditions (p < .01 for both psychological and social domains), behaving aggressively (p < .01 for psychological domain, p < .05 for social domain), and having spoken sentences at 6 years of age (p < .05 for psychological domain).

The results of multiple regression analysis are summarized in Tables 6 and 7. Mother's support being helpful emerged as significantly predictive of higher QoL for both the psychological and social domains (β = 0.32, p < .01; β = 0.32, p < .001, respectively). In addition, having received early diagnosis before 4 years of age was also significantly associated with higher psychological

Table 7. Summary of a stepwise multiple regression analysis investigating the predictive variables of demographic characteristics, medical conditions, past history, and current family support on QOL 'social relationships' domain scores of the ASD participants (N = 154)

Variables entered	Standardized coefficients (β)	*t*	*p*-value
Aggressive behaviors: absent	0.18	2.15	.05
Mother's support being helpful	0.32	3.69	.001

Adjusted R^2=0.14. Excluded variables by a stepwise procedure were gender, age, residential status, marital status, education, employment, comorbid psychiatric conditions.

QoL ($\beta = 0.22$, $p < .05$), and not having aggressive behaviors was significantly associated with higher social QoL ($\beta = 0.18$, $p < .05$).

Discussion

The present study investigated long-term outcomes for adults with HFASD living in the community in Japan, focusing on subjective aspects such as QoL, and also identified past and current environmental factors that had (pseudo) predictive value. Our major findings are the following.

First, as expected, the self-reported QoL in the psychosocial domain of our sample with HFASD over 18 years of age was found to be significantly lower than the gender- and age-matched healthy Japanese population. The QoL was not found to be related to parent-reported performance level, age, or conventionally used outcome indicators such as residential, marital, educational, and employment status. Although these conventional indicators are certainly important to consider as long-term outcomes, psychosocial QoL in our adults with HFASD appeared not to be related to them. Thus, our findings suggest that the QoL reported by adults with HFASD might be measuring an additional independent aspect that should be considered in judging long-term outcomes in populations with HFASD, which is in line with Renty and Roeyers (2006) and Ruble and Dalrymple (1996).

Second, receiving diagnosis before 4 years of age and mother's support that met current needs were determined to be factors associated with better psychological QoL for adults with HFASD. This finding supports our prediction and is partially consistent with Renty and Roeyers (2006) in that support variables had significant impact on long-term outcomes in HFASD. In Renty and Roeyers (2006), perceived informal support indicative of availability, but not received formal or informal support indicative of actual transfer of advice, aid, and affect, was found to have predictive value; both support characteristics were measured using validated scales. On the other hand, family support characteristics in the current study were not measured using such standard scales, and were instead judged by the facility staff who knew the person well and therefore knew to what degree the family support was actually helpful to the person. This is different from subjectively perceived availability or objectively measured actual transfer of family support. The question was intended to ask how family support met the participant's actual needs from an objective viewpoint. However, validation of this is required.

To our knowledge, the present study is the first to associate early diagnosis with better psychological QoL in adults with HFASD. Only 29 cases out of our sample (18.8%) were diagnosed before 4 years of age, and 22 cases among them used some services during childhood. On the other hand, parental concerns about development were reported for a majority of the sample (66.9%). Why parental concerns did not lead to early diagnosis may be explained by a lack of healthcare or educational professionals with accurate knowledge and wide experience with HFASD at that time in Japan. Moreover, socioeconomic status (SES) could be associated with age of diagnosis: according to birth cohort data from individuals with autism born in California between 1992 and 2001, children

of high SES parents were diagnosed earlier (Fountain et al., 2011). The role of SES in our Japanese participants with HFASD is unknown and this remains a topic for future study.

The finding that mother's support was the best predictor of psychosocial QoL of individuals with HFASD has to be interpreted with caution, because it suggests a bidirectional but not causal association. However, the obvious significance of mother's support but not father's in our study may be related to the Japanese socio-cultural environment in relation to child rearing. For example, Japanese mothers have traditionally been viewed as overprotective and overindulgent toward their children (Doi, 1973). Although it is not clear how such a cultural bias in childrearing practices influences familial attitudes toward a child with HFASD, future intervention must target parenting behavior and assist mothers with formal and informal social support after an early diagnosis of ASD.

Third, our results suggest that adults with aggressive behaviors might experience lower psychosocial QoL. This could be attributable to satisfactory social relationships being disrupted by violent behavior toward others or anger being reflected on others or self, although the causal relationship is not clear.

Study limitations. There are several methodological limitations in the present study. First, our sample (N = 154) who returned a complete set of self-, parent-, and facility staff-report questionnaires were not representative of all persons with HFASD, although the male:female ratio was 4:1, which is similar to the epidemiological data in Japan (Honda et al., 2005). In regard to the 48 individuals not included in the analyses because of incomplete data, although they were older and more educated than our sample of 154, we confirmed, based on the information that was available, that they did not differ from our sample in psychological and social QoL scores. Moreover, being male, absence of comorbid psychiatric conditions, and mother's support being helpful were significantly associated with higher QoLs, as in our sample. However, whether early diagnosis before 4 years or having aggressive behaviors was similarly predictive of QoL in the 48 individuals not included could not be confirmed. Second, diagnostic status and IQ level of our sample was based on reports by parents and facility staff and was not confirmed using standard procedures. Third, we chose to focus on the domains of 'social relationships' and 'psychological health' because we considered that they best reflected psychosocial QoL, although the 'physical health' and 'environment' domains also reflect psychosocial functioning in everyday life to some degree. We based this decision on the findings of previous studies. Health-related QoL studies on ASD found that individuals with ASD scored lower in most domains than healthy populations (Jennes-Coussens et al., 2006; Kamp-Becker et al., 2010; Kuhlthau et al., 2010), but children with ASD had significantly lower scores for psychosocial health but not physical health than other clinical populations with chronic conditions (Kuhlthau et al., 2010), and adolescents and young adults with HFASD had higher scores than patients with schizophrenia-spectrum disorders except for the 'social relationship' domain (Kamp-Becker et al., 2010). Future research should aim to clarify the relationships between the various QoL domains in ASD. Fourth, past history was retrospectively obtained only from parents and was not based on a review of the clinical records, so there is a chance that the history has been influenced by parents' recall or memory bias.

Clinical implications. Despite these methodological limitations, this study points to some important clinical issues. First, clinicians can help children maximize their chances for high long-term QoL by changing environmental factors and treating comorbid psychiatric conditions related to aggressive behaviors, both of which may affect psychosocial well-being and QoL, even if the autistic core symptoms are largely not changeable. Bastiaansen et al. (2005) demonstrated that the QoL of a subgroup of child psychiatric patients improved although the level of psychopathology remained high during a 1-year follow-up period. Therefore, improving QoL should be included as one of the goals in treating individuals with HFASD.

Second, the present study provides evidence for the long-term significance of early detection and intervention for children with HFASD. Although there has been controversy about the positive and negative effects of early diagnosis for parents (Johnson et al., 2007), it may be important for clinicians to convey to parents – and empower them – that prognosis is not deterministic and may be changed by appropriate treatment and family support (Tantam, 2000).

Conclusions

This study demonstrated that self-reported QoL by adults with HFASD can be an important subjective aspect of long-term outcomes. Environmental factors, such as mother's support being helpful and early diagnosis, were associated with better QoL, and aggressive behaviors were associated with poorer QoL in adulthood, whereas expressive language level in preschool years, a conventional outcome indicator, did not predict QoL levels. To improve long-term QoL, professionals need to detect autistic symptoms in the early years, evaluate the needs of the child and family, provide consistent support, and comprehensively monitor all aspects of mental health. Future outcome studies should be conducted prospectively to determine predictive factors at each developmental stage and at the same time try to determine the mediators and moderators that modify the developmental trajectories for children with ASD.

Funding

This study was funded in part by the Japanese Ministry of Health, Labor, and Welfare (H19-SHOGAI-008), and an Intramural Research Grant (20B-5) for Neurological and Psychiatric Disorders of NCNP.

Acknowledgements

We thank all the participants for their co-operation. We also thank to Hisami Nishida and many facility staff and clinicians at the Support Centers for Persons with Developmental Disorders, Institutions for Persons with Autism, and Centers for Mental Health and Welfare Services.

References

Bastiaansen D, Koot HM, Ferdinand RF and Verhulst FC (2004) Quality of life in children with psychiatric disorders: self-, parent, and clinician report. *Journal of the American Academy of Child and Adolescent Psychiatry* 43(2): 221–230.

Bastiaansen D, Koot HM and Ferdinand RF (2005) Psychopathology in children: improvement of quality of life without psychiatric symptom reduction? *European Child and Adolescent Psychiatry* 14(7): 364–370.

Doi T (1973) *The anatomy of dependence.* Tokyo: Kodansya International.

Fountain C, King MD and Bearman PS (2011) Age of diagnosis for autism: individual and community factors across 10 birth cohorts. *Journal of Epidemiological Community Health* 65: 503–510.

Gerber F, Baud MA, Giroud M and Carminati GG (2008) Quality of life of adults with pervasive developmental disorders and intellectual disabilities. *Journal of Autism and Developmental Disorders* 38: 1654–1665.

Honda H, Shimizu Y, Imai M and Nitto Y (2005) Cumulative incidence of childhood autism: a total population study of better accuracy and precision. *Developmental Medicine and Child Neurology* 47: 10–18.

Howlin P, Goode S, Hutton J and Rutter M (2004) Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry* 45: 212–229.

Jennes-Coussens M, Magill-Evans J and Koning C (2006) The quality of life of young men with Asperger syndrome: a brief report. *Autism* 10: 403–414.

Johnson C, Myers S and the Council on Children with Disabilities of the American Academy of Pediatrics (2007) Identification and evaluation of children with autism spectrum disorders. *Pediatrics* 120: 1183–1215.

Kamio Y and Inokuchi E (2009) Psychiatric practice's role for individuals with developmental disorders: Current trend and future issues. *Journal of Japanese Association of Psychiatric Hospitals* 28: 14–20.

Kamio Y, Tobimatsu S and Fukui H (2011) Developmental disorders. In: Decety J and Cacioppo J (eds) *The Oxford Handbook of Social Neuroscience.* Oxford: Oxford University Press, 848–858.

Kamp-Becker I, Schröder J, Remschmidt H and Bachmann CJ (2010) Health-related quality of life in adolescents and young adults with high functioning autism-spectrum disorder. *GMS Psycho-Social-Medicine* Aug 31; 7. pii: Doc03.

Kobayashi R, Murata T and Yoshinaga K (1992) A follow-up study of 201 children with autism in Kyushu and Yamaguchi areas, Japan. *Journal of Autism and Developmental Disorders* 22: 395–411.

Koyama T, Kamio Y, Inada N, Adachi J, Uno Y and Kasahara M (2009) Laifusutêji ni okeru syujyunoyoin to chokiyogo tono kannren nikansuru kenkyu: Chokiyogo ni kansuru zenkokuchosa wo motoni [Examinations on the association between long-term outcome and related factors across life stage]. In: Kamio Y (ed) [*Annual Report of Research Supported by Health and Labour Sciences Research Grants*]. Tokyo: National Center of Neurology and Psychiatry.

Kuhlthau K, Orlich F, Hall TA, Sikora D, Kovacs EA, Delahaye J et al. (2010) Health-related quality of life in children with autism spectrum disorders: Results from the Autism Treatment Network. *Journal of Autism Developmental Disorders* 40: 721–729.

Landa RJ (2008) Diagnosis of autism spectrum disorders in the first 3 years of life. *Nature Clinical Practice Neurology* 4: 138–147.

Mawhood L, Howlin P and Rutter M (2000) Autism and developmental receptive language disorder-a comparative follow-up in early adult life. I: Cognitive and language outcomes. *The Journal of Child Psychology and Psychiatry* 41: 547–559.

Ministry of Education, Culture, Sports, Science and Technology (2009) *School Basic Survey 2009.* Nikkei Insatu. Available at: http://www.mext.go.jp/b_menu/toukei/001/08121201/1282588.htm.

Nakane Y, Tazaki M and Miyaoka E (1999) WHOQOL-BREF survey of general population. *Iryo To Shakai (Journal of Health Care and Society)* 9: 123–131.

National Institute of Population and Social Security Research (2007) Report on the thirteenth Japanese National Fertility Survey in 2005. Volume II: Attitudes towards marriage and the family among Japanese singles, March 2007. Available at: http://www.ipss.go.jp/syoushika/bumken/DATA/pdf/132542.pdf.

Persson B (2000) Brief report: A longitudinal study of quality of life and independence among adult men with autism. *Journal of Autism and Developmental Disorders* 30: 61–66.

Renty J and Roeyers H (2006) Quality of life in high-functioning adults with autism spectrum disorder. *Autism* 10: 511–524.

Ruble L and Dalrymple N (1996) An alternative view of outcome in autism. *Focus on Autism and Other Developmental Disabilities* 11: 3–14.

Saldaña D, Álvarez RM, Lobatón S, Lopez AM, Moreno M and Rojamo M (2009) Objective and subjective quality of life in adults with autism spectrum disorder in southern Spain. *Autism* 13: 303–316.

Tantam D (2000) Adolescence and adulthood of individuals with Asperger syndrome. In: Klin A, Volkmar FR and Sparrow SS (eds) *Asperger Syndrome.* New York: Guilford, 367–399.

Tazaki M and Nakane Y (2007) *WHOQOL26 Tebiki Kaiteiban.* Tokyo: Kaneko Shobo.

Tsatsanis K (2003) Outcome research in Asperger syndrome and autism. *Child and Adolescent Psychiatric Clinics of North America* 12: 47–63.

The WHOQOL Group (1995) World Health Organization quality of life assessment (WHOQOL): Position paper from the World Health Organization. *Social Science & Medicine* 41: 1403–1409.

Neurology Asia 2013; 18(1) : 45 – 47

Neuropsychiatric comorbidities in autism spectrum disorders without intellectual disability

Yoko Kamio *MD PhD*, Aiko Moriwaki *PhD*, Eiko Inokuchi *MD*

National Institute of Mental Health, National Center of Neurology and Psychiatry, Tokyo, Japan

Abstract

Epilepsy and autism spectrum disorder tend to co-occur in the population with intellectual disability. However, in the autistic population without intellectual disability, the prevalence of epilepsy is also much greater than in the general population. The special health needs in children having autism spectrum disorder without intellectual disability, namely those with high-functioning autism spectrum disorder have become recognized in recent years, yet comorbid neuropsychiatric symptoms such as anxiety, attention-deficit/hyperactivity disorder, and epilepsy still sometimes remain undiagnosed and untreated. Heightened awareness of such comorbidities will help these children to access appropriate treatment. Whether the epilepsy associated with high-functioning autism spectrum disorder is the same or different from that associated with intellectual disability, and whether the autistic profile associated with epilepsy in high-functioning autism spectrum disorder is the same or different from that without epilepsy, should be answered by future studies.

INTRODUCTION

Evidence from a community-based study[1] and numerous clinical reports indicate that a high proportion of individuals with autism spectrum disorder (ASD) suffer from one or more comorbid neuropsychiatric disorders. An association between epilepsy and ASD is well recognized, and comorbidity tends to be accompanied by intellectual disability. According to a meta-analysis of 23 studies[2], the pooled prevalence rates of epilepsy are 21.4% for individuals with ASD (defined as autism and/or pervasive developmental disorder) and intellectual disability, and 8% for those without intellectual disability, both of which are notably higher than the rate of 0.5% for the general population. In addition, sex seems to be another factor influencing the prevalence rate of epilepsy in the autistic population; epilepsy is more prevalent in autistic females than in autistic males, suggesting a close association between epilepsy and a female-predominant subgroup with ASD.[2]

Recently, heightened awareness of milder autistic conditions without intellectual disability has led to a higher overall prevalence rate of ASD of 2.6%[3], and has highlighted the clinical needs of children with ASD who have been undiagnosed until school age. In fact, a recent U.S. study found that most children with ASD were first identified as having ASD after age 5.[4] Further, it is reported that being older at first diagnosis of ASD is one of associated factors to reduce quality of life (QOL) for adults with high-functioning ASD, together with having a comorbid psychiatric disorder and being female by a nationwide survey conducted in Japan.[5] Taken together, it is clear that the early identification and treatment of ameliorable comorbid neuropsychiatric disorders such as depression and epilepsy is important to improving the QOL for autistic individuals with or without accompanying intellectual disability.

