宇野洋太、吉田友子	広汎性発達障害と注意欠如・多動性障害、学習障害	市川宏伸	専門医のための精神科臨床リュミエール19広汎性発達障害	中山書店	東京	2010	68-75
著者氏名	論文タイトル名	書籍全体の編集者名	書籍名	出版社名	出版地	出版年	ページ
宇野洋太、内山登紀夫	TEACCHによる療育	市川宏伸	専門医のための精神科臨床リュミエール19広汎性発達障害	中山書店	東京	2010	141-148
吉川徹	構造化による指導法ＴＥＡＣＣＨプログラムによる支援	松村暢隆・石川裕之・佐野亮子・小倉正義	ワードマップ 認知的個性―違いが活きる学びと支援―	新曜社	東京	2010	280-284

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Tsuchiya KJ, Matsumoto K, Yagi A, Inada N, Kuroda M, Inokuchi E, Koyama T, Kamio Y, Tsujii M, Sakai S, Mohri I, Taniike M, Iwanaga R, Ogasahara K, Miyachi T, Nakashima S, Tani I, Ohnishi M, Inoue M, Nomura K, Hagiwara T, Uchiyama T, Ichikawa H, Uchida H, Kobayashi S, Miyamoto K, Nakamura K, Suzuki K, Mori N, Takei N.	Reliability and Validity of Autism Diagnostic Interview-Revised, Japanese Version	J Autism Dev Disord		Epub ahead of print	2012
Uno Y, Uchiyama T.	The combined measles, mumps, and rubella vaccines and the total number of vaccines are not associated with development of autism spectrum disorder. The first case-control study in Asia.	Vaccine	30	4292-4298	2012
Ito H, Tani I, Yukihiro R, Adachi J, Hara K, Ogasawara M, Inoue M, Kamio Y, Nakamura K, Uchiyama T, Ichikawa H, Sugiyama T, Hagiwara T, Tsujii M	Validation of an interview-based rating scale developed in Japan forpervasive developmental disorders.	Research in Autism Spectrum Disorders	6(4)	1265–1272	2012
伊藤大幸,行廣隆次, 内山登紀夫, 黒田美保他.	日本版Vineland-II適応行動尺度の開発　不適応行動尺度の信頼性・妥当性に関する報告	精神医学	54 (9)	889-898	2012
内山登紀夫	広汎性発達障害とｽﾍﾟｸﾄﾗﾑ概念	精神科治療学	27(4)	443-451	2012
田中恭子、内山登紀夫	発達障がい児への支援の基本的な考え方	小児看護	35	534-540	2012

研究成果の刊行に関する一覧表(6 / 13)

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
内山登紀夫	大人の自閉症ｽﾍﾟｸﾄﾗﾑ障害の診断	治療	94(8)	1376-1380	2012
蜂矢百合子, 内山登紀夫	3歳から就学年齢までの場合	小児内科	44(5)	714-718	2012
吉田香織, 内山登紀夫	広汎性発達障害の心理社会的支援をめぐって.	Pharma Medica	30(4)	33-36	2012
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		6-27	2013
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.	Acta Psychiatrica Scandinavica		DOI 10. 1111/ acps. 12034	online
Takahashi H, Hashimoto R, Iwase M, Ishii R, Kamio Y, Takeda M.	Prepulse inhibition of startle response: recent advances in human studies of psychiatric disease.	Clinical Psychopharmacology and Neuroscience	9	102-110	2012
Katagiri M, Kasai T, Kamio Y, Murohashi H	Individuals with Asperger's Disorder Exhibit Difficulty in Switching Attention from a Local Level to a Global Level.	J Autism Dev Disord	43	395-403	2013
神尾陽子.	子どもの社会性の発達の障害	特集子どもの社会性の形成・発達の基礎基盤, 子どもと発育発達	10	161-165.	2012
高橋秀俊,深津玲子,神尾陽子	成人 ASD の社会参加に向けて.	精神科	21	687-691	2012
稲田尚子, 神尾陽子	早期アセスメントと早期支援. 特集「発達障害支援」	臨床心理学	12	628-633	2012
Anitha A, Nakamura K, Thanseem I, Yamada K, Iwayama Y, Toyota T, Matsuzaki H, Miyachi T, Yamada S, Tsujii M, Tsuchiya KJ, Matsumoto K, Iwata Y, Suzuki K, Ichikawa H, Sugiyama T, Yoshikawa T, Mori N.	Brain region-specific altered expression and association of mitochondria-related genes in autism.	Mol Autism	3(1)	12	2012
Anitha A, Nakamura K, Thanseem I, Matsuzaki H, Miyachi T, Tsujii M, Iwata Y, Suzuki K, Sugiyama T, Mori N.	Downregulation of the Expression of Mitochondrial Electron Transport Complex Genes in Autism Brains.	Brain Pathol.	10		2012
Anitha A, Thanseem I, Nakamura K, Yamada K, Iwayama Y, Toyota T, Iwata Y, Suzuki K, Sugiyama T, Tsujii M, Yoshikawa T, Mori N.	Protocadherin α (PCDHA) as a novel susceptibility gene for autism.	J Psychiatry Neurosci.	37(6)		2012