PREVALENCE OF EPILEPSY IN AUTISM SPECTRUM DISORDER WITHOUT INTELLECTUAL DISABILITY

Previous studies have reported a relatively lower prevalence of epilepsy in high-functioning ASD compared to that in ASD with intellectual disability, but the prevalence is still greater than that found in the general population. A large cohort study in the U.K. identified epilepsy in 8.7% (2/23) of children with Asperger syndrome compared to 16.7% (5/30) of children with childhood autism.[6] A large population-based survey in Finland found epilepsy, defined according to the International League Against Epilepsy (ILAE), in 18.2% (34/187) of children with autistic disorder and 12.1% (11/91) of those with IQ >70.[7] However, no firm statistical conclusion can be drawn about the type of epilepsy from these data. A clinical

Address correspondence to: Yoko Kamio *MD PhD*, Department of Child and Adolescent Mental Health, National Institute of Mental Health, National Center of Neurology and Psychiatry, 4-1-1 Ogawa-Higashi, Kodaira, Tokyo 187-8553, Japan. E-mail: kamio@ncnp.go.jp

study of 100 boys with Asperger syndrome made a conservative comorbidy estimate of epilepsy in 4 patients.[8] Another clinical study comparing 26 patients with Asperger syndrome and 16 patients with high-functioning autism found no significant differences between the groups in electroencephalogram abnormalities (8.7% vs 13.3%), epilepsy (7.7% vs 6.3%) or clinical variables.[9]

Recently, we conducted a small-scale study in the west of Tokyo to determine the comorbid neuropsychiatric disorders associated with high-functioning ASD (HFASD).[10] The target population was primary school children aged 6-12 years in mainstream classes (n=1,374), of which 775 participants were screened using teacher-report autism questionnaires. Following semi-structured diagnostic interviews with all screen-positives and randomly selected screen negatives, 7 children were identified as having definite ASD and 3 as having broader ASD. None had intellectual disability. One or more diagnoses according to the Text revision of the Diagnostic and Statistical Manual of Mental Disorders, the fourth edition (DSM-IV-TR) was found in 72% of children with definite HFASD and 100% of children with broader ASD, findings consistent with those of the UK study.[1] The distribution pattern of comorbid disorders is also similar; anxiety or phobic disorders and oppositional or conduct disorders being the most common, with a prevalence of up to 40%. Most of these children were undiagnosed and had received no professional health interventions in terms of these comorbid disorders. Epilepsy was found in one boy in the HFASD sample (1/7, 14.3%). He was diagnosed as having complex partial seizure upon his first seizure at age 4 and has been treated using valproic acid. He is currently seizure free but has attention problems. In addition, two girls in our sample (one with Asperger syndrome, one with broader ASD) had repeated generalized seizures over the last 1-3 years, although they were not diagnosed with epilepsy. Since some individuals with ASD are still at a risk of developing epilepsy after puberty, the rate of 14.3% should not be overestimated. Our sample was small, however, the results emphasize that there does seem to be a high rate of children who develop epilepsy in autistic population without intellectual disability.

Many issues regarding the association between epilepsy and HFASD remain unanswered. Pediatric neurologists may want to know whether the epilepsy associated with HFASD is the same as that associated with ASD plus intellectual disability; or from another viewpoint, child psychiatrists may want to know how the autistic profile associated with epilepsy in HFASD is different from that without epilepsy.

CONCUSIONS

Recently, the special health needs of children with HFASD have been recognized. The prevalence of epilepsy is much higher in children with HFASD than it is in the general population. Children with HFASD are likely to have additional psychiatric symptoms such as anxiety and attention-deficit/hyperactivity disorder, but such health problems are often undiagnosed and untreated. Comprehensive neuropsychiatric evaluations of children with HFASD or children with epilepsy will lead to early identification of treatable health problems and the provision of appropriate treatment. Some behavioral problems in ASD can be improved with antiepileptic drugs.

Given that there are approximately 2-3% of children having ASD, and 10% of children with subthreshold autistic traits[11], an approach using quantitatively measured autistic traits may be also helpful to explore the association between epilepsy and autism.

ACKNOWLEDGEMENTS

This work was supported by grants [H19-KOKORO-006 and H20-KOKORO-004] from the Japanese Ministry of Health, Labor, and Welfare.

REFERENCES

1. Simonoff E, Pickles A, Charman T, *et al*. Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity, and associated factors in a population-derived sample. *J Am Acad Child Adolesc Psychiatry* 2008; 47:921-9.
2. Amiet C, Gourfinkel-An I, Bouzamondo A, *et al*. Epilepsy in autism is associated with intellectual disability and gender: Evidence from a meta-analysis. *Biol Psychiatry* 2008; 64:577-82.
3. Kim YS, Leventhal BL, Koh YJ, *et al*. Prevalence of autism spectrum disorders in a total population sample. *Am J Psychiatry* 2011; 168:904-12
4. Pringle BA, Colpe LJ, Blumberg SJ, Avila RM, Kogan MD. Diagnostic History and Treatment of School-Aged Children with Autism Spectrum Disorder and Special Health Care Needs. NCHS data brief, no 97. Hyattsville, MD: National Center for Health Statistics. 2012.
5. Kamio Y, Inada N, Koyama T. A nationwide survey on quality of life and associated factors of adults with high-functioning autism spectrum disorders.

Autism, first published on March 7, 2012 as doi:10.1177/1362361312436848
6. Williams E, Thomas K, Sidebotham H, Emond A. Prevalence and characteristics of autistic spectrum disorders in the ALSPAC cohort. *Dev Med Child Neurol* 2008; 50:672-7.
7. Kielinen M, Rantala H, Timonen E, Linna S-L, Moilanen I. Associated medical disorders and disabilities in children with autistic disorder. *Autism* 2004; 8:49-60.
8. Cederlund M, Gillberg C. One hundred males with Asperger syndrome: a clinical study of background and associated factors. *Dev Med Child Neurol* 2004; 46:652-60.
9. Kurita H. A comparative study of Asperger syndrome with high-functioning atypical autism. *Psychiatr Clin Neurosci* 1997; 51:67-70.
10. Kamio Y, Inokuchi E, Moriwaki A, *et al*. Prevalence of developmental disorders and the associated factors in general child population. In: KAMIO Y, ed. Annual report of research supported by health and labour sciences research grants [in Japanese]. National Center of Neurology and Psychiatry. Tokyo, 2011Kamio Y, Inada N, Moriwaki A, *et al*. Quantitative autistic traits ascertained in a national survey of 22,529 Japanese schoolchildren. *Acta Psychiatrica Scandinavica*, DOI 10.1111/acps.12034

Acta Psychiatrica Scandinavica

Acta Psychiatr Scand 2013: 128: 45–53
All rights reserved
DOI: 10.1111/acps.12034

ACTA PSYCHIATRICA SCANDINAVICA

Quantitative autistic traits ascertained in a national survey of 22 529 Japanese schoolchildren

Kamio Y, Inada N, Moriwaki A, Kuroda M, Koyama T, Tsujii H, Kawakubo Y, Kuwabara H, Tsuchiya KJ, Uno Y, Constantino JN. Quantitative autistic traits ascertained in a national survey of 22 529 Japanese schoolchildren.

Objective: Recent epidemiologic studies worldwide have documented a rise in prevalence rates for autism spectrum disorders (ASD). Broadening of diagnostic criteria for ASD may be a major contributor to the rise in prevalence, particularly if superimposed on an underlying continuous distribution of autistic traits. This study sought to determine the nature of the population distribution of autistic traits using a quantitative trait measure in a large national population sample of children.
Method: The Japanese version of the Social Responsiveness Scale (SRS) was completed by parents on a nationally representative sample of 22 529 children, age 6–15.
Results: Social Responsiveness Scale scores exhibited a skewed normal distribution in the Japanese population with a single-factor structure and no significant relation to IQ within the normal intellectual range. There was no evidence of a natural 'cutoff' that would differentiate populations of categorically affected children from unaffected children.
Conclusion: This study provides evidence of the continuous nature of autistic symptoms measured by the SRS, a validated quantitative trait measure. The findings reveal how paradigms for diagnosis that rest on arbitrarily imposed categorical cutoffs can result in substantial variation in prevalence estimation, especially when measurements used for case assignment are not standardized for a given population.

Y. Kamio[1], N. Inada[1], A. Moriwaki[1], M. Kuroda[1,2], T. Koyama[1], H. Tsujii[1], Y. Kawakubo[2], H. Kuwabara[2], K. J. Tsuchiya[3], Y. Uno[4], J. N. Constantino[5]

[1]Department of Child and Adolescent Mental Health, National Institute of Mental Health, National Center of Neurology and Psychiatry, Tokyo, Japan, [2]Department of Child Neuropsychiatry, Graduate School of Medicine, University of Tokyo, Tokyo, Japan, [3]Research Center for Child Mental Development, United Graduate School of Child Development, School of Medicine, Hamamatsu University, Hamamatsu, Japan, [4]Department of Psychiatry and Psychiatry for Parents and Children, Graduate School of Medicine, Nagoya University, Nagoya, Japan and [5]Departments of Psychiatry and Pediatrics, School of Medicine, Washington University, St. Louis, MO, USA

Key words: autism; questionnaire; prevalence; classification; diagnosis

Yoko Kamio, National Institute of Mental Health, National Center of Neurology and Psychiatry, 4-1-1 Ogawa-Higashi, Kodaira, Tokyo 187-8553, Japan.
E-mail: kamio@ncnp.go.jp

Accepted for publication October 1, 2012

Significant outcomes

- In a large Japanese child population, behaviorally measured autistic traits are continuously distributed without any apparent deflection in the distribution plot that would signal a natural cutoff for categorical diagnoses. This is similar to the distribution pattern in US and European samples.
- Autistic traits measured quantitatively by parents differ slightly by culture, suggesting the need to interpret autism spectrum disorder (ASD) severity ratings with the use of culturally calibrated norms.
- Many children who do not meet the diagnosis of ASD exhibit elevations in autistic traits measured quantitatively, suggesting the need to reconsider current diagnostic systems that assume discontinuity between affected and unaffected populations.

Limitations

- The response rate of this nationwide survey was 29%.
- There is a possibility of bias that would differentiate respondents vs. non-respondents.
- High-scoring children in the sample as a whole were not confirmed using diagnostic instruments, although quantitatively measured autistic traits were extensively clinically confirmed for a separate smaller sample.

Introduction

Although to date the designation of pervasive developmental disorders in children – and the services to which affected children are entitled – rest on categorical case definitions, the concept of an autistic *spectrum*, along which the number and intensity of autistic features vary continuously from mild to severe, dates back to early epidemiological research by Wing and Gould (1). Wing (2) subsequently developed the concept of the autistic continuum, broadening the case designation beyond classic autism to encompass the mildest (but most prevalent) of the autism spectrum disorders (ASDs), pervasive developmental disorder not otherwise specified (PDD-NOS) assigned by diagnostic and statistical manual of mental disorders: text revision (DSM-IV-TR) (3). Several lines of subsequent research (4–7) now strongly suggest that the autism spectrum extends beyond this PDD-NOS subcategory to include subclinical levels of symptomatology, which are known to aggregate in the undiagnosed members of families with multiple-incidence autism. Very recently, Lord et al. (8) observed that diagnostic assignments of autistic disorder, Asperger's disorder, and PDD-NOS made by expert clinicians varied considerably across sites, despite the fact that distributions of scores on validated measures were similar. They concluded that current taxonomies should be revised to place priority on characterizing the dimensions of ASD while controlling for IQ and language level.

Clarifying the nature of the population distribution of autistic traits and symptoms across cultures has substantial implications for understanding a rise in prevalence over time (9) and for establishing the 'boundaries' of clinical affectation. A recent Korean study (10) suggested the highest ever reported prevalence for categorically defined ASD in a total population sample; in that study, symptom counts were found to be continuously distributed in the population.

Aims of the study

This study determined whether autistic traits would be continuously distributed in a population-based sample to establish the appropriate epidemiologic framework for interpreting the rise in estimated autism spectrum disorders prevalence over time.

Material and methods

Participants

The participants comprised a normative sample (n = 22 529) of schoolchildren, a child psychiatric clinical sample (n = 417), and typically developing (TD) children (n = 61). The normative sample was exclusively assessed using the Japanese version of the Social Responsiveness Scale (SRS) (11). The latter two samples were more extensively assessed using standard diagnostic batteries for the purpose of validation and calibration of the Japanese version of the SRS.

In regard to the normative sample, questionnaires were distributed by mail to the caregivers of all students attending mainstream classes at primary or secondary schools in the 10 geographical areas making up Japan in 2010 (n = 87 548 caregivers). One hundred and forty-eight primary schools and 71 secondary schools participated in this study. All of them were community schools where >93% of children living in the community attend, according to the annual report of Japan's Ministry of Education, Culture, Sports, Science and Technology, 2010 (12). Questionnaires were returned for 25 779 children aged 6–15 years (response rate 29.4%). Questionnaires with missing answers were excluded so that all analysis was based on a complete data set, leaving a final normative sample of 22 529 participants (11 455 boys) with SRS data provided by their mothers (n = 20 430), fathers (n = 1728), both parents (n = 166), other caregivers (n = 119) or unspecified (n = 86). Each of the 9 grade levels comprised a minimum of 754 participants of each sex, and both sexes were proportionally represented (Table 1).

The clinical sample consisted of 257 children diagnosed with ASD (ASD group) and 157 children with psychiatric diagnoses other than ASD (non-ASD group) (Table 2). They were patients who visited one of 10 child psychiatric clinics dur-

Table 1. Social Responiveness Scale total raw score distributions in the normative sample by sex and age (grade)

	Sex						
	Males		Females				
Grade	N	Mean (SD)	N	Mean (SD)	t	P	d
1	1655	37.3 (18.2)	1473	33.0 (16.7)	44.3	0.000	0.25
2	1521	36.2 (18.2)	1394	32.1 (16.3)	37.8	0.000	0.24
3	1384	35.4 (19.2)	1432	31.2 (16.4)	39.0	0.000	0.24
4	1375	33.7 (18.4)	1386	30.2 (16.3)	26.2	0.000	0.20
5	1449	33.0 (18.5)	1287	31.0 (17.5)	8.6	0.003	0.11
6	1203	31.9 (19.6)	1229	29.9 (17.8)	6.9	0.009	0.11
7	1072	32.3 (19.1)	1070	30.3 (17.8)	6.7	0.010	0.11
8	1007	32.7 (20.2)	1049	29.8 (18.2)	12.7	0.000	0.15
9	789	31.7 (20.7)	754	28.9 (18.6)	9.2	0.002	0.14
Total	11 455	34.1 (19.1)	11 074	30.9 (17.2)	13.4	0.000	0.18
Total children	22 529		32.5 (18.3)				

Grade 1 children are usually 6–7 years old. Most grade 1 participants were 7 years old at the time of the survey.

ing 2008–2010 and whose caregivers gave informed consent to participate in this study. Their existing clinical diagnoses were confirmed according to DSM-IV-TR criteria (3) based on all of the clinical information available to our research team, which included experienced child psychiatrists and licensed clinical psychologists. Among the 257 children of the ASD group, 229 were subcategorized with 100% diagnostic agreement: 96 with autistic disorder, 65 with Asperger's disorder, 68 with PDD-NOS, and 28 were unspecified. Children in the non-ASD group were diagnosed with adjustment disorder, attention deficit hyperactivity disorder, anxiety disorder, eating disorder, schizophrenia, somatoform disorder, conduct disorder, mood disorder, or mental retardation. Moreover, 61 children recruited from local communities comprised a TD group and were confirmed in diagnostic interviews with the children and their parents to have no history of neuropsychiatric conditions.

The intellectual levels of the children in the clinical sample ranged from normal intelligence to severe mental retardation based on cognitive testing carried out at clinics [various versions of the Wechsler Intelligence Scale and the Revised Kyoto Scale of Psychological Development (13)] or educational/administrative records. The proportions of children with normal intelligence in the ASD and non-ASD groups were not significantly different ($\chi^2 = 1.42$, n.s.).

Measures

The social responsiveness scale. The SRS (11) is a 65-item questionnaire of autistic traits for use with 4- to 18-year-olds that can be completed in 15 min by any adult who has observed the child over time in naturalistic social settings. The SRS was developed to assess autistic symptoms or quantitative traits and has subsequently undergone extensive validation in US samples for use in subclinical and clinical child populations(4, 14–17) as well as in general child populations for behavioral genetic research (18–20). It also demonstrated satisfactory internal consistency (Cronbach's $\alpha > 0.95$), inter-rater reliability between parents and teachers ($r = 0.78$, $P < 0.01$), and concurrent validity with an interview-based instrument(21) ($r = 0.86$, $P < 0.05$ for preschoolers; $r = 0.48$, $P < 0.05$ for children aged 7–12; $r = 0.77$, $P < 0.001$ for adolescents aged 13–18) for Japanese children(22, 23) and also for German children(24). The Japanese version was used in this study. Higher scores on the SRS indicate higher degrees of social impairment. The 65 SRS items were further categorized into five treatment subscales (social awareness, social cognition, social communication, social motivation, autistic mannerisms) (11). The SRS total scores are generally unrelated to IQ in the normal range and distinguish children with ASD from those with other types of psychopathology (16).

Table 2. Comparison of Social Responsiveness Scale total raw score between the United States and Japan

	Country						
	Japan		US				
Grade	N	Mean (SD)	N	Mean (SD)	t	P	d
1	3102	35.3 (17.6)	71	29.6 (25.6)	1.87	0.06	0.318
2	2891	34.2 (17.4)	92	34.9 (26.9)	0.25	0.80	0.041
3	2786	33.2 (18.0)	109	35.7 (26.8)	0.97	0.33	0.136
4	2739	31.9 (17.5)	227	35.3 (24.9)	2.02	0.04	0.188
5	2703	32.0 (18.0)	214	34.5 (25.3)	1.42	0.16	0.134
6	2408	30.8 (18.7)	211	31.7 (21.5)	0.59	0.56	0.049
7	2123	31.3 (18.4)	161	31.1 (20.6)	0.12	0.90	0.008
8	2040	31.1 (19.1)	137	31.9 (23.7)	0.39	0.70	0.040
9	1532	30.2 (19.7)	124	38.9 (29.2)	3.26	0.00	0.422
Total	22 344	32.5 (18.2)	1626*	33.6 (24.7)	1.76	0.08	0.051

Grade 1 children are usually 6–7 years old. Most grade 1 participants were 7 years old at the time of the survey.
*US data were cited from the SRS manual (p. 28) (11).

The autism diagnostic interview-revised. The Autism Diagnostic Interview-Revised (ADI-R) (25) is a parent-report interview and is a research standard for establishing a diagnosis of autism. To meet the ADI-R criteria for autism, the cutoff must be reached in each domain of reciprocal social interaction, communication, and restricted, repetitive, and stereotyped patterns of behavior. The Japanese version of the ADI-R was used in this study, which has demonstrated good reliability and validity for Japanese children (26).

Ethical issues

The study protocol was approved by the Ethics Committee of the National Center of Neurology and Psychiatry, Japan. Written informed consent to participate was obtained from the caregivers of each child participant.

Data analysis

Following examination of the SRS distribution as a function of age and sex, a cross-cultural comparison of SRS total scores provided by parents was performed between previously reported US norms (the SRS manual, p. 28) (11) and the obtained Japanese scores using t-tests. Factor analysis was performed using principal components analysis

(PCA) on children in the ASD, non-ASD, and TD groups, and the most parsimonious model was subsequently examined by confirmatory factor analysis (CFA) in the normative sample. To address discriminant validity, comparisons of the SRS scores across diagnostic groups were made using analysis of variance (ANOVA) methods with Bonferroni correction whenever appropriate. Intraclass correlation coefficient (ICC) was computed for associations between SRS scores, full scale IQ, and ADI-R algorithm scores. In addition, a receiver operating characteristics (ROC) analysis was conducted to determine the cutoff points for primary and secondary screening; for the former, the cutoff point was where the sum of sensitivity and specificity was the largest, and for the latter, it was where the likelihood was the largest for children in the ASD, non-ASD, and TD groups, for boys and girls separately. Analysis was performed using SPSS 18.0J for Windows (SPSS Japan Inc., Tokyo, Japan), with AMOS 17.0J for Windows (SPSS Japan Inc., Tokyo, Japan) used for the confirmatory factor analysis.

Results

Population distribution

Social Responsiveness Scale score distribution among 6- to 15-year-old children in the Japanese general population is shown in Fig. 1, and mean SRS total raw scores by age group are presented for boy and girl subsamples in Table 1. To investigate the effects of age (grade) and sex on SRS scores, a 2-way ANOVA (grade × sex) was conducted on the total raw scores. The interaction was significant ($F_{8,180,224} = 2.00$, $P < 0.05$, $\eta^2 = 0.00$), and the main effects of grade ($F_{8,180,224} = 20.03$, $P < 0.001$, $\eta^2 = 0.01$) and sex ($F_{8,180,224} = 157.37$, $P < 0.001$, $\eta^2 = 0.01$) were significant, although the effect size indicates that the differences in the SRS scores by grade and sex were modest.

Mean SRS score of each age group was within 0.2 standard deviations of the entire sample means for boys and girls respectively (boys 30.3–37.9, girls 27.5–34.3). Boys scored higher than girls across the entire age range, with the maximum sex difference seen for the youngest subgroup at grade 1 ($t = 44.24$, $P < 0.001$, $d = 0.25$). Therefore, we standardized the Japanese version of the SRS on each of the boy and girl subsamples across the age range (27).

Table 2 shows our Japanese normative data together with the original US parent and teacher rating data (the SRS manual, p. 28) (11) derived from five different studies. Japanese children scored similarly to their US counterparts, except those in grades 4 and 9; here, Japanese children had significantly lower mean SRS scores than their US counterparts.

Factor structure. PCA suggested a one-factor solution for the 475 children comprising the clinical and TD groups (Table 3). Seven items (items 24, 29, 35, 37, 44, 49, 51) with factor loadings >0.600 represented all three of the DSM-IV-TR criterion domains for autism. When 22 items with factor loadings <0.400 were excluded, the first factor explained 34.8% of variance in SRS scores in this sample, consistent with the original US and German data for child psychiatric patients. When performed with the mean scores of the five treatment subscales, rather than the mean scores of 65 items, PCA gave a one-factor solution accounting for 77.2% in this sample.

Next, the single-factor model suggested by PCA and by extensive prior research on the SRS (20, 24) was subjected to CFA using data from the normative sample. The comparative fit index, the goodness of fit index, the adjusted goodness of fit index, and root mean square error of approximation were 0.677, 0.739, 0.722, and 0.055 for all 65 items, 0.811, 0.854, 0.840, and 0.055 for 43 items with factor loadings >0.400 derived from PCA on the exploratory set, and 0.989, 0.987, 0.962, and 0.083 for the five treatment subscales. These findings lend support to the notion of a unitary factor influencing the multiple aspects of dysfunction that characterize autistic symptomatology in children in the general population.

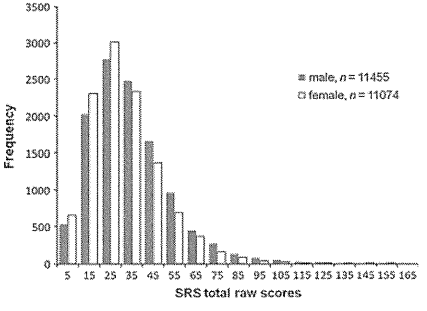

Fig. 1. Distribution of Social Responsiveness Scale (SRS) total raw scores rated by caregivers in the general sample of 6- to 15-year-old children.

Table 3. Principal components analysis of social responsiveness scale data

	ASD, non-ASD, and TD groups (n = 475)		
Component	Total	% of variance	Cumulative%
1	18.928	29.120	29.120
2	3.851	5.925	35.045
3	3.152	4.850	39.895
4	1.926	2.963	42.858
5	1.701	2.616	45.474

ASD, autism spectrum disorders; TD, typical development.
The clinical sample consisted of participants with ASD (n = 257) and non-ASD (n = 157).

Other psychometric properties

Table 4 indicated that the mean SRS total score of the ASD group was significantly higher than that of the clinical non-ASD (boys $t = 4.87$, $P < 0.001$, $d = 0.65$, girls $t = 4.68$, $P < 0.001$, $d = 0.83$) and TD (boys $t = 11.73$, $P < 0.001$, $d = 2.29$, girls $t = 11.80$, $P < 0.001$, $d = 2.66$) groups. The differences in SRS score were not pronounced among the ASD subcategories: the score did not discriminate between Asperger's disorder and PDD-NOS for either sex, as previously reported (23). As shown in Fig. 2, the SRS scores of both ASD and non-ASD groups were distributed widely and with significant overlap with the general population distribution. Table 5 shows the raw score cutoffs for the 99th, 97.5th, 95th, and 90th percentile values by sex for our normative sample and the proportion of boys and girls with diagnosed ASD who fell within the respective percentile cutoffs. In general, a higher proportion of diagnosed females were at the more extreme percentile rankings in comparison with males.

Social Responsiveness Scale score did not correlate with IQ (ICC = −0.23, n.s.) for 118 participants with IQs > 70 for whom formal test data were available (ASD 46, non-ASD 11, TD 61), although the subgroup with mental retardation tended to score higher. With regard to autistic symptoms, SRS score was significantly correlated with ADI-R total score (ICC = 0.66, $P < 0.001$; Fig. 3), as well as scores for the social interaction domain (ICC = 0.68, $P < 0.001$), communication domain (ICC = 0.58, $P < 0.001$), and restricted and repetitive behavior domain (ICC = 0.50, $P < 0.001$) for a subsample for whom data from both the SRS and ADI-R were available (n = 36; ASD 20, non-ASD 10, TD 6; mean age 8.0 years, range 4–18 years).

Receiver operating characteristics analysis informed two sets of cutoff points depending on the purpose of use. When used for primary screening of the general child population such as at school entrance, an optimal cutoff point was 53.5 for boys (sensitivity 0.91, specificity 0.48) and 52.5 for girls (sensitivity 0.89, specificity 0.41). For secondary screening of children referred to clinical settings, where a much higher rate of ASD is expected, the cutoff point of 109.5 for boys (sensitivity 0.23, specificity 0.96, likelihood ratio 6.14) and 102.5 for girls (sensitivity 0.32, specificity 0.95, likelihood ratio 5.73) increases the positive predictive value for ASD diagnosis up to 80.4% for boys and 79.2% for girls, given that the prevalence in Japanese child psychiatric clinics is 40%. Primary and secondary screening cutoffs correspond to a SRS T-score of 60 and 90 for boys and 62 and 92 for girls respectively.

Discussion

We conclude from these data involving a nationwide representative sample of schoolchildren that autistic traits measured by the Japanese version of the SRS are distributed continuously in the population; that the clinical validity of the measurements (in essence, their relevance to autism) appeared strong; and that the findings of this cross-cultural study recapitulate and extend what has been observed in smaller epidemiologic studies of autistic traits in other countries.

The results of this study of quantitative autistic traits – the largest of its kind – add substantial evidence in support of the continuous nature of autistic traits in the general population. This does not mean that individual cases of autism are never discretely or categorically determined. It has long been known, for example, that there exist categorical, relatively rare causes of autistic syndromes (e.g., fragile X syndrome, Rett syndrome, and tuberous sclerosis) caused by single gene abnormalities. The notion of an autistic continuum remains consistent with the existence of such discrete entities. The same is true for mild to moderate intellectual disability, which constitutes the extreme end of a normal distribution (the so-called 'bell curve') but comprises a number of discrete syndromes (including but not limited to Down syndrome, Fragile X syndrome, etc.) in the severe end of the symptom distribution. Similarly, segments of the autistic continuum may be comprised of small clusters of discrete disorders (e.g., SHANK 1 mutations, 15q duplications, 16p11.2 deletions) that contribute to intervals at the pathological end of the distribution (for example 75–85, 90–110), but overlap in severity with other cases that represent quantitative accumulations of inherited liability transmitted by polygenic mechanisms or by gene–environment interactions. The causes of cases

Table 4. Social Responsiveness Scale total raw score means of the ASD, non-ASD, and TD groups

	ASD	nonASD	TD	ASD subcategory			
				Autism	Asperger's disorder	PDD-NOS	Unspecified
N (Male/female)	257 (203 : 54)	157 (78 : 79)	61 (30: 31)	96 (77 : 19)	65 (48 : 17)	68 (54 : 14)	28 (24 : 4)
Age (years)							
Mean (SD) Range	10.0 (3.9) 4–18	12.1 (3.7) 4–18	9.61 (2.5) 6–18	9.0 (4.2) 4–18	10.7 (3.1) 4–17	10.0 (4.1) 4–18	11.68 (3.67) 6–17
Intellectual level (N)							
Normal	181	118	57	57	64	59	1
Borderline	14	9	4	8	1	3	2
Mild MR	10	12	0	5	0	3	2
Moderate MR	7	3	0	2	0	1	4
Severe MR	12	8	0	2	0	0	10
MR (unknown level)	33	7	0	22	0	2	9
SRS Mean (SD) Range							
Males	87.6 (27.4) 15–158†	69.7 (27.9) 13–141†	27.4 (16.6) 6–72†	89.5 (24.0) 48–139‡	82.4 (26.8) 15–132	78.4 (26.5) 24–144‡	
Females	86.1 (27.9) 21–153§	62.1 (28.9) 12–134§	21.4 (16.2) 2–65§	91.4 (27.2) 21–133	91.0 (31.4) 38–153	74.7 (25.3) 40–114	
Total	87.3 (27.4) 15–158¶	65.9 (29.1) 12–141¶	24.3 (16.5) 2–72¶	89.8 (24.5) 21–139**	84.6 (28.1) 15–153	77.7 (26.1) 24–144**	

SRS, Social Responsiveness Scale; ASD, autism spectrum disorders; TD, typical development; PDD-NOS, pervasive developmental disorder not otherwise specified; MR, mental retardation.
†ASD > non-ASD, TD ($t = 4.87$, $P < 0.001$, $d = 0.65$, $t = 11.73$, $P < 0.001$, $d = 2.29$, respectively), non-ASD > TD ($t = 7.79$, $P < 0.001$, $d = 1.67$).
‡Autism > PDD-NOS ($t = 2.48$, $P < 0.05$, $d = 0.44$).
§ASD > non-ASD, TD ($t = 4.68$, $P < 0.001$, $d = 0.83$; $t = 11.80$, $P < 0.001$, $d = 2.66$, respectively), non-ASD > TD ($t = 7.17$, $P < 0.001$, $d = 1.52$).
¶ASD > non-ASD, TD ($t = 7.53$, $P < 0.001$, $d = 0.76$; $t = 17.19$, $P < 0.001$, $d = 2.45$, respectively), non-ASD > TD ($t = 10.51$, $P < 0.001$, $d = 1.59$).
**Autism > PDD-NOS ($t = 3.05$, $P < 0.05$, $d = 0.48$).

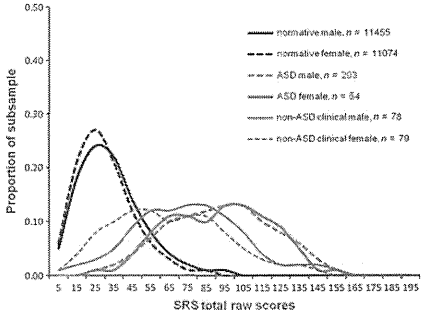

Fig. 2. Distribution of Social Responsiveness Scale (SRS) total raw scores in child psychiatric patients with and without autistic spectrum disorders (ASD).

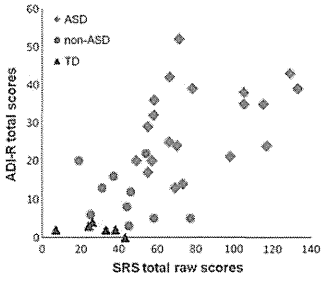

Fig. 3. Social Responsiveness Scale (SRS) total raw scores as a function of Autism Diagnostic Interview-Revised (ADI-R) total scores for children with autism spectrum disorders (ASD), non-ASD, and typical development (TD).

Table 5. Proportion of children with autism spectrum disorders (ASD) corresponding to the 99th, 97.5th, 95th, and 90th percentile values among the ASD group of the Japanese clinical sample

Normative sample (n = 22 529)			ASD group (n = 257)			
	Raw score cutoff		N (%)			
			Males		Females	
Percentile value	Males	Females	(n = 203)		(n = 54)	
≥ 99	98	87	70	34.5%	28	51.9%
≥ 97.5	81	73	117	57.6%	36	66.7%
≥ 95	70	63	147	72.4%	42	77.8%
≥ 90	58	53	173	85.2%	44	81.5%

represented by any given score in the distribution may be independent, partially overlapping, or fully overlapping with the underlying causes of other cases at the same level of severity. The result is a continuous distribution encompassing both discrete and quantitative pathways to affectation across a wide range of severity (28–32). We note that in a recent large general population twin study, Robinson et al. (33) demonstrated overlap in causal influence on autistic symptomatology at each of the first, second, and fifth percentiles of severity in the population.

In our study, there was no evidence of a natural cutoff that differentiated children categorically affected from those unaffected by ASD. The parent-report Japanese SRS cutoff scores for secondary screening derived from our ROC analysis, 109.5 for boys and 102.5 for girls, would comprise approximately 0.5% of our normative sample. On the other hand, the ASD primary screening cutoff with the highest sensitivity, 53.5 for boys and 52.5 for girls, encompassing 10.9% of our normative sample, identifies subthreshold conditions in children that might warrant clinical attention (11). Taken together, these findings complement a recent Korean study (10), in which categorical screening and diagnostic confirmation revealed (and validated) what a continuous distribution of symptom counts. In our normative sample, a parent-report Japanese SRS raw score of 74 for boys and 80 for girls would cut off approximately 3.74%, 1.47% of each gender-specific population distribution, which is very near the prevalence for ASD reported in the Korean study (2.64%) (10).

Our observation of higher quantitative autistic trait scores in males than in females confirms across cultures a subtle but statistically robust gender difference (11, 18, 24). The sex distribution pattern has potentially profound implications for sex disparities universally observed at the extreme end of the distribution (i.e., in clinical ASD cases), where such disparities would be expected to be accentuated, as is true for any normally distributed trait such as height. The magnitude of the sex difference in our sample (d = 0.18) was smaller than that in the US data set (11) (d = 0.37) but similar to the German normative sample (24) (d = 0.16). Accentuation of the gender difference in the US data set could potentially relate to its being derived from a twin sample, given that male twins score higher than non-twins (34). Japanese children diagnosed with ASD were rated as having somewhat lower quantitative trait scores than their US and German counterparts. Such cross-cultural differences could be partly explained by cultural differences in responding to Likert-type rating, on which Japanese informants have a higher tendency to use the midpoint on the scales and US informants a higher tendency to use the extreme values (35).

The results of the exploratory factor analysis for the clinical sample replicate those of previous

studies (17, 18), and the results of the confirmatory factor analysis for a very large general population underscore the presence of a primary underlying factor that influences the symptoms representing all three DSM-IV-TR criterion domains of autism. Factor structure has important implications for understanding the core neuropsychological mechanisms underlying autistic traits and symptoms, which are relevant to not only the pursuit of biomarkers and genetic susceptibility factors related to ASD but also diagnostic paradigms (20, 31).