研究成果の刊行に関する一覧表（ 7 / 13 ）

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Egawa J, Watanabe Y, Nunokawa A, Endo T, Kaneko N, Tamura R, Sugiyama T, Someya T.	A detailed association analysis between the tryptophan hydroxylase 2 (TPH2) gene and autism spectrum disorders in a Japanese population.	Psychiatry Res.	196(2-3)	320-2	2012
吉田友子	自閉症スペクトラムを告知するということ	精神神経学雑誌	115 (5-7)	未定	2013
黒田美保，稲田尚子	Autism Diamostic Observation Schedule（自閉症診断観察検査）日本語版の開発状況と今後の課題	精神医学	54(4)	427-433	2012
稲田尚子, 神尾陽子	自閉症スペクトラム幼児に対する早期支援の有効性に対する客観的評価：成果と考察.	乳幼児医学・心理学研究，特集「自閉症スペクトラム障害の早期療育への前方視的研究」	20(2)	73-81	2012
Yoko Kamio, Naoko Inada and Tomonori Koyama	A nationwide survey on quality of life and associated factors of adults with high-functioning autism spectrum disorders	Autism	Published online	doi:10.1177/1362361312436848	2012
Egawa J, Watanabe Y, Nunokawa A, Endo T, Kaneko N, Tamura R, Sugiyama T, Someya T.	A detailed association analysis between the tryptophan hydroxylase 2 (TPH2) gene and autism spectrum disorders in a Japanese population.	Psychiatry Res.		Epub ahead of print.	2012
森本武, 杉山登志郎, 東誠	広汎性発達障害における双極性障害の臨床的検討	小児の精神と神経	52(1)	35-44	2012
井上雅彦、岡田涼、野村和代、安達潤、辻井正次、大塚晃、市川宏伸	強度行動障害における自閉性障害との関連性　日本自閉症協会評定尺度（ＰＡＲＳ）短縮版による分析	精神医学	54(5)	473-481	2012
安達　潤	特別支援教育の現在とスクールカウンセラーの役割	こころの科学	163	71-74	2012
安達　潤	PARS：評定の視点と活用の留意点	児童青年精神医学とその近接領域	53(3)	299-305	2012
寺崎真一郎、安達潤	被虐待体験を持つ幼児への対人トラブルの軽減を目的とした支援について	子どもの虐待とネグレクト	14(2)	183-194	2012
安達潤、齊藤真善、萩原拓、神尾陽子	アイトラッカーを用いた高機能広汎性発達障害者における会話の同調傾向の知覚に関する実験的検討	児童青年精神医学とその近接領域	53(5)	561-576	2012
内山登紀夫	思春期から成人期の広汎性発達障害　思春期から成人期の自閉症スペクトラム	児童青年精神医学とその近接領域	52(4)	431-436	2011
Tanaka, K, Uchiyama,T, Endo, F	Informing children about their sibling's diagnosis of autism spectrum disorder: An initial investigation into current practices	Research in Autism Spectrum Disorders	5(4)	1421-1429	2011

研究成果の刊行に関する一覧表(8 / 13)