There are two major limitations in this study. First, the response rate was low (29%), although it is keeping with what is expected from population-based surveys. Second, high-scoring children in 22 529 Japanese schoolchildren were not confirmed using any diagnostic instruments, although quantitatively measured autistic traits were extensively clinically confirmed for the separate smaller sample.

In the present study, although the instrument capably distinguished children diagnosed with ASD from children diagnosed with other psychiatric conditions, the score distribution for both clinical groups overlapped. A possible interpretation of this observation, given that autistic traits exhibit considerable independence in causation from many forms of psychopathology in genetic epidemiologic research (15, 36), is that autistic traits, when present, exacerbate other types of psychopathology when they cooccur with autistic traits as comorbid conditions. For some neurodevelopmental conditions, however, it has also become increasingly clear that there are elements of genetic causation that genuinely overlap with the genetic cause of autism; these include ADHD, tic disorders, and developmental coordination disorders, among others (37).

In conclusion, our study provides strong evidence of the continuous nature of autistic symptomatology in the general population, as has been reported in previous studies (1, 18, 19, 37). The findings underscore the notion that paradigms for categorical case assignment are superimposed on a continuous distribution, which can result in substantial variation in prevalence estimation, especially when the measurements used in case assignment are not standardized for a given population (i.e. by gender, informant, culture, etc.). In other words, these data illustrate that when imposing an arbitrary, non-standardized cutoff for diagnosis, small, clinically insignificant changes in the cutoff value can result in significant changes in prevalence, especially when operating at the steeper slopes of the distribution. Our results support the importance, validity, and feasibility of determining standardized quantitative ratings of autistic traits and symptoms across cultures, the implementation of which has the potential to advance international collaborative research on autism and related conditions. Finally, these results call for a rational approach to revising systems of diagnosis and service delivery that currently perpetuate the notion of discontinuity between ASD-affected and unaffected populations.

Acknowledgements

This study was supported by Research Grants from the Ministry of Health, Labour and Welfare of Japan to Dr. Kamio (H19-KOKORO-006 and H20-KOKORO-004) and by a grant from the National Institute of Child Health and Human Development to Dr. Constantino (HD42541). We would like to thank Drs. Norio Ozaki and Seiji Koishi for translating the SRS, Drs. Akiko Takaki, Miyako Shirakawa, Tokio Uchiyama, Masahiro Oshima, and Eiko Inokuchi for data collection, and Drs. Shoji Tanaka and Hisateru Tachimori for helpful discussion.

Declaration of interest

Dr. Constantino receives royalties for commercial distribution of the SRS, which is published by Western Psychological Services. No royalties were generated from use of the scale for this research study, and the study was exclusively designed to address scientific questions in the domains of epidemiology and public health.

The authors have no conflict of interests to declare with respect to this article.

References

1. Wing L, Gould J. Severe impairments of social interaction and associated abnormalities in children: epidemiology and classification. J Autism Dev Disord 1979;**9**:11–29.
2. Wing L. The continuum of autistic characteristics. In: Schopler E, Mesibov GB eds. Diagnosis and assessment in autism, New York: Plenum, 1988:91–110.
3. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 4th edn. Washington, DC: Text rev. American Psychiatric Association, 2000.
4. Constantino JN, Zhang Y, Frazier T, Abbacchi AA, Law P. Sibling recurrence and the genetic epidemiology of autism. Am J Psychiatry 2010;**167**:1349–1356.
5. Bailey A, Palferman S, Heavey L, Le Couteur A. Autism: the phenotype in relatives. J Autism Dev Disord 1998;**28**:369–392.
6. Piven J, Palmer P, Jacobi D, Childress D, Arndt S. Broader autism phenotype: evidence from a family history study of multiple-incidence autism families. Am J Psychiatry 1997;**154**:185–190.
7. Szatmari P, Maclean JE, Jones MB et al. The familial aggregation of the lesser variant in biological and nonbiological relatives of PDD probands: a family history study. J Child Psychol Psychiatry 2000;**41**:579–586.
8. Lord C, Petkova E, Hus V et al. A multisite study of the clinical diagnosis of different autism spectrum disorders. Arch Gen Psychiatry 2012;**69**:306–313.
9. Fombonne E, Simmons H, Ford T, Meltzer H, Goodman R. Prevalence of pervasive developmental disorders in the

British Nationwide Survey of Child Mental Health. J Am Acad Child Adolesc Psychiatry 2001;**40**:820–827.
10. Kim YS, Leventhal BL, Koh YJ et al. Prevalence of autism spectrum disorders in a total population sample. Am J Psychiatry 2011;**168**:904–912.
11. Constantino JN, Gruber CP. Social responsiveness scale (SRS). Los Angeles: Western Psychological Services, 2005.
12. Ministry of Education, Culture, Sports, Science and Technology. Annual report on school basic survey, 2010. Tokyo: Nikkei, 2010.
13. Koyama T, Osada H, Tsujii H, Kurita H. Utility of the Kyoto Scale of Psychological Development in cognitive assessment of children with pervasive developmental disorders. Psychiatry Clin Neurosci 2009;**63**:241–243.
14. Constantino JN, Przybeck T, Friesen D, Todd RD. Reciprocal social behavior in children with and without pervasive developmental disorders. J Dev Behav Pediatr 2000;**21**:2–11.
15. Constantino JN, Hudziak JJ, Todd RD. Deficits in reciprocal social behavior in male twins: evidence for a genetically independent domain of psychopathology. J Am Acad Child Adolesc Psychiatry 2003;**42**:458–467.
16. Constantino JN, Davis SA, Todd RD et al. Validation of a brief quantitative measure of autistic traits: comparison of the social responsiveness scale with the autism diagnostic interview-revised. J Autism Dev Disord 2003;**33**:427–433.
17. Constantino JN, Lavesser PD, Zhang Y, Abbacchi AM, Gray T, Todd RD. Rapid quantitative assessment of autistic social impairment by classroom teachers. J Am Acad of Child Adolesc Psychiatry 2007;**46**:1668–1676.
18. Constantino JN, Todd RD. Autistic traits in the general population: a twin study. Arch Gen Psychiatry 2003;**60**:524–530.
19. Constantino JN, Todd RD. Genetic structure of reciprocal social behavior. Am J Psychiatry 2000;**157**:2043–2045.
20. Constantino JN, Gruber CP, Davis S, Hayes S, Passanante N, Przybeck T. The factor structure of autistic traits. J Child Psychol Psychiatry 2004;**45**:719–726.
21. Ito H, Tani I, Yukihiro R et al. Validation of an interview-based rating scale developed in Japan for pervasive developmental disorders. Res Autism Spectr Disord 2012;**6**:1265–1272.
22. Kamio Y, Tsujii H, Inada N et al. Jiheisyo no cyosoukish-inndannhou oyobi mishinndannseijinnsyorei no kan-bennashinndannho no kaihatsu ni kansuru kenkyu [Screening for undiagnosed children and adults with autism spectrum disorders]. In: Okuyama M ed. Hattasushogaisya no atarashi shindan chiryoho no kaihatsu nikansuru kenkyu: Heisei 20 nendo sokatsubuntan hokokusyo [Annual research report on development of diagnosis and treatment for individuals with developmental disorders, supported by Health and Labour Sciences Research Grants]. Tokyo: National Center for Child Health and Development, 2009:25–35.
23. Kamio Y, Tsujii H, Inada N et al. Validation of the Japanese version of the social responsiveness scale: comparison with PDD-Autism Society Japan Rating Scales (PARS) [in Japanese]. Seishin Igaku 2009;**51**:1101–1109.
24. Bölte S, Poustka F, Constantino JN. Assessing autistic traits: cross-cultural validation of the Social Responsiveness Scale (SRS). Autism Res 2008;**1**:354–363.
25. Rutter M, Le Couteur A, Lord C. Autism diagnostic interview-revised (ADI-R). Los Angeles: Western Psychological Services, 2003.
26. Tsuchiya KJ, Matsumoto K, Yagi A et al. Reliability and validity of the Autism Diagnostic Interview-Revised-Japanese Version. J Autism Dev Disord 2013;**43**:643–662.
27. Moriwaki A, Koyama T, Kamio Y. Nihongoban SRS no hyojyunka [Standardization of the Japanese version of the Social Responsiveness Scale]. In: Kamio Y ed. Annual report of research supported by health and labour sciences research grants [in Japanese], Tokyo: National Center of Neurology and Psychiatry, 2011:49–68.
28. Constantino JN. The quantitative nature of autistic social impairment. Pediatr Res 2011;**69**:55R–62R.
29. Lund J, Jensen J. Dimensional approach to infantile autism in mentally retarded adults. Acta Psychiatr Scand 1989;**80**:389–394.
30. Mandy W, Charman T, Gilmour J, Skuse D. Toward specifying pervasive developmental disorder-not otherwise specified. Autism Res 2011;**4**:121–131.
31. Virkud YM, Todd RD, Abbacchi AM, Zhang Y, Constantino JN. Familial aggregation of quantitative autistic traits in multiplex versus simplex autism. Am J Med Genet B Neruropsychiatr Genet 2009;**150B**:328–334.
32. Zappella M. Reversible autism and intellectual disability in children. Am J Med Genet Part C 2012;**160C**:111–117.
33. Robinson EB, Koenen KC, Mccormick MC et al. Evidence that autistic traits show the same etiology in the general population and at the quantitative extremes (5%, 2.5%, and 1%). Arch Gen Psychiatry 2011;**68**:1113–1121.
34. Ho A, Todd RD, Constantino JN. Brief report: autistic traits in twins vs. non-twins-a preliminary study. J Autism Dev Disord 2005;**35**:129–133.
35. Chen C, Lee S, Stevenson HW. Response style and cross-cultural comparisons of rating scales among East Asian and North American students. Psychol Sci 1995;**6**:170–175.
36. Reiersen AM, Constantino JN, Volk HE, Todd RD. Autistic traits in a population-based ADHD twin sample. J Child Psychol Psychiatry 2007;**48**:464–472.
37. Lichtenstein P, Carlström E, Råstam M, Gillberg C, Anckarsäter H. The genetics of autism spectrum disorders and related neuropsychiatric disorders in childhood. Am J Psychiatry 2010;**167**:1357–1363.

Hindawi Publishing Corporation
Autism Research and Treatment
Volume 2013, Article ID 373240, 9 pages
http://dx.doi.org/10.1155/2013/373240

Hindawi

Clinical Study

Utility of Teacher-Report Assessments of Autistic Severity in Japanese School Children

Yoko Kamio, Aiko Moriwaki, and Naoko Inada

Department of Child and Adolescent Mental Health, National Institute of Mental Health, National Center of Neurology and Psychiatry, 4-1-1 Ogawa-Higashi, Kodaira, Tokyo 187-8553, Japan

Correspondence should be addressed to Yoko Kamio; kamio@ncnp.go.jp

Received 2 June 2013; Revised 28 October 2013; Accepted 11 November 2013

Academic Editor: Herbert Roeyers

Recent studies suggest that many children with milder autism spectrum disorder (ASD) are undiagnosed, untreated, and being educated in mainstream classes without support and that school teachers might be the best persons to identify a child's social deviance. At present, only a few screening measures using teacher ratings of ASD have been validated. The aim of this study was to examine the utility of teacher ratings on the Social Responsiveness Scale (SRS), a quantitative measure of ASD. We recruited 130 participants aged 4 to 17 years from local schools or local pediatric outpatient clinics specializing in neurodevelopmental disorders that included 70 children with ASD. We found that the teacher-report SRS can be reliably and validly applied to children as a screening tool or for other research purposes, and it also has cross-cultural comparability. Although parent-teacher agreement was satisfactory overall, a discrepancy existed for children with ASD, especially for girls with ASD. To improve sensitivity in children at higher risk, especially girls, we cannot overstate the importance of using standardized norms specific to gender, informant, and culture.

1. Introduction

The current professional consensus is that early diagnosis and subsequent early treatment of autism spectrum disorder (ASD) can facilitate development and learning [1, 2], reduce the need for treatment later in life [3, 4], and improve longterm prognosis in adulthood [5, 6]. However, not all families with children with ASD necessarily get timely access to treatment and other support. Delayed identification and diagnosis of ASD have been associated with subtypes of ASD [7–10], cognitive level [10, 11], gender [11, 12], and demographic factors such as low socioeconomic status [8, 10]. Diagnosis of ASD tends to be delayed in children having both milder autistic symptoms and above-average general cognitive ability, especially in girls. For example, reported age at first diagnosis of Asperger, syndrome ranges from 7 to 11 years [9, 10, 12, 13]. In a Japanese nationwide survey of adults with high-functioning ASD, the median age at first diagnosis was 10.3 years [6].

Recent epidemiological studies [14, 15] have revealed that most mainstreamed children with ASD were undiagnosed and untreated. Although most of these children might have had few diagnosable symptoms during preschool to draw the attention of primary health professionals, school teachers should be the best persons to identify any overt social deviance [16, 17].

At present, many quantitative behavioral measures of ASD have been created and validated in both primary care and clinical settings. However, these measures were largely validated for use by parents, not teachers, except in the case of the Autism Spectrum Screening Questionnaire (ASSQ), the Social Communication Questionnaire (SCQ), and the Social Responsiveness Scale (SRS). The ASSQ is a 27-item questionnaire that was originally developed as a first-stage population screening instrument in a prevalence study of Asperger, syndrome in mainstream schools with teachers as target raters [18], and it has been validated as a general population screen [19, 20]. The reliability and validity of both parent and teacher ASSQ ratings in a clinical setting have also been reported, although parent-teacher agreement was low to moderate for children with high-functioning ASD [21]. The SCQ [22] is a 40-item screening instrument that has been

investigated mainly as a parent-report screen. In one study of children with ASD and their siblings [17], the teacher-report SCQ-Current version was moderately correlated with the parent-report SCQ-Lifetime version, whereas it was strongly correlated with the teacher-report SRS. The SRS was developed as a quantitative measure of autistic traits in children [23], and the parent-report SRS has been extensively validated for the general child population [24–27] as well as for clinical samples [24, 28–32] not only in the USA but also in Europe, South America, and Asia. On the other hand, the literature on the utility of the SRS as a screening tool assessed by teachers is still limited [17, 31, 33]. Constantino et al. [34] demonstrated that the teacher-report SRS exhibited strong correlations with the parent-report SRS ($r = 0.72$), and the combined use of both parent and teacher reports resulted in extremely high sensitivity and specificity for a diagnosis of ASD in 271 children with ASD and 171 children without ASD, including 52 child psychiatric patients and 119 unaffected siblings. Schandling et al. [17] examined the utility of parent- and teacher-report SCQ and SRS in 1,663 children with ASD and 1,712 unaffected siblings from 1,655 families and showed that the screening properties of the teacher-report SRS were superior to those of the teacher-report SCQ-Current. In their study, the teacher-report SRS was more congruent with clinician-observed behaviors than with parent-reported behaviors and raised the possibility that behaviors exhibited by the children with ASD are contextually related and might be more congruent across classroom and clinical settings [17]. Fombonne et al. [31] examined the psychometric properties of the SRS-Spanish version in a Mexican sample consisting of 140 children with ASD and 319 community controls and found that the teacher-report SRS was an excellent screening tool similar to the parent-report SRS. In addition, they noted that the parent-teacher correlation of the SRS was much higher in the ASD sample compared with the control group.

Although some evidence exists on the SRS as a screening tool assessed by teachers, its utility has not been examined in an Asian population. Further, the reason for the discrepancy between parent and teacher reports on this scale is unclear.

Thus, the main aim of this study was to examine the utility of the SRS-Japanese version as a teacher-report screening tool for ASD. To this end, we examined test-retest reliability and discriminant/convergent validity of the teacher-report SRS, parent-teacher correlations or discrepancies on the SRS, and screening cutoffs in Japanese children aged 4 to 17 years.

2. Materials and Methods

2.1. Participants. This study involved 130 children consisting of 70 children with ASD (51 boys, mean age 8.6 [3.7], range 4–17 years) and 60 children without ASD (39 boys, mean age 8.0 [2.5], range 5–15 years; 24 with any neuropsychiatric diagnosis other than ASD; and 36 typically developing [TD] children). Seventy-eight children (23 with ASD, 19 with any neuropsychiatric diagnosis, and 36 with TD) currently participated in our ongoing community-based longitudinal study of child mental health at the National Center of Neurology and Psychiatry (NCNP), Japan. All research participants were attending mainstream classes at local schools. We also recruited 20 children from a local special school for children with learning disabilities (15 with mental retardation [MR] and ASD, 5 with MR only). In addition, we recruited 32 patients diagnosed with ASD from three local pediatric outpatient clinics specializing in neurodevelopmental disorders.

The gender ratio did not significantly differ between children with ASD and those without ASD ($\chi^2 = 0.94$, ns). Mean age did not significantly differ between groups ($t = 1.16$, ns).

2.2. Measures

2.2.1. The Social Responsiveness Scale. The SRS is a 65-item questionnaire of autistic traits for use with 4–18-year-olds that can be completed in 15 minutes by parents or teachers who have observed the child over time in naturalistic social settings [23]. The SRS was developed to assess autistic symptoms or quantitative traits and has subsequently undergone extensive validation in general and clinical child populations in the USA and other countries. The 65 SRS items can be categorized into five subscales (social awareness, social cognition, social communication, social motivation, and autistic mannerisms). Each item is scored on a 4-point scale, and total score ranges from 0 to 195, with higher scores indicating higher degrees of social impairment. We used the teacher version in the present study and also the parent version as a subsample. The Japanese version of the parent SRS exhibited a skewed normal distribution in the general population with a single-factor structure, had no relation to IQ within the normal intellectual range [27], and demonstrated satisfactory discriminant and convergent validity [27, 35]. Both the parent- and teacher-report SRS were standardized on boys and girls separately [36].

2.2.2. The Autism Diagnostic Interview-Revised (ADI-R). The Autism Diagnostic Interview-Revised (ADI-R) [37] is a parent-report interview and a research standard for establishing a diagnosis of autism. The algorithm generates scores in each of three domains: reciprocal social interaction; communication; and restricted, repetitive, and stereotyped patterns of behavior. We used total scores of three domains of the Japanese version of the ADI-R [38] for the analysis in this study.

2.2.3. The Autism Diagnostic Observation Schedule (ADOS). The ADOS [39] is a semistructured behavioral assessment of social interaction, communication, and stereotyped behaviors. The algorithm generates scores in each of the three domains. We used total scores of the social and communication domains of the Japanese version of the ADOS [40] for the analysis in this study.

2.3. Procedure. The study protocol was approved by the Ethics Committee of the NCNP. A written informed consent was obtained from the parents of each child participant, and the study was conducted from 2010 to 2012.

First, parents were informed about the study by a letter from the investigators, which was distributed by the investigators themselves, a principal teacher, child psychiatrist, or pediatrician. Second, after providing the written consent, parents asked classroom teachers to complete the SRS on their children. Among all returned questionnaires, we excluded 16 teacher reports (11.0%) that had one or more missing answers, leaving 130 teacher reports on 130 children. Among these, we obtained 109 parent reports on 109 children (57 with ASD, 52 without ASD [19 clinical, 33 TD]).

Our research team conducted diagnostic interviews at the NCNP for 78 children, at the special school for 20 children, and at clinics for 32 children.

ASD diagnoses were confirmed according to DSM-IV-TR criteria based on all available clinical information by our research team that included experienced child psychiatrists and licensed clinical psychologists. To corroborate each ASD diagnosis, we evaluated the severity of autistic symptoms using either the Japanese versions of the Autism Diagnostic Interview-Revised (ADI-R) [38], the Autism Diagnostic Observation Schedule (ADOS) [40], the Diagnostic Interview for Social and Communication Disorders [41], or other semistructured interviews developed and validated in Japan [42]. Among 70 children with ASD, 55 were subcategorized with 100% diagnostic agreement based on available information among our research team: 24 with autistic disorder, 10 with Asperger's disorder, and 21 with pervasive developmental disorder, not otherwise specified. For 15 children, we reached complete agreement on a diagnosis of ASD, although we could not reach agreement on the subcategory.

The non-ASD diagnoses of 24 children were attention deficit hyperactivity disorder (ADHD), oppositional defiant disorder, specific phobia, social phobia, obsessive-compulsive disorder, enuresis, tic disorder, or mental retardation. These diagnoses were confirmed by diagnostic interviews with children and their parents using the Kiddie Schedule for Affective Disorder and Schizophrenia Present and Lifetime (K-SADS-PL), Japanese version. By parent interview, we confirmed the typical development of 36 children as having no history of neurological or psychiatric disorders.

We judged intellectual level based on cognitive testing (i.e., various versions of the Wechsler Intelligence Scale or other measures) for 115 children and educational/administrative records for 15 children. Intellectual level ranged from normal intelligence to severe MR (normal to borderline 105, mild MR 8, moderate MR 6, severe MR 4, and unknown MR 7). The proportion of children with normal intelligence did not significantly differ between children with ASD (53/70) and those without ASD (52/60) ($\chi^2 = 2.5$, ns).

2.4. Data Analysis. To address discriminant validity, we compared mean total and mean subscale SRS scores by gender between children with ASD ($n = 70$) and those without ASD ($n = 60$). To examine test-retest reliability, we calculated the intraclass correlation coefficient (ICC) for a subsample ($n = 23$). To examine convergent validity, we computed Pearson's correlation coefficients between the SRS and ADI-R, ADOS, or full scale IQ scores on three subsamples ($n = 49, 56, 115$). To examine the teacher-parent discrepancy, we calculated ICC and compared mean total and mean subscale SRS scores by group (ASD versus non-ASD) and by gender using a paired t-test on a subsample ($n = 109$) that included both teacher and parent ratings. Finally, we conducted a receiver operating characteristics (ROC) analysis to compare the area under the curve (AUC) for the parent- and teacher-report SRS for a subsample ($n = 109$), and determined the cutoff scores that maximized sensitivity and specificity for the teacher-report SRS for the total sample.

All analysis was performed using SPSS 18.0J for Windows.

3. Results

3.1. Discriminant Validity. Table 1 presents the mean raw teacher-report SRS scores for the total sample ($N = 130$; ASD 70, non-ASD 60 [non-ASD diagnosis 24, TD 36]) by gender. Total scores and the five subscale scores were significantly higher in children with ASD than in those without ASD for both genders, except for social awareness and social motivation subscales in girls, where the mean subscale scores did not significantly differ between girls with ASD and those without ASD.

3.2. Test-Retest Reliability. Among 130 children, 23 (ASD 1, non-ASD diagnosis 3, TD 19) were assessed by their classroom teachers on two occasions with a mean interval of 40.0 days (12–131 days). Test-retest reliability was shown to be excellent for the total score (time 1: mean SRS 63.2 [22–103]; time 2: mean SRS 61.4 [24–119]; ICC = 0.87; $P < 0.001$).

3.3. Convergent Validity. SRS total score was significantly positively correlated with ADI-R total score ($r = 0.30$, $P < 0.05$) in a subsample with available data from both the SRS and ADI-R ($n = 56$, ASD 21, non-ASD diagnosis 14, TD 21; 36 boys) and also significantly correlated with ADOS total score ($r = 0.30$, $P < 0.05$) in a subsample with available data from both the SRS and ADOS ($n = 49$, ASD 20, non-ASD diagnosis 16, TD 13; 35 boys). In 115 children with available IQ data, the SRS score did not significantly correlate with IQ ($r = -0.14$) for 97 children with IQs ≥ 70 (ASD 46, non-ASD diagnosis 16, TD 35; 66 boys), whereas it significantly correlated with IQ in 18 children with IQs < 70 (ASD 11, non-ASD diagnosis 7) ($r = -0.58$, $P < 0.05$).

3.4. Parent-Teacher Correlation and Discrepancy. Among 130 total participants, 109 participants (ASD 57, non-ASD 52, including non-ASD diagnosis 19 and TD 33) were rated by both their teachers and parents at almost the same time (Table 2). For this subsample (73 boys, mean age 10.9 [2.8], range 7–14 years), ICCs showed moderate to large agreement between teachers and parents for all 109 children (73 boys and 36 girls; ICCs = 0.48, 0.50, and 0.40, resp.; Table 3). Among five subscales, ICCs ranged from moderate to large (ICCs = 0.29–0.53, P values < 0.05), except for the social awareness subscale in girls (ICC = 0.08, ns).

Table 4 shows that children with ASD of either gender were rated significantly higher by parents than by teachers on the total scores. Among five subscales, significant differences in ratings between parents and teachers were found only for autistic mannerisms in boys with ASD, whereas subscale ratings on social cognition, social communication, and autistic mannerisms were significantly different in girls with ASD. On the other hand, children without ASD of either gender were rated similarly by parents and teachers on the total scale and on all subscales.

For children with ASD, we found a significant gender difference in teacher ratings on the SRS only on the social awareness subscale, where teachers rated girls significantly lower than boys ($t = 2.10$, $P < 0.05$). By contrast, we found no significant gender differences in parent ratings for this sample. On the other hand, for children without ASD, we observed no significant gender differences in both parent and teacher ratings (Table 4). That is, the gender difference was strongest in teacher reports on social awareness in the ASD group. Thus, teachers tended to rate boys and girls with ASD lower compared to parents, and teachers tended to rate girls with ASD lower compared to boys with ASD.

3.5. ASD Cutoff Scores. ROC analyses of 109 children who were rated by both parents and teachers informed the AUC for each parent and teacher report on the SRS; among this sample, the teacher-report SRS accurately classified 73.2% of boys ($P < 0.005$) and 70.8% of girls ($P < 0.05$), whereas the parent-report SRS accurately classified 90.0% of boys ($P < 0.005$) and 94.8% of girls ($P < 0.005$) (Figures 1(a) and 1(b)). Therefore, the parent-report SRS appears to be more accurate than the teacher-report SRS as a screening tool. For the total sample, Youden's index was computed to determine the cutoff points that maximized the sum of sensitivity and specificity of the teacher-report SRS, 58.0 for boys (sensitivity 0.725, specificity 0.667, false-negative rate 0.275, false-positive rate 0.333, and positive likelihood ratio 2.177) and 43.0 for girls (sensitivity 0.789, specificity 0.667, false-negative rate 0.211,

Table 1: Mean raw SRS scores of teacher ratings of children with ASD and without ASD ($N = 130$).

Subscale	Boys ($n = 90$)			Girls ($n = 40$)		
	ASD ($n = 51$) M (SD)	Non-ASD ($n = 39$) M (SD)	t	ASD ($n = 19$) M (SD)	Non-ASD ($n = 21$) M (SD)	t
Awareness	11.6 (0.5)	7.4 (3.6)	4.8[a]	8.7 (3.9)	7.1 (4.3)	1.3[d]
Cognition	14.8 (6.2)	9.3 (6.0)	4.2[a]	12.6 (4.9)	7.3 (5.3)	3.3[b]
Communication	27.3 (10.2)	16.8 (10.0)	4.9[a]	24.7 (12.7)	13.1 (10.7)	3.1[b]
Motivation	11.0 (5.2)	8.4 (5.7)	2.2[c]	10.8 (5.2)	8.4 (6.0)	1.4[d]
Mannerisms	13.9 (8.5)	7.2 (5.4)	4.5[a]	11.3 (8.4)	5.5 (6.4)	2.5[c]
Total	78.6 (29.9)	49.1 (26.9)	4.8[a]	68.2 (28.8)	41.5 (30.2)	2.9[a]

Note. SRS: Social Responsiveness Scale; ASD: autism spectrum disorder.
[a] $P < 0.001$. [b] $P < 0.01$. [c] $P < 0.05$. [d] ns.

Table 2: Demographic characteristics of 109 children rated by both teacher and parent.

	ASD ($n = 57$)	Non-ASD ($n = 52$)	
		Neuropsychiatric diagnosis ($n = 19$)	TD ($n = 33$)
Boy : girl	41 : 16	12 : 7	20 : 13
Age (years)			
Mean (SD), range	8.60 (3.90), 4–17	8.26 (2.77) 5–15	7.67 (2.13) 5–12
Intellectual level (n)			
Normal	34	9	33
Borderline	14	5	0
Mild MR	3	4	0
Moderate MR	2	0	0
Severe MR	2	1	0
MR (unknown level)	2	0	0
IQ*	$n = 49$	$n = 19$	$n = 32$
Mean (SD), range	91.2 (26.8), 31–148	82.7 (23.3), 27–113	109.7 (13.8), 85–146

Note. Between the ASD and non-ASD groups, no significant differences existed in gender ratio ($\chi^2 = 0.25$, ns) or age ($t = 1.2$, ns). The proportion of intellectual level did not differ significantly by group ($\chi^2 = 9.4$, ns). For 100 children with available IQ data, mean IQ did not significantly differ between groups (91.2 [26.8] for ASD, 99.7 [22.0] for non-ASD). Among the ASD and two non-ASD groups, no significant differences existed in gender ratio ($\chi^2 = 0.51$, ns) or age ($F = 0.84$, ns). The proportion of intellectual level differed significantly by group ($\chi^2 = 28.5$, $P < 0.005$). *For 100 children with available IQ data, mean IQ of the ASD group ($n = 49$) and that of the non-ASD neuropsychiatric diagnosis group ($n = 19$) were lower than that of the TD group ($n = 32$) ($t = 4.1, 4.6$, respectively, P values < 0.001), whereas no significant difference existed between the former two groups ($t = 1.2$, ns). MR: mental retardation; ASD: autism spectrum disorder; TD: typically developing.

Table 3: Intraclass correlation coefficients ($N = 109$).

Teacher rating	Parent rating					
	Awareness	Cognition	Communication	Motivation	Mannerisms	Total
Awareness	Total 0.38[a] Boys 0.50[a] Girls 0.08[d]					
Cognition		Total 0.45[a] Boys 0.46[a] Girls 0.41[b]				
Communication			Total 0.45[a] Boys 0.45[a] Girls 0.38[c]			
Motivation				Total 0.48[a] Boys 0.47[a] Girls 0.53[a]		
Mannerisms					Total 0.38[a] Boys 0.38[a] Girls 0.29[c]	
Total						Total 0.48[a] Boys 0.50[a] Girls 0.40[b]

Note. This subsample ($N = 109$) comprises 57 children with ASD and 52 children without ASD.
[a]$P < 0.001$. [b]$P < 0.01$. [c]$P < 0.05$.[d] ns.

Table 4: Mean raw SRS scores of parent and teacher ratings of children with ASD and without ASD ($N = 109$).

Rater	Boys ($n = 73$)			Girls ($n = 36$)		
	Parent Mean (SD)	Teacher Mean (SD)	P	Parent Mean (SD)	Teacher Mean (SD)	P
ASD ($n = 57$)		Boys ($n = 41$)			Girls ($n = 16$)	
Awareness	11.9 (3.4)	11.2 (4.3)	ns	10.2 (2.6)	8.6 (4.0)	ns
Cognition	16.2 (6.4)	14.2 (6.0)	ns	16.9 (4.8)	11.6 (4.7)	<0.005
Communication	30.4 (11.3)	26.6 (9.8)	=0.06	30.6 (9.2)	22.4 (12.0)	<0.05
Motivation	12.3 (5.8)	10.6 (5.3)	ns	12.5 (5.4)	10.8 (5.6)	ns
Mannerisms	16.4 (7.8)	12.7 (8.4)	<0.05	14.9 (7.4)	8.9 (6.5)	<0.05
Total	87.2 (31.3)	75.3 (29.2)	<0.05	85.1 (25.3)	62.3 (27.1)	<0.01
Non-ASD ($n = 52$)		Boys ($n = 32$)			Girls ($n = 20$)	
Awareness	6.6 (5.1)	7.4 (3.4)	ns	6.6 (3.3)	7.0 (4.4)	ns
Cognition	8.3 (4.4)	9.5 (6.0)	ns	6.6 (3.9)	7.5 (5.3)	ns
Communication	13.6 (7.0)	17.3 (10.2)	ns	10.8 (6.5)	13.7 (10.6)	ns
Motivation	7.3 (3.7)	8.7 (5.8)	ns	6.3 (4.9)	8.5 (6.1)	ns
Mannerisms	6.2 (4.5)	7.5 (5.5)	ns	4.2 (4.0)	5.8 (6.5)	ns
Total	42.0 (18.7)	50.5 (27.6)	ns	34.3 (19.9)	42.5 (30.6)	ns

Note. SRS: Social Responsiveness Scale; ASD: autism spectrum disorder.

false-positive rate 0.333, and positive likelihood ratio 2.369). These optimal cutoff scores were found to correspond to a T-score of 60 for each boy and girl according to T-score norms that were created for the Japanese standardization sample [36]. Because no natural cutoff was found that differentiated children diagnosed with ASD from those without ASD in the Japanese general and clinical samples for the parent-report SRS [27], these cutoff scores of teacher-report SRS would identify many subthreshold conditions and at the same time miss many true-positive children. Compared to the previously reported optimal cutoff scores on the parent-report SRS (boys, sensitivity 0.91, specificity 0.48, and positive likelihood ratio 1.75; girls, sensitivity 0.89, specificity 0.41, and positive likelihood ratio 1.51) [27], the optimal cutoff scores on the teacher-report SRS would seem to result in a higher false-negative rate (boy, 0.28 versus 0.09, girl, 0.21 versus 0.11,

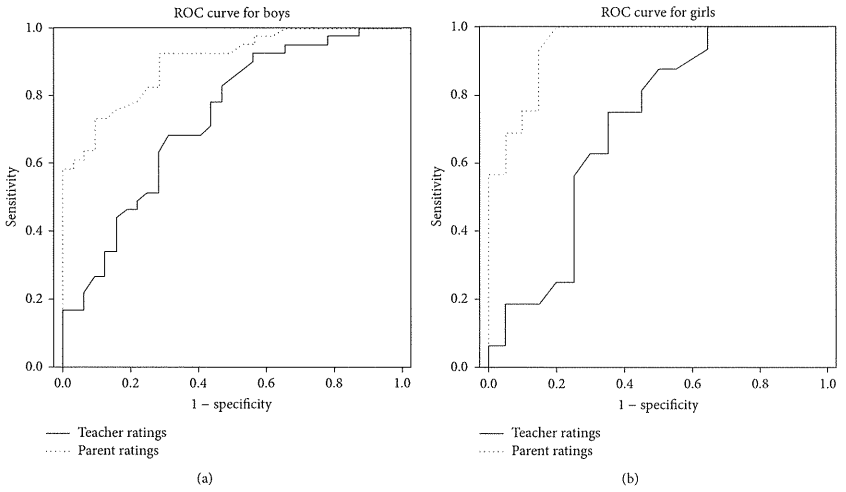

Figure 1: (a) Receiver operating characteristics (ROC) curve demonstrating sensitivity and specificity of both teacher and parent ratings for boys ($n = 73$). (b) Receiver operating characteristics (ROC) curve demonstrating sensitivity and specificity of both teacher and parent ratings for girls ($n = 36$).

teacher, and parent, resp.) and a lower false-positive rate (boy, 0.33 versus 0.52, girl, 0.33 versus 0.59, teacher, and parent, resp.). As addressed by Constantino et al. [34], when both parent and teacher rate a child as having a T-score of ≥60, the positive likelihood ratio would improve up to 3.730 in our sample, which exceeds the teacher-report SRS alone or the parent-report SRS alone.