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Kuroda M, Wakabayashi A, Uchiyama T, Yoshida Y, Koyama T, Kamio Y	Determining differences in social cognition between high-functioning autistic disorder and other pervasive developmental disorders using new advanced "mind-reading" tasks	Research in Autism Spectrum Disorders	5(1)	554-561	2011
神尾陽子	20年後を見据えた精神医学・心身医学研究の展望特集. 児童精神医学研究の将来展望.	学術の動向	7	15-19	2011
神尾陽子	教育講演　児童期から成人期へ：レジリエンスという視点.	児童青年精神医学とその近接領域	52(4)	379-384	2011
神尾陽子	災害時に見えてくる、これからの子どものメンタルヘルス対策に必要なこと（巻頭言）.	精神医学	53	934-935	2011
Kuwano Y, Kamio Y, Kawai T, Katsuura S, Inada N, Takaki A, Rokutan K	Autism-Associated Gene Expression in Peripheral Leucocytes Commonly Observed between Subjects with Autism and Healthy Women Having Autistic Children.	PLoS ONE	6(9)	e24723	2011
Takahashi H, Hashimoto R, Iwase M, Ishii R, Kamio Y, Takeda M.	Prepulse Inhibition of Startle Response: Recent Advances in Human Studies of Psychiatric Disease.	Clinical Psychopharmacology and Neuroscience	9 (3)	102-1102	2011
Kuroda M, Wakabayashi A, Uchiyama T, Yoshida Y, Koyama T, Kamio Y	Determining Differences in Social Cognition between High-Functioning Autistic Disorder and Other Pervasive Developmental Disorders Using New Advanced "Mind-Reading" Tasks	Research in Autism Spectrum Disorder	5	554-561	2011
Inada N, Koyama T, Inokuchi E, Kuroda M, & Kamio Y	Reliability and validity of the Japanese version of the Modified Checklist for Autism in Toddlers (M-CHAT)	Research in Autism Spectrum Disorder	5	330-336	2011
KoyamaT, nokuchi E, Inada N, Kuroda M, Moriwaki A, Katagiri M, Noriuchi M, Kamio Y	Utility of the Japanese version of the Checklist for Autism in Toddlers (CHAT-J) for predicting pervasive developmental disorders at age 2	Psychiatry and Clinical Neurosciences	64	330-332	2011
吉川徹、　金田昌子	広汎性発達障害と解離性障害	児童青年精神医学とその近接領域	52(2)	178-185	2011
明翫 光宜 ， 辻井 正次	子どもたちの「できること」を伸ばす--発達障害のある子どものスキル・トレーニング実践(10)怒りと不安をコントロールする	こころの科学	155	129-134	2011
南谷 奈穂 ， 辻井 正次	子どもたちの「できること」を伸ばす--発達障害のある子どものスキル・トレーニング実践(11)相互交渉のスキルを学ぶ	こころの科学	156	124-129	2011

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Iwata K, Matsuzaki H, Miyachi T, Shimmura C, Suda S, Tsuchiya KJ, Matsumoto K, Suzuki K, Iwata Y, Nakamura K, Tsujii M, Sugiyama T, Sato K, Mori N.	Investigation of the serum levels of anterior pituitary hormones in male children with autism.	Mol Autism.	19	2-16	2011
Thanseem I, Nakamura K, Anitha A, Suda S, Yamada K, Iwayama Y, Toyota T, Tsujii M, Iwata Y, Suzuki K, Matsuzaki H, Iwata K, Sugiyama T, Yoshikawa T, Mori N.	Association of transcription factor gene LMX1B with autism.	PLoS One.	6(8)	E23738	2011
Suzuki K, Matsuzaki H, Iwata K, Kameno Y, Shimmura C, Kawai S, Yoshihara Y, Wakuda T, Takebayashi K, Takagai S, Matsumoto K, Tsuchiya KJ, Iwata Y, Nakamura K, Tsujii M, Sugiyama T, Mori N.	Plasma cytokine profiles in subjects with high-functioning autism spectrum disorders.		6(5)	e20470	2011
Suzuki K, Sugihara G, Ouchi Y, Nakamura K, Tsujii M, Futatsubashi M, Iwata Y, Tsuchiya KJ, Matsumoto K, Takebayashi K, Wakuda T, Yoshihara Y, Suda S, Kikuchi M, Takei N, Sugiyama T, Irie T, Mori N.	Reduced acetylcholinesterase activity in the fusiform gyrus in adults with autism spectrum disorders.	Arch Gen Psychiatry.	68(3)	306-313	2011
Nakamura K, Iwata Y, Anitha A, Miyachi T, Toyota T, Yamada S, Tsujii M, Tsuchiya KJ, Iwayama Y, Yamada K, Hattori E, Matsuzaki H, Matsumoto K, Suzuki K, Suda S, Takebayashi K, Takei N, Ichikawa H, Sugiyama T, Yoshikawa T, Mori N.	Replication study of Japanese cohorts supports the role of STX1A in autism susceptibility.	Prog Neuropsychopharmacol Biol Psychiatry.	35(2)	454-458	2011
Marui T, Funatogawa I, Koishi S, Yamamoto K, Matsumoto H, Hashimoto O, Jinde S, Nishida H, Sugiyama T, Kasai K, Watanabe K, Kano Y, Kato N.	The NADH-ubiquinone oxidoreductase 1 alpha subcomplex 5 (NDUFA5) gene variants are associated with autism.	Acta Psychiatr Scand.	123(2)	118-124	2011
杉山登志郎	そだちの凸凹（発達障害）とそだちの不全（子ども虐待）	日本小児看護学会誌	20(3)	103-107	2011
杉山登志郎	【子どもの虐待と脳の発達】子ども虐待と子どもの発達	子どものこころと脳の発達	2(1)	5-13	2011