4. Discussion

The main aim of this study was to examine the utility of the teacher-report SRS as an ASD screening tool for Japanese children. In this study, the teacher-report SRS demonstrated excellent test-retest reliability and satisfactory discriminant and convergent validity for measuring autistic severity in children aged 4 to 17 years. Overall, there were moderate to large parent-teacher correlations on the total and subscale ratings. Thus, our findings showed that the teacher ratings on the Japanese version of the SRS can be reliably and validly applied to Japanese children at school or in clinical settings as a screening tool of ASD clinical settings.

Our results suggest overall good agreement on SRS measurements in terms of severity of autistic symptoms between teachers and parents; the correlations fall within the range reported in previous studies for the SRS (0.24–0.82) [17, 29–31, 33, 34, 43]. Our result is satisfactory compared to other psychiatric domains [43]. However, it is difficult to compare ours with the correlations reported by previous studies because of differences in sample size (26–3375), the proportion of children with ASD included in the total sample (0–69.5%), and how control children were sampled (siblings from families who registered participation in autism research, community schools, and clinical non-ASD psychiatric patients); there appears to be no systematic tendency explaining the wide variation. For example, in Fombonne et al. [31], parent-teacher correlations for total SRS were stronger in children with ASD than in control children, but the opposite was found in Kanne et al. [43]. Based on data from Japan, the correlation for the non-ASD sample (Pearson's $r = 0.78$) [35] decreased to an ICC of 0.48 when calculated for the sample that included children with ASD (52.3%) in this study. Further studies are needed to answer this issue.

Despite overall good agreement, teachers tended to rate both boys and girls with ASD lower than did parents, although the teacher-parent discrepancy was not pronounced in children without ASD. Such discrepancy relating to the type of children (ASD versus non-ASD) was consistently found in previous studies [17, 34, 43, 44] except in a study based on a Mexican sample [31]. In the present study, teacher-parent discrepancy was pronounced, especially for girls with ASD (teacher 62.3 versus parent 85.1); parent ratings were significantly higher than teacher ratings not only on the total score but also on 3 (social cognition, social communication, and autistic mannerisms) of 5 subscales. One possible interpretation could be an effect of situational context as suggested by Szatmari et al. [45] and Posserud et al. [19]. How children with ASD behave can change depending

on the situation, such as the degree of structurization, and it is likely that autistic behaviors of higher-functioning children with ASD are observed less often at school than at home if the school environment meets a child's needs. Shanding et al. [17] raised the possibility that teachers and clinicians similarly observe and report behaviors exhibited by children with ASD based on the stronger association between teacher ratings on the SRS and the ADOS compared to that between the teacher SRS and the ADI-R. Szatmari et al. [45] warned that this discrepancy between home and school might lead to higher parental stress. Thus, we should exercise caution when interpreting information from parents and teachers in diagnosis and assessment.

Regarding gender differences, it appears that teachers tend to rate girls with ASD lower than boys with ASD, whereas they rate girls without ASD higher than boys without ASD, although these differences reached statistical significance only on the social awareness subscale of the teacher report. Similar gender differences were reported in Norway for total population data using the ASSQ [19]. By contrast, in a Mexican sample [31], affected girls scored higher than affected boys on the teacher-report SRS, whereas control boys scored higher than control girls. However, closer inspection revealed a similar gender difference related to the social awareness subscale between ours and Fombonne et al. [31]. In both studies, teacher and parent ratings for girls did not agree on this subscale, and gender differences in teacher ratings were pronounced on this subscale. In this study, teacher ratings on this subscale also did not discriminate girls with ASD from those without ASD. The poor reliability and validity of this subscale might be related to the measurement of social awareness, which is not overt and is difficult to observe from the outside. Lai et al. [46] reported that women with ASD showed fewer autistic features than males but perceived themselves as having more autistic features, perhaps because they are better at hiding their autistic features, or perhaps because of greater self-awareness. Our finding of gender differences, if replicated, emphasizes the need for both a deeper understanding of gender differences in ASD and the establishment of a gender-specific norm.

The ROC analysis demonstrated that teacher ratings on the SRS classified both boys and girls with moderate accuracy, although the parent-report SRS appears to be more accurate than the teacher-report SRS as a screening tool. The optimal cutoff for boys was 58.0 in this study, which fell within the range of 51.5 to 64.5 proposed in previous studies of the teacher-report SRS [17, 31, 34, 44], whereas that for girls was 43.0 in our sample, which fell below the range. If this great discrepancy in cutoff scores between boys and girls is replicated in a different Japanese sample, the importance of establishing gender-specific norms for this population should be emphasized again. In this study, either sensitivity or specificity values were lower compared to those in studies that included only children with ASD and typically developing children [17, 31], which is consistent with studies that included children with non-ASD clinical conditions [34, 44]. Children with non-ASD psychiatric diagnoses such as ADHD or mood disorders tended to have high SRS scores [47, 48], and there is an overlap in SRS scores of children with ASD and those of children with non-ASD psychiatric diagnoses [27]. That is, the sensitivity or specificity values in our sample might be associated with the type of non-ASD controls, including children with non-ASD psychiatric diagnoses whose mean SRS scores were expected to be higher than those of the normative sample.

Regarding cultural differences in teacher ratings of children with ASD, our female sample with ASD scored similar to children with ASD (86.5% male) in a large-sized study by Schanding et al. [17], whereas our male sample with ASD scored higher. However, our sample with ASD of either gender scored lower than children with ASD in other studies [31, 34]. This variance might be partly explained by the sampling method rather than culture-related differences, taking the heterogeneity of ASD into account. As for gender differences found in this study, little evidence exists, except that in a Mexican sample [31], to draw any conclusion about it from a cultural perspective. If our findings on gender differences are replicated in samples representing different cultures, we should consider culture-free gender differences. Or, if our findings are limited to a Japanese sample, we should consider any cultural factor such as social expectations of the female role in public settings, especially in terms of social awareness. Again, cross-cultural validity of the teacher-report SRS would be guaranteed if it is applied to children according to culturally calibrated gender-specific norms.

The major limitation of this study is its small sample size. Further, we did not use the same assessment battery to determine ASD status for children diagnosed with ASD. Comorbid psychiatric disorders were not assessed using diagnostic measures for 23 children with ASD. The strength of this study is that ASD was excluded for all of the non-ASD children.

5. Conclusions

In conclusion, this study provided evidence that the teacher-report SRS is a useful measurement of autistic severity with good reliability and validity and recapitulated what has been observed in studies conducted in other countries. Although parent-teacher agreement on the SRS was satisfactory, characteristic discrepancies specific to ASD diagnostic status and gender between informants should be kept in mind when interpreting the SRS from only one-sided informants. To improve sensitivity for children who are at higher risk, especially girls who are likely to remain undiagnosed, we emphasize the importance of combining information from multiple informants and using standardized norms specific to gender, informant, and culture for screening, clinical, or research purposes.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This study was supported by research grants from the Ministry of Health, Labour, and Welfare of Japan to Dr. Yoko Kamio (H20-KOKORO-004 and ID11103316). The authors would like thank Drs. Tokio Uchiyama, Miho Kuroda, and Hiroshi Fujino for data collection.

References

[1] R. J. Landa and L. G. Kalb, "Long-term outcomes of toddlers with autism spectrum disorders exposed to short-term intervention," *Pediatrics*, vol. 130, supplement 2, pp. S186–S190, 2012.

[2] S. J. Rogers, A. Estes, C. Lord et al., "Effects of a brief Early Start Denver Model (ESDM)-based parent intervention on toddlers at risk for autism spectrum disorders: a randomized controlled trial," *Journal of American Academy of Child and Adolescent Psychiatry*, vol. 51, no. 10, pp. 1052–1065, 2012.

[3] G. Dawson, S. Rogers, J. Munson et al., "Randomized, controlled trial of an intervention for toddlers with autism: the early start Denver model," *Pediatrics*, vol. 125, no. 1, pp. e17–e23, 2008.

[4] D. Fein, M. Barton, I.-M. Eigsti et al., "Optimal outcome in individuals with a history of autism," *Journal of Child Psychology and Psychiatry*, vol. 54, no. 2, pp. 195–205, 2013.

[5] E. Fernell, M. A. Eriksson, and C. Gillberg, "Early diagnosis of autism and impact on prognosis: a narrative review," *Clinical Epidemiology*, vol. 5, pp. 33–43, 2013.

[6] Y. Kamio, N. Inada, and T. Koyama, "A nationwide survey on quality of life and associated factors of adults with high-functioning autism spectrum disorders," *Autism*, vol. 17, no. 1, pp. 15–26, 2013.

[7] S. Chakrabarti and E. Fombonne, "Pervasive developmental disorders in preschool children: confirmation of high prevalence," *American Journal of Psychiatry*, vol. 162, no. 6, pp. 1133–1141, 2005.

[8] H. Coo, H. Ouellette-Kuntz, M. Lam et al., "Correlates of age at diagnosis of autism spectrum disorders in six Canadian regions," *Chronic Diseases and Injuries in Canada*, vol. 32, no. 2, pp. 90–100, 2012.

[9] P. Howlin and A. Asgharian, "The diagnosis of autism and Asperger syndrome: findings from a survey of 770 families," *Developmental Medicine and Child Neurology*, vol. 41, no. 12, pp. 834–839, 1999.

[10] D. S. Mandell, M. M. Novak, and C. D. Zubritsky, "Factors associated with age of diagnosis among children with autism spectrum disorders," *Pediatrics*, vol. 116, no. 6, pp. 1480–1486, 2005.

[11] P. T. Shattuck, M. Durkin, M. Maenner et al., "Timing of identification among children with an autism spectrum disorder: findings from a population-based surveillance study," *Journal of the American Academy of Child and Adolescent Psychiatry*, vol. 48, no. 5, pp. 474–483, 2009.

[12] S. Begeer, D. Mandell, B. Wijinker-Holmes et al., "Sex differences in the timing of identification among children and adults with autism spectrum disorders," *Journal of Autism and Developmental Disorders*, vol. 43, no. 5, pp. 1151–1156, 2013.

[13] C. Tandir and N. M. Mukaddes, "Referral pattern and special interests in children and adolescents with asperger syndrome: a Turkish referred sample," *Autism*, 2012.

[14] S. Baron-Cohen, F. J. Scott, C. Allison et al., "Prevalence of autism-spectrum conditions: UK school-based population study," *British Journal of Psychiatry*, vol. 194, no. 6, pp. 500–509, 2009.

[15] Y. S. Kim, B. L. Leventhal, Y. Koh et al., "Prevalence of autism spectrum disorders in a total population sample," *American Journal of Psychiatry*, vol. 168, no. 9, pp. 904–912, 2011.

[16] B. A. Pringle, L. J. Colpe, S. J. Blumberg, R. M. Avila, and M. D. Kogan, "Diagnostic history and treatment of school-aged children with autism spectrum disorder and special health care needs," NCHS Data Brief 97, National Center for Health Statistics, Hyattsville, Md, USA, 2012.

[17] G. T. Schanding Jr., K. P. Nowell, and R. P. Goin-Kochel, "Utility of the Social Communication Questionnaire-Current and Social Responsiveness Scale as teacher-report screening tools for autism spectrum disorders," *Journal of Autism and Developmental Disorders*, vol. 42, no. 8, pp. 1705–1716, 2012.

[18] S. Ehlers and C. Gillberg, "The epidemiology of Asperger syndrome. A total population study," *Journal of Child Psychology and Psychiatry*, vol. 34, no. 8, pp. 1327–1350, 1993.

[19] M. Posserud, A. J. Lundervold, and C. Gillberg, "Autistic features in a total population of 7–9-year-old children assessed by the ASSQ (Autism Spectrum Screening Questionnaire)," *Journal of Child Psychology and Psychiatry and Allied Disciplines*, vol. 47, no. 2, pp. 167–175, 2006.

[20] M. Posserud, A. J. Lundervold, and C. Gillberg, "Validation of the autism spectrum screening questionnaire in a total population sample," *Journal of Autism and Developmental Disorders*, vol. 39, no. 1, pp. 126–134, 2009.

[21] S. Ehlers, C. Gillberg, and L. Wing, "A screening questionnaire for Asperger syndrome and other high-functioning autism spectrum disorders in school age children," *Journal of Autism and Developmental Disorders*, vol. 29, no. 2, pp. 129–141, 1999.

[22] M. Rutter, A. Bailey, and C. Lord, *Manual for the Social Communication Questionnaire*, Western Psychological Services, Los Angeles, Calif, USA, 2003.

[23] J. N. Constantino and C. P. Gruber, *Social Responsiveness Scale (SRS)*, Western Psychological Services, Los Angeles, Calif, USA, 2005.

[24] J. N. Constantino and R. D. Todd, "Genetic structure of reciprocal social behavior," *American Journal of Psychiatry*, vol. 157, no. 12, pp. 2043–2045, 2000.

[25] J. N. Constantino and R. D. Todd, "Autistic traits in the general population: a twin study," *Archives of General Psychiatry*, vol. 60, no. 5, pp. 524–530, 2003.

[26] J. N. Constantino, J. J. Hudziak, and R. D. Todd, "Deficits in reciprocal social behavior in male twins: evidence for a genetically independent domain of psychopathology," *Journal of the American Academy of Child and Adolescent Psychiatry*, vol. 42, no. 4, pp. 458–467, 2003.

[27] Y. Kamio, N. Inada, A. Moriwaki et al., "Quantitative autistic traits ascertained in a national survey of 22, 529 Japanese schoolchildren," *Acta Psychiatrica Scandinavica*, vol. 128, pp. 45–55, 2013.

[28] S. Bölte, F. Poustka, and J. N. Constantino, "Assessing autistic traits: cross-cultural validation of the social responsiveness scale (SRS)," *Autism Research*, vol. 1, no. 6, pp. 354–363, 2008.

[29] J. N. Constantino, S. A. Davis, R. D. Todd et al., "Validation of a brief quantitative measure of autistic traits: comparison of the social responsiveness scale with the Autism Diagnostic Interview-Revised," *Journal of Autism and Developmental Disorders*, vol. 33, no. 4, pp. 427–433, 2003.

[30] J. N. Constantino and C. P. Gruber, *Social Responsiveness Scale (SRS-2)*, Western Psychological Services, Los Angeles, Calif, USA, 2nd edition, 2012.

[31] E. Fombonne, C. Marcin, R. Bruno, C. M. Tinoco, and C. D. Marquez, "Screening for autism in Mexico," *Autism Research*, vol. 5, no. 3, pp. 180–189, 2012.

[32] J. Wang, L. Lee, Y. Chen, and J. Hsu, "Assessing autistic traits in a Taiwan preschool population: cross-cultural validation of the Social Responsiveness Scale (SRS)," *Journal of Autism and Developmental Disorders*, vol. 42, no. 11, pp. 2450–2459, 2012.

[33] J. N. Constantino, T. Przybeck, D. Friesen, and R. D. Todd, "Reciprocal social behavior in children with and without pervasive developmental disorders," *Journal of Developmental and Behavioral Pediatrics*, vol. 21, no. 1, pp. 2–11, 2000.

[34] J. N. Constantino, P. D. Lavesser, Y. Zhang, A. M. Abbacchi, T. Gray, and R. D. Todd, "Rapid quantitative assessment of autistic social impairment by classroom teachers," *Journal of the American Academy of Child and Adolescent Psychiatry*, vol. 46, no. 12, pp. 1668–1676, 2007.

[35] Y. Kamio, H. Tsujii, N. Inada et al., "Screening for undiagnosed children and adults with autism spectrum disorders," in *The Annual Research Report on the Development of Diagnosis and Treatment for Individuals with Developmental Disorders*, M. Okuyama, Ed., pp. 25–35, National Center for Child Health and Development, Tokyo, Japan, 2009, (Japanese).

[36] A. Moriwaki, T. Koyama, and Y. Kamio, "Standardization of the Japanese version of the Social Responsiveness Scale," in *The Annual Report of Research Supported by Health and Labour Sciences Research Grants*, Y. Kamio, Ed., pp. 49–68, National Center of Neurology and Psychiatry, Japan Tokyo, Japan, 2011, (Japanese).

[37] C. Lord, M. Rutter, and A. Le Couteur, "Autism diagnostic interview-revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders," *Journal of Autism and Developmental Disorders*, vol. 24, no. 5, pp. 659–685, 1994.

[38] K. J. Tsuchiya, K. Matsumoto, A. Yagi et al., "Reliability and validity of the Autism Diagnostic Interview-Revised Japanese Version," *Journal of Autism and Developmental Disorders*, vol. 43, no. 3, pp. 643–662, 2013.

[39] C. Lord, S. Risi, L. Lambrecht et al., "The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism," *Journal of Autism and Developmental Disorders*, vol. 30, no. 3, pp. 205–223, 2000.

[40] M. Kuroda, N. Inada, R. Yukihiro et al., "Autism Diagnosis Observation Schedule-Generic (ADOS-G): reliability and validity of the Japanese version of ADOS-G, module 1-4," in *The Annual Report of Research Supported by Health and Labour Sciences Research Grants*, T. Uchiyama, Ed., pp. 31–38, Fukushima University, Fukushima, Japan, 2013, (Japanese).

[41] T. Uchiyama, T. Yoshikawa, Y. Uno, and T. Nakayama, "The Diagnostic Interview for Social and Communication Disorders-11 (DISCO-11): a study of the Japanese version of the Diagnostic Interview for Social and Communication Disorders-11 (DISCO-11)," in *The Annual Report of Research Supported by Health and Labour Sciences Research Grants*, T. Uchiyama, Ed., pp. 13–18, Fukushima University, Fukushima, Japan, 2013, (Japanese).

[42] H. Ito, I. Tani, R. Yukihiro et al., "Validation of an interview-based rating scale developed in Japan for pervasive developmental disorders," *Research in Autism Spectrum Disorders*, vol. 6, no. 4, pp. 1265–1272, 2012.

[43] S. M. Kanne, A. M. Abbacchi, and J. N. Constantino, "Multi-informant ratings of psychiatric symptom severity in children with autism spectrum disorders: the importance of environmental context," *Journal of Autism and Developmental Disorders*, vol. 39, no. 6, pp. 856–864, 2009.

[44] F. J. Aldridge, V. M. Gibbs, K. Schmidhofer, and M. Williams, "Investigating the clinical usefulness of the social responsiveness scale (SRS) in a tertiary level, autism spectrum disorder specific assessment clinic," *Journal of Autism and Developmental Disorders*, vol. 42, no. 2, pp. 294–300, 2012.

[45] P. Szatmari, L. Archer, S. Fisman, and D. L. Streiner, "Parent and teacher agreement in the assessment of pervasive developmental disorders," *Journal of Autism and Developmental Disorders*, vol. 24, no. 6, pp. 703–717, 1994.

[46] M.-C. Lai, M. V. Lombardo, G. Pasco et al., "A behavioral comparison of male and female adults with high functioning autism spectrum conditions," *PLoS ONE*, vol. 6, no. 6, Article ID e20835, 2011.

[47] A. M. Reiersen, J. N. Constantino, H. E. Volk, and R. D. Todd, "Autistic traits in a population-based ADHD twin sample," *Journal of Child Psychology and Psychiatry and Allied Disciplines*, vol. 48, no. 5, pp. 464–472, 2007.

[48] K. E. Towbin, A. Pradella, T. Gorrindo, D. S. Pine, and E. Leibenluft, "Autism spectrum traits in children with mood and anxiety disorders," *Journal of Child and Adolescent Psychopharmacology*, vol. 15, no. 3, pp. 452–464, 2005.

J Autism Dev Disord (2013) 43:643–662
DOI 10.1007/s10803-012-1606-9

ORIGINAL PAPER

Reliability and Validity of Autism Diagnostic Interview-Revised, Japanese Version

Kenji J. Tsuchiya · Kaori Matsumoto · Atsuko Yagi · Naoko Inada · Miho Kuroda · Eiko Inokuchi · Tomonori Koyama · Yoko Kamio · Masatsugu Tsujii · Saeko Sakai · Ikuko Mohri · Masako Taniike · Ryoichiro Iwanaga · Kei Ogasahara · Taishi Miyachi · Shunji Nakajima · Iori Tani · Masafumi Ohnishi · Masahiko Inoue · Kazuyo Nomura · Taku Hagiwara · Tokio Uchiyama · Hironobu Ichikawa · Shuji Kobayashi · Ken Miyamoto · Kazuhiko Nakamura · Katsuaki Suzuki · Norio Mori · Nori Takei

Published online: 18 July 2012

Abstract To examine the inter-rater reliability of Autism Diagnostic Interview-Revised, Japanese Version (ADI-R-JV), the authors recruited 51 individuals aged 3–19 years, interviewed by two independent raters. Subsequently, to assess the discriminant and diagnostic validity of ADI-R-JV, the authors investigated 317 individuals aged 2–19 years, who were divided into three diagnostic groups as follows: autistic disorder (AD), pervasive developmental disorder not otherwise specified, and other psychiatric diagnosis or no diagnosis, according to the consensus clinical diagnosis. As regards inter-rater reliability, intra-class correlation coefficients of greater than 0.80 were obtained for all three domains of ADI-R-JV. As regards discriminant validity, the mean scores of the three domains was significantly higher in individuals with AD than in those of other diagnostic groups. As regards diagnostic validity, sensitivity and specificity for correctly diagnosing AD were 0.92 and 0.89, respectively, but sensitivity was 0.55 for individuals younger than 5 years. Specificity was consistently high regardless of age and intelligence. ADI-R-JV was shown to be a reliable tool, and has sufficient discriminant validity and satisfactory diagnostic validity for correctly diagnosing AD, although the diagnostic validity appeared to be compromised with respect to the diagnosis of younger individuals.

Keywords Autism · ADI-R · Reliability · Validity · Japan

Electronic supplementary material The online version of this article (doi:10.1007/s10803-012-1606-9) contains supplementary material, which is available to authorized users.

K. J. Tsuchiya · M. Tsujii · K. Suzuki · N. Mori · N. Takei
Department of Child Development, United Graduate School of Child Development at Hamamatsu, Handayama 1 Higashiku, Hamamatsu 431-3192, Japan

K. J. Tsuchiya (✉) · K. Matsumoto · M. Tsujii · S. Nakajima · I. Tani · M. Ohnishi · K. Suzuki · N. Mori · N. Takei
Research Center for Child Mental Development, Hamamatsu University School of Medicine, Handayama 1 Higashiku, Hamamatsu 431-3192, Japan
e-mail: tsuchiya@hama-med.ac.jp

A. Yagi · K. Nakamura · N. Mori
Department of Neurology and Psychiatry, Hamamatsu University School of Medicine, Handayama 1 Higashiku, Hamamatsu 431-3192, Japan

A. Yagi
Shizuoka Center for Child and Family Welfare, Shizuoka, Japan

N. Inada · M. Kuroda · E. Inokuchi · T. Koyama · Y. Kamio
Department of Child and Adolescent Mental Health, National Institute of Mental Health, National Center of Neurology and Psychiatry, Tokyo, Japan

M. Kuroda
Department of Psychology, Shukutoku University, Chiba, Japan

M. Kuroda
Department of Child Neuropsychiatry, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

T. Koyama
Kokeijuku Preparatory School, Kumamoto, Japan

M. Tsujii
Chukyo University School of Contemporary Sociology, Toyota, Japan

S. Sakai · I. Mohri · M. Taniike
Division of Developmental Neuroscience, United Graduate School of Child Development, Osaka University, Osaka, Japan

S. Sakai · I. Mohri · M. Taniike
Research Center for Child Mental Development, United Graduate School of Child Development, Osaka University, Osaka, Japan

R. Iwanaga
Division of Physical and Occupational Therapy, Department of Health Sciences, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan

Introduction

Autistic disorder (AD) is defined by irregularities in three behavioral domains, namely, deficits in reciprocal social interaction, deficits in communication, and restricted and repetitive behaviors and interests (American Psychiatric Association 2000). AD is classified as an autism spectrum disorder (ASD), an umbrella term that encompasses AD and pervasive developmental disorder not otherwise specified (PDDNOS). The reported prevalence estimates of AD or ASD have been increasing (Fombonne 2009; Williams et al. 2006), with the prevalence of ASD now thought to be between 1 and 2 per 100 school children in the United Kingdom (Baron-Cohen et al. 2009) and in Japan (Kawamura et al. 2008), and even higher in South Korea (Kim et al. 2011). The observed change in prevalence estimates has been suggested to be an artifact due to increased awareness of ASDs, changes in diagnostic precision, and recent trends toward earlier diagnosis (Kočovská et al. 2012; Parner et al. 2008; Waterhouse 2008). Such observations have hastened the worldwide demand for reliable and valid methods of identifying ASD.

A number of questionnaires, interviews, and observation schedules have been developed to assist clinicians and researchers in the diagnostic assessment of specific behaviors found in individuals with AD or ASD. Among these instruments, Autism Diagnostic Interview-Revised (ADI-R (Lord et al. 1994)) is a structured, investigator-based interview directed to caregivers for the detection of AD in a research context. ADI-R has been widely used, and its reliability and validity have been examined in the original as well as in non-English versions (Cicchetti et al. 2008; Hill et al. 2001; Lampi et al. 2010; Lord et al. 1994; Mildenberger et al. 2001).

Discussions of ADI-R have accumulated, particularly as regards its diagnostic validity. Despite the fact that ADI-R provides a good to excellent level of sensitivity for diagnosing and predicting AD among varying samples (de Bildt et al. 2004; Gray et al. 2008; Lampi et al. 2010; Lord et al. 1994, 2006; Tomanik et al. 2007), studies have pointed out compromised diagnostic validity in certain types of examinees, such as younger children, because some symptoms are not evident at an early age (Cox et al. 1999; Rutter et al. 2003). This observation is of particular relevance among individuals with ASD other than AD (Gilchrist et al. 2001). On the other hand, as the algorithm-based diagnosis with ADI-R is made with reference to current as well as past behaviors, caregivers of examinees tend to report fewer symptoms when examinees are in adolescence or early adulthood (McGovern and Sigman 2005). Furthermore, depending on the level of function, ADI-R diagnoses of AD among children exhibiting a cognitive delay are less likely to conform to clinical or other types of research-related diagnosis (de Bildt et al. 2004), such as those based on Autism Diagnostic Observation Schedule (ADOS (Lord et al. 2000)). It should be noted that the use of ADOS alone has limited predictability (Lord et al. 2006). Considering these pitfalls, some groups have recommended that not a single source but rather multiple sources of information, including both ADI-R and ADOS, should be consulted when establishing a diagnosis of ASD or AD (Le Couteur et al. 2008; Lord et al. 2006), particularly in a research context. It follows

K. Ogasahara
Comprehensive Educational Science Division, Tokyo Gakugei University, Tokyo, Japan

T. Miyachi
Department of Pediatrics, Nagoya City University Hospital, Nagoya, Japan

I. Tani
Faculty of Humanities, Tokai Gakuen University, Aichi, Japan

M. Inoue · K. Nomura
Graduate School of Medical Science, Center for Clinical Psychology, Tottori University Faculty of Medicine, Tottori, Japan

K. Nomura
Department of Child Psychiatry, Hamamatsu University School of Medicine, Handayama 1 Higashiku, Hamamatsu 431-3192, Japan

T. Hagiwara
Division of Special Education, Department of Education and Development, Hokkaido University of Education, Asahikawa, Japan

T. Uchiyama
Faculty of Human Development and Culture, Fukushima University, Fukushima, Japan

H. Ichikawa
Tokyo Metropolitan Children's Medical Center, Tokyo, Japan

S. Kobayashi · K. Miyamoto
Department of Pediatrics, Kosai City Hospital, Kosai, Japan

N. Takei
King's College of London, Institute of Psychiatry, London, UK

that the foundation of reliability and validity of ADI-R is important in countries such as Japan, where such diagnostic tools have not been readily available.

ADI-R in particular was unavailable in Japan until 2005, when the present authors translated the WPS Edition of ADI-R (Rutter et al. 2003) into Japanese, at which time the back-translation was confirmed to be congruent with the original version by the developers of ADI-R. However, the reliability and validity of the Japanese version had remained unexamined to date.

Therefore, in the present study, the authors aimed to test the inter-rater reliability and discriminant and diagnostic validity of ADI-R, Japanese Version (ADI-R-JV). The inter-rater reliability was assessed using two types of agreement measures: the weighted Kappa (Kw) and intra-class correlation coefficient (ICC) of diagnostic algorithm item scores of two independent interviewers. Discriminant validity was assessed by comparing mean scores of diagnostic algorithm items/subdomains/domains between individuals with and without a consensus clinical diagnosis. Diagnostic validity in this study refers to agreement between the algorithm diagnosis based on ADI-R-JV and a consensus clinical diagnosis. The sensitivity, specificity, positive predictive value, and negative predictive value were calculated to assess this agreement.

For our assessment, we hypothesized the following.

1. Good to excellent inter-rater reliability in terms of the Kw and ICC of ADI-R-JV would be observed, which would be consistent with the published literature (Cicchetti et al. 2008; Hill et al. 2001; Lord et al. 1994).
2. The discriminant validity of ADI-R-JV would be sufficient, with higher mean scores of diagnostic algorithm items among individuals with AD than among those without AD (Lampi et al. 2010; Lord et al. 1994). That is, it was expected that AD scores > non-ASD scores, and AD scores > PDDNOS scores.
3. The diagnostic validity of ADI-R-JV would be satisfactory yet compromised among younger individuals and individuals with intellectual disabilities (Cox et al. 1999; de Bildt et al. 2004; Rutter et al. 2003).

Methods

Participants and Diagnostic Procedure

Reliability Study

To enroll study subjects, we recruited participants from 3 research sites, namely, 2 developmental, university-affiliated clinics and 1 research center. Basically, these clinics are open for referrals from local health practitioners. Participants were selected on the basis of the cumulative number of participants thus far enrolled (targeted N = 30), age (kindergarteners or school-age children/adolescents under 20 years of age), clinical diagnosis (confirmed or suspected diagnosis of ASD), and the provision of consent to participate in the study voluntarily, including videotaping. Thus, purposive sampling was incorporated into the study design.

For the reliability study, we recruited 35 individuals who were referred to one of our research sites between December 1, 2006 and November 30, 2010 (Table 1). Among them, 31 individuals had been already suspected of having ASD by their local health practitioners and had been referred to our institutions for a more definitive diagnosis. Soon after participating in this study, these participants underwent a clinical assessment based on DSM-IV-TR (American Psychiatric Association 2000) assessment, conducted by one of the authors. After the detailed clinical assessments were complete and comprehensive caregiver interviews were conducted in order to collect the developmental history of the participants, our research team provided consensus clinical diagnoses based on DSM-IV-TR. Our research team included clinical experts with more than 3 years of experience in pediatrics or in child neurodevelopmental practices and in assessing individuals with ASD (5 certified clinical psychologists, 3 child psychiatrists, and 4 pediatricians were involved). A total of 31 individuals were confirmed to have a consensus clinical diagnosis of ASD, namely, AD (N = 12) or PDDNOS (N = 19). The remaining 4 individuals were referred to our research sites on the basis of suspected intellectual impairment, and they were confirmed not to have a diagnosis of ASD according to the same diagnostic procedures as those used for the confirmed ASD cases.

The 35 clinically referred individuals were also examined with respect to cognitive measures. For those subjects who were age 5 or older, the Japanese version of the Wechsler Intelligence Scale for Children, third edition (WISC-III: (Wechsler et al. 1992)) or the Tanaka-Binet intelligence scale (Tanaka Institute of Education 1987) was used to estimate the intelligence quotient (IQ). For individuals younger than 5 years old, a standardized developmental test, the Kyoto Scale of Psychological Development (Koyama et al. 2009), was adopted to estimate development quotient (DQ). Among the 31 individuals with ASD, 6 had a full-scale IQ/DQ of lower than 70. Among the 4 non-ASD clinical individuals, all had a full-scale IQ/DQ of lower than 70.

In addition to the clinically referred individuals, 16 kindergarteners and school-age children exhibiting typical development were also invited to participate in the study as

Table 1 Reliability study: characteristics of the sample studied

	Clinically referred individuals [N = 35]	Control individuals [N = 16]	Statistics
Age in years			
Range	3–18	3–14	
Median	5.0	5.0	
Mean (SD)	8.7 (5.2)	7.0 (3.8)	t(49) = 1.16, $p = 0.25$
Gender (F:M)	5:30	4:12	Chi-square(1) = 0.84, $p = 0.36$
Full scale IQ/DQ[a]			
Number of individuals with cognitive delay (IQ/DQ < 70)	10 (29 %)	0 (0 %)	Chi-square(1) = 5.67, exact $p = 0.02$
Range	42–118	86–124	
Median	81	102.5	
Mean (SD)	81.9 (22.6)	102.0 (11.6)	t(44) = 2.85, $p < 0.001$
DSM-IV-TR diagnosis			
Autistic disorder	11 (31 %)	0	
Autistic disorder + mental retardation	1 (3 %)	0	
Pervasive developmental disorder, not otherwise specified	14 (20 %)	0	
Pervasive developmental disorder, not otherwise specified + Mental retardation	5 (14 %)	0	
Mental retardation	4 (11 %)	0	
Major depressive disorder	0	1 (6 %)	
Adjustment disorder	0	1 (6 %)	
No psychiatric diagnosis	0	14 (88 %)	
ADI-R score (based on data derived from a first examiner)			
Domain A			
Range	5–28	0–7	
Median	18	3.5	
Mean (SD)	15.9 (6.6)	3.3 (2.8)	t(49) = 7.16, $p < 0.001$
Domain BV[a]	[N = 23]	[N = 14]	
Range	3–14	0–8	
Median	7	2	
Mean (SD)	7.3 (3.6)	3.3 (2.9)	t(35) = 6.94, $p < 0.001$
Domain BNV[b]	[N = 12]	[N = 2]	
Range	1–12	0–1	
Median	8	0.5	
Mean (SD)	6.9 (4.5)	0.5 (0.7)	t(12) = 1.96, $p = 0.07$
Domain C			
Range	0–11	0–4	
Median	3	0.5	
Mean (SD)	3.5 (2.5)	1.3 (1.5)	t(49) = 3.35, $p = 0.002$

[a] 5 Individuals, all aged 6 years or older, in the control individuals have no data on IQ/DQ. The school records of these participants were carefully checked and we regarded their histories as equivalent to a lack of cognitive delay

[b] Verbal subjects (defined as a score of 0 on item 30 "overall level of language")

[c] Non-verbal subjects (defined as a score of 1 or 2 on item 30)

control individuals. The control groups was recruited via a notice published in newspapers local to three of our research sites, where the clinically referred individuals for the reliability study had also been enrolled. The characteristics of these control individuals are given in Table 1. Considering the male predominance among clinically

referred children, boys were intentionally oversampled. The control subjects underwent clinical assessment based on DSM-IV-TR in an interview conducted by one of the authors, and the results were later confirmed by our research team according to the same procedures as those described above. Among the control subjects, 1 individual had a diagnosis of major depressive disorder, and 1 had a diagnosis of adjustment disorder. All 16 control individuals were also examined either using WISC-III, the Kyoto Scale of Psychological Development, or the Tanaka-Binet intelligence scale, depending on the subject's mental age, and none of the control subjects were confirmed to have any cognitive delays.