研究成果の刊行に関する一覧表(10 / 13)

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
杉山登志郎	震災の中で生きる子ども】何ができるか　子どもにおける大震災の後遺症を減らすための対応	精神科治療学	25(12)	1639-1645	2011
杉山登志郎	【アスペルガー症候群のいま】アスペルガー症候群の最新理解　アスペルガー症候群再考	そだちの科学	17	2-11	2011
小野真樹、杉山登志郎、栗山貴久子,東誠,浦野葉子	高機能広汎性発達障害における同胞併発例の検討　攻撃性とその相互作用について	小児の精神と神経	51(3)	217-230	2011
杉山登志郎	【自閉症スペクトラムの生物学】自閉症スペクトラムとは	分子精神医学	11(4)	264-268	2011
杉山登志郎	【性的虐待】性的虐待の実態とケア	子どもの虐待とネグレクト	13(2)	209-215	2011
杉山登志郎	【医学的観点からみた子ども虐待～子ども虐待を理解し､実践的なかかわりをめざして～】虐待の発見　虐待による精神疾患の特徴	チャイルドヘルス	14(9)	1532-1535	2011
杉山登志郎	子ども虐待と精神医学	児童青年精神医学とその近接領域	52(3)	250-263	2011
海野千畝子, 小山内文, 杉山登志郎	心療科病棟における性的安全の文化の創造に関する研究,(その2)　性的虐待対応看護師チーム（SAR）による性的安全プログラム	小児の精神と神経	51(1)	51-58	2011
杉山登志郎	【貧困とそだち】子どもの貧困とそだち　日本社会的養護と子どもの貧しさ	そだちの科学	16	8-14	2011
杉山登志郎	広汎性発達障害疑い例へのプライマリケアでの対応	日本医事新法	4543	59-60	2011
杉山登志郎	【施設保護を受けた子のトラウマ】発達障害とアタッチメント障害	トラウマティック・ストレス	9(1)	25-31	2011
杉山登志郎	子ども虐待　子どもの命とこころを守る	心と社会	42(1)	12-15	2011
杉山登志郎、原仁、山根希代子、藤坂龍司、野邑健二、今本繁、富永亜由美、並木典子、明翫光宜、野村香代、天野美鈴、有光興記、田島志保、加藤康子、藤田佑里子	早期療育の成果に関する前方向視的研究	乳幼児医学・心理学研究	20(2)	115-125	2011
Shimmura C, Suda S, Tsuchiya KJ, Hashimoto K, Ohno K, Matsuzaki H, Iwata K, Matsumoto K, Wakuda T, Kameno Y, Suzuki K, Tsujii M, Nakamura K, Takei N, Mori N	Alteration of plasma glutamate and glutamine levels in children with high-functioning autism.	PLoS One	6(10)	e25340	2011

研究成果の刊行に関する一覧表(11 / 13)