In sum, the enrolled participants comprised two groups (Table 1): 35 clinically referred individuals and 16 control individuals. The mean age of these two groups did not differ significantly (8.7 [SD 5.2] vs. 7.0 [SD 3.8]; $t(49) = 1.15$, $p = 0.25$), and the F:M ratio did not differ (F:M = 5:30 vs. 4:12; Chi-square (1) = 0.84, $p = 0.35$), although the mean IQ/DQ differed significantly (81.9 [SD 22.6] vs. 102.0 [SD 11.6]; $t(44) = 4.9$, $p < 0.001$).

Validity Study

To collect a sufficient number of clinically referred individuals in this sub-study, 6 additional research sites were involved (4 developmental, university-affiliated clinics, 1 pediatric clinic at a general hospital, and 1 privately run clinic for child psychiatry), together with the three research sites also involved in the reliability study. The mode of purposive selection of study participants was the same as that adopted in the reliability study except that in the validity study, the targeted number of participants was larger (N = 200), and the recruitment period was longer (September 1, 2006 and March 31, 2011). To capture any differences between the two recruitment methods used for the two sub-studies, we compared 35 clinically referred individuals enrolled in the reliability study and an additional 200 clinically referred individuals (not shown in the Table). This comparison did not reveal any significant difference in the F:M ratio (F:M = 5:30 vs. 42:158; Chi-Square(1) = 0.84, $p = 0.36$), no significant difference in mean age (mean = 8.7 (SD 5.2) vs. 10.5 (SD 4.9) years; $t(233) = 0.61$, $p = 0.54$), and no significant difference in mean DQ/IQ (81.9 (SD 22.6) versus 89.2 (SD 24.8); $t(233) = 1.62$, $p = 0.11$) between the two groups of individuals. Therefore, we regarded these two groups as basically the same in terms of background characteristics. We then combined the two groups and considered them as feasible for the analysis. A total of 235 clinically referred individuals were enrolled in the validity study.

To establish the group of control individuals, 66 kindergarteners and school-age children exhibiting typical development were also invited to participate in this study. Participants were recruited through a notice placed in local newspapers that serve the regions of the nine research sites at which the 235 clinically referred individuals were also enrolled. As a group, these individuals were identical in terms of mean age, F:M ratio, and mean IQ/DQ to the 16 control individuals enrolled in the reliability study, and as such, they were combined as a single control group of individuals. As a result, for the validity study, we investigated 235 clinically referred individuals and 82 control individuals (Appendix Table 2 in supplementary materials). The mean age of the 235 clinically referred individuals was older than that of the 82 control individuals (10.3 (SD 4.9) vs. 6.5 (SD 3.8) years; $t(315) = 6.42$, $p < 0.001$), and the mean full-scale IQ/DQ of the clinically referred individuals (86.6 (SD 23.0) vs. 100.2 (SD 13.3); $t(310) = 4.65$, $p < 0.001$) was lower than that of the control individuals. There were significantly more male individuals among the clinically referred individuals than among the control individuals (F:M = 47:188 vs. 34:48; Chi-Square(1) = 14.7, $p < 0.001$; see Appendix Table 2 in supplementary materials).

As was done in the reliability study, 235 clinically referred individuals and 82 control individuals underwent a clinical assessment based on DSM-IV-TR (American Psychiatric Association 2000) conducted by one of the authors, and diagnoses, if any, were confirmed by our research team and were established as a DSM-IV-TR-based consensus clinical diagnosis. Among the 235 clinically referred individuals, 227 were confirmed to have ASD, namely, AD (N = 138) or PDDNOS (N = 89) as the consensus clinical diagnoses. The remaining 8 individuals were assessed as not having ASD. Among the 82 control individuals, none had a diagnosis of ASD; however, 1 had a diagnosis of major depressive disorder, 1 had social phobia, 1 had attention deficit/hyperactive disorder not otherwise specified, and 1 had adjustment disorder. To measure IQ/DQ, WISC-III, Tanaka-Binet intelligence scale, or Kyoto Scale of Psychological Development was employed. Among the 82 control individuals, 12 had no IQ/DQ records; the school records of these participants were carefully checked and we regarded their histories as equivalent to a lack of cognitive delay.

Finally, the 235 clinically referred individuals and 82 control individuals were combined and re-grouped into the three following diagnostic groups based on a consensus clinical diagnosis (Table 2): 138 individuals with AD, 89 with PDDNOS, and 90 with non-ASD. Group comparisons of mean age across the three groups revealed a significantly higher value in the AD group than in the other two groups (AD 11.7 [SD 4.3], PDDNOS 8.5 [SD 5.1], non-ASD 6.4 [SD 3.7]; $F(2, 314) = 42.1$, $p < 0.001$). Likewise, the F:M ratio of the three groups showed a significant difference

Table 2 Validity study: characteristics of the sample studied

	(1) AD [N = 138]	(2) PDDNOS [N = 89]	(3) Non-ASD [N = 90]	Statistics
Age in years				
Range	2–19	2–19	2–17	
Median	11.8	8.0	5.0	
Mean (SD)	11.7 (4.3)	8.5 (5.1)	6.4 (3.7)	$F(2, 314) = 42.9$, $p < 0.001$ 1 > 3: $p < 0.001$ 2 > 3: $p < 0.001$ 1 > 2: $p < 0.001$
Gender (F:M)	18:120	25:64	38:52	Chi-square(2) = 24.8, $p < 0.001$
Number of individuals with cognitive delay (IQ/DQ < 70)	18 (13 %)	9 (10 %)	8 (9 %)	Chi-Square(2) = 1.1, $p = 0.59$
DSM-IV-TR diagnosis				
Autistic disorder	120 (87 %)	0	0	
Autistic disorder + mental retardation	18 (13 %)	0	0	
Pervasive developmental disorder, not otherwise specified	0	80 (90 %)	0	
Pervasive developmental disorder not otherwise specified + mental retardation	0	9 (10 %)	0	
Mental retardation	0	0	8 (9 %)	
Major depressive disorder	0	0	1 (1 %)	
Social phobia	0	0	1 (1 %)	
Attention deficit/hyperactive disorder, not otherwise specified	0	0	1 (1 %)	
Adjustment disorder	0	0	1 (1 %)	
No psychiatric diagnosis	0	0	78 (87 %)	
Full scale IQ/DQ[a]				
Range	41–140	42–131	45–132	
Median	87.5	90	93	
Mean (SD)	88.4 (22.8)	87.9 (20.7)	90.8 (23.1)	$F(2, 302) = 0.2$, $p = 0.82$
ADI-R score				
Domain A				
Range	8–30	3–28	0–11	
Median	20	13	1	
Mean (SD)	19.9 (5.3)	14.8 (6.4)	2.3 (2.7)	$F(2, 314) = 330.6$, $p < 0.001$ 1 > 3: $p < 0.001$ 2 > 3: $p < 0.001$ 1 > 2: $p < 0.001$
Domain BV[b]	[N = 116]	[N = 68]	[N = 79]	
Range	3–25	2–21	0–12	
Median	14	8.5	1	
Mean (SD)	14.3 (4.1)	9.7 (4.4)	2.5 (3.2)	$F(2, 260) = 210.9$, $p < 0.001$ 1 > 3: $p < 0.001$ 2 > 3: $p < 0.001$ 1 > 2: $p < 0.001$
Domain BNV[c]	[N = 22]	[N = 21]	[N = 11]	
Range	0–14	1–12	0–9	
Median	10	6	1	

Table 2 continued

	(1) AD [N = 138]	(2) PDDNOS [N = 89]	(3) Non-ASD [N = 90]	Statistics
Mean (SD)	12.6 (4.9)	9.0 (4.4)	2.3 (2.5)	$F(2, 51) = 21.0$, $p < 0.001$ 1 > 3: $p < 0.001$ 2 > 3: $p < 0.005$ 1 > 2: $p = 0.02$
Domain C				
Range	0–12	0–12	0–9	
Median	5	2	0	
Mean (SD)	5.5 (2.4)	2.9 (2.5)	1.1 (1.8)	$F(2, 314) = 106.6$, $p < 0.001$ 1 > 3: $p < 0.001$ 2 > 3: $p < 0.001$ 1 > 2: $p < 0.001$

NS not significant

[a] 12 individuals, all aged 6 years or older, in the Non-ASD group have no data on IQ/DQ. The school records of these participants were carefully checked and we regarded their histories as equivalent to a lack of cognitive delay

[b] Verbal subjects (defined as a score of 0 on item 30 "overall level of language")

[c] Non-verbal subjects (defined as a score of 1 or 2 on item 30)

(AD 18:120, PDDNOS 25:64, Non-ASD 38:52; Chi-Square(2) = 24.8, $p < 0.001$). The mean IQ/DQ did not differ across the three groups (AD 90.8 [SD 23.0], PDD-NOS 87.9 [SD 20.1], Non-ASD 88.3 [SD 88.3]; F(2, 302) = 0.2, $p = 0.82$), and the proportion of individuals with an IQ/DQ of less than 70 did not show any statistically significant departures from the expected values (AD 13 %, PDDNOS 10 %, Non-ASD 9 %, Chi-Square(2) = 1.07, $p = 0.59$).

With ADI-R-JV, an algorithm diagnosis of AD was provided if the sum scores of all of four domains (A, B, C, and D) met the criteria (equal to or exceeding the cutoff for each domain) as described in the original guidelines (Rutter et al. 2003).

Interviews Using ADI-R-JV

All caregivers of participants in this study were interviewed using ADI-R-JV within a 2-month period after the participants had taken part in the study. These interviews were conducted either by one of the present authors (KJT, KM, AY, SS) who established the research reliability of the original ADI-R together with the developers based on intensive training sessions at the training sites, namely, the interviewers reached more than 90 % exact agreement with the ADI-R trainers (Risi et al. 2006), or by the authors who were supervised by the authors KJT, KM, AY, or SS when the interview using ADI-R-JV was conducted. In this study, the same standard of agreement was achieved across all members of the research team who conducted ADI-R-JV. In total, 8 of the present authors were entitled to conduct interviews using ADI-R-JV, and thus were regarded as ADI-R-JV interviewers for the current study.

For the reliability study, all ADI-R-JV interviews were first conducted by one of four interviewers (KJT, KM, AY, SS), and all interviews were videotaped. Each tape was assessed independently by another rater from the same group of four interviewers, and all combinations of the four raters were equally likely. For the validity study, only one out of 8 interviewers conducted an ADI-R-JV interview, and that interviewer was blind to the consensus clinical diagnosis of the examinee. All 8 interviewers assessed participants at each research site on a random basis.

Analyses

Construction of ADI-R-JV Diagnostic Algorithm

ADI-R diagnostic algorithm consists of the following 4 domains: (A) Qualitative abnormalities in reciprocal social interaction; (B) Qualitative abnormalities in communication; (C) Restricted, repetitive, stereotyped patterns of behavior; and (D) Abnormality of development evident at or before 36 months. Domains A, B, and C correspond to the three groups of symptoms described in the DSM-IV-TR (American Psychiatric Association 2000). Domain A consists of 4 subdomains covering 16 algorithm items; domain B consists of 4 subdomains covering 13 algorithm items; domain C consists of 4 subdomains covering 8 algorithm items; and domain D has no subdomain and covers 5 algorithm items. Our analyses focused on each of 42 algorithm items, 12 subdomains and 3 domains (A, B, and C); we did not total up domain D scores and thus did not analyze this, since this is the summary code for evidence of abnormality within the first 3 years. The assessment of domain B was further divided into two types of assessments according to verbal skills of the examined individuals; subdomains B1, B4, B2 (V), and B3 (V), covering 13 algorithm items, were used for verbal individuals, whereas only B1 and B4 were used for non-verbal individuals (including pre-speech infants).

An algorithm-based diagnosis of AD was provided if all of scores of four domains (A, B, C, and D) were equal to or exceeded the following cut-off points: 10 points for domain A; 8 points for domain BV (domain B for verbal subjects) or 7 points for domain BNV (domain B for non-verbal subjects); 3 points for domain C; and 1 point for domain D.

Reliability Study

We first calculated the weighted kappa (Kw) value for each of the 42 algorithm items; scores on the algorithm items took only one of three values (0, 1, or 2). We adopted the quadratic weighting system, that is, $w_{ij} = 1 - (i - j)^2/(k - 1)^2$ (Fleiss and Cohen 1973). This allowed Kw and the intraclass correlation coefficient (ICC) to be considered as equivalent to each other. We also calculated the ICC for each of 12 subdomains and 4 domains; the summed scores of subdomains and domains could take a number of values, and thus the ICC was preferred over the Kw. As regards judgments of the clinical level of significance, we followed the criteria provided in previous studies (Cicchetti 1994; Cicchetti and Sparrow 1981), i.e., items showing Kw $\geq$ 0.75 and subdomains/domains showing ICC $\geq$ 0.75 were regarded as excellent, 0.60 $\leq$ Kw $<$ 0.75 and 0.60 $\leq$ ICC $<$ 0.75 were considered good, and 0.40 $\leq$ Kw $<$ 0.60 and 0.40 $\leq$ ICC $<$ 0.60 were considered fair, while Kw $<$ 0.40 and ICC $<$ 0.40 exhibited poor inter-rater reliability. Considering the difference in age distribution of the three diagnostic groups of participants, analyses were first conducted on all the enrolled participants, and then a subsequent analysis was conducted separately for three age bands: below 5 years (<5:0 years); 5 years 0 months to 9 years 11 months (5:0–9:11 years); and 10 years and older.

Validity Study—Discriminant Validity

We compared the mean scores for 42 algorithm items, 12 subdomains, and 3 domains (A, B, and C) among the three diagnostic groups of participants (AD, PDDNOS, and non-ASD) using one-way ANOVA analysis with a post hoc comparison after Bonferroni's correction. We also examined whether differences in the mean scores of items, subdomains, and domains would be smaller if the analyses were limited to younger individuals (<5 years of age) or individuals exhibiting cognitive delay (IQ/DQ < 70).

Validity Study—Diagnostic Validity

To assess whether the provided diagnosis based on ADI-R-JV was diagnostically valid, we estimated the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of ADI-R-JV. In this study, sensitivity referred to the proportion of individuals judged to have an ADI-R-JV algorithm-based diagnosis of AD among those with a consensus clinical diagnosis of AD. Specificity was the proportion of those judged not to have AD based on ADI-R-JV among those with a non-AD consensus clinical diagnosis or with no psychiatric diagnosis (i.e., subjects without a consensus clinical diagnosis of AD). PPV was the proportion of subjects with a consensus clinical diagnosis of AD among those with an algorithm-based diagnosis of AD, and NPV was the proportion of subjects with a consensus clinical diagnosis of non-AD among those with an algorithm-based diagnosis of non-AD. According to previously reported criteria (Cicchetti et al. 1995), we judged the clinical significance of sensitivity, specificity, and PPV and NPV values to be "fair" if results for these measures were equal to or exceeded 70 %, good if they were ≥80 %, and excellent if they were ≥90 %. We also examined whether results for these would be lower if the analysis were limited to that of younger individuals (<5 years of age) or individuals with an intellectual disability (IQ/DQ < 70).

Ethical Issues

The study protocol followed the ethical guidelines of the most recent Declaration of Helsinki (Edinburgh 2000) and was approved by the Institutional Ethical Review Boards at each research site. All participants, together with their caregivers, were given a complete description of the study, and the caregivers were asked to provide written informed consent to participate. As regards clinically referred individuals, they were initially contacted at one of the participating research sites, where we provided caregivers with routine feedback, which included our clinical observations and assessments. Then, by the time ADI-R-JV interview was conducted, we had formed a clinical consensus diagnosis, arrived at by experts in our research team. After ADI-R-JV interview with the caregivers had been conducted, we formulated a best-estimate diagnosis based on