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
長峰 伸治・加藤 麻登佳・辻井 正次	定型発達児･者との比較による高機能広汎性発達障害児･者の対人交渉方略の検討	聖隷クリストファー大学看護学部紀要	19	27-40	2011
明翫光宜・飯田愛・森一晃・堀江奈央・稲生慧・中島俊思・辻井正次	広汎性発達障害児を対象とした「気分は変えられる」プログラム作成の試み	小児の精神と神経	51(4)	377-385	2011
川上ちひろ・辻井正次	思春期広汎性発達障害男児への性教育プログラムの検討：試行的実践からの分析	小児保健研究	70(3)	402-411	2011
辻井正次	子どもたちの「できること」を伸ばす--発達障害のある子どものスキル・トレーニング実践(12・最終回)楽しい生活のために必要なこと	こころの科学,	157	116-121	2011
松岡弥玲・岡田涼・谷伊織・大西将史・中島俊思・辻井正次	養育スタイル尺度の作成：発達的変化とADHD傾向との関連から	発達心理学研究	22(2)	179-188	2011
明翫光宜・望月知世・内田裕之・辻井正次	広汎性発達障害児の人物画研究(1)：DAM項目による身体部位表現の分析	小児の精神と神経	51(2)	157-168	2011
原幸一・神谷美里・辻井正次	高機能広汎性発達障害児のバウムテストの発達特徴	発達障害研究	33(3)	314-321	2011
井上雅彦・岡田涼・野村和代・上田暁史・安達潤・辻井正次・大塚晃・市川宏伸	知的障害者入所更生施設利用者における強度行動障害とその問題行動の特性に関する分析	精神医学	53(7)	639-645	2011
辻井正次	発達障害のある子どもたちの家庭と学校(5)特別支援学級に在籍すること･通常学級に在籍すること	子どもの心と学校臨床	5	89-97	2011
宮地泰士・神谷美里・野村香代・吉橋由香・辻井正次	広汎性発達障害児本人への診断説明(告知)に関する親の意識と実態調査	精神科治療学	26(11)	1465-1472	2011
辻井正次	発達障害への支援～ライフステージに応じて～青年期の支援(生活・就労支援)	第52回日本児童青年精神医学会総会抄録集		121	2011
辻井正次	成人期アスペルガー症候群の社会適応支援—ライフプランニング・スキルに関連して	第107回日本精神神経学会総会プログラム・抄録集		S.338	2011
黒田美保・稲田尚子	Autism Diagnostic Observation Schedule (自閉症診断観察検査) 日本語版の開発状況と今後の課題	精神医学	54	427-433	2011
安達潤	地域での発達障害に対する一貫した支援のあり方について	児童青年精神医学とその近接領域	52(3)	280-288	2011
安達潤	ＰＡＲＳ短縮版の作成と評定における観点	乳幼児医学・心理学研究	20(2)	83-88	2011

研究成果の刊行に関する一覧表（ 12 / 13 ）

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Thanseem I, Nakamura K, Miyachi T, Toyota T, Yamada S, Tsujii M, Tsuchiya KJ, Anitha A, Iwayama Y, Yamada K, Hattori E, Matsuzaki H, Matsumoto K, Iwata Y, Suzuki K, Suda S, Kawai M, Sugihara G, Takebayashi K, Takei N, Ichikawa H, Sugiyama T,et al.	Further evidence for the role of MET in autism susceptibility	Neurosci Res	68(2)	137-141	2010
Munesue T, Yokoyama S, Nakamura K, Anitha A, Yamada K, Hayashi K, Asaka T, Liu HX, Jin D, Koizumi K, Islam MS, Huang JJ, Ma WJ, Kim UH, Kim SJ, Park K, Kim D, Kikuchi M, Ono Y, Nakatani H, Suda S, Miyachi T, Hirai H, Salmina A, Pichugina YA, Soumarokov AA, Takei N, Mori N, Tsujii M, Sugiyama T, et al.	Two genetic variants of CD38 in subjects with autism spectrum disorder and controls	Neurosci Res	67(2)	181-191	2010
Liu X, Kawamura Y, Shimada T, Otowa T, Koishi S, Sugiyama T, et al.	Association of the oxytocin receptor (OXTR) gene polymorphisms with autism spectrum disorder (ASD) in the Japanese population	J Hum Genet	55(3)	137-141	2010
Fujita-Shimizu A, Suzuki K, Nakamura K, Miyachi T, Matsuzaki H, Kajizuka M, Shinmura C, Iwata Y, Suda S, Tsuchiya KJ, Matsumoto K, Sugihara G, Iwata K, Yamamoto S, Tsujii M, Sugiyama T,et al.	Decreased serum levels of adiponectin in subjects with autism	Prog Neuropsychopharmacol Biol Psychiatry	16；34(3)	455-458	2010
杉山登志郎	タイムスリップ現象再考	精神科治療学	25(12)	1639-1645,	2010
杉山登志郎,	性的虐待へのケア	日本小児科学会雑誌	114(10)	1526-1533	2010
杉山登志郎	小児慢性患者におけるこころの臨床のニード	そだちの科学	15	62-65	2010
野村和代, 鈴木将文, 井上雅彦, 杉山登志郎	強度行動障害の再検討(その 1) 強度行動障害特別処遇事業における対象事例の支援・経過についての分析	小児の精神と神経	50(3)	291-296	2010
杉山登志郎, 川村昌代, 橋詰由加里, 大隅香苗	強度行動障害の再検討(その 2) 厚生科学研究における強度行動障害研究の再検討	小児の精神と神経	50(3)	247-257	2010
川村昌代, 杉山登志郎	発達精神病理学的視点からみた広汎性発達障害	臨床心理学	増刊 2	25-30	2010
杉山登志郎	広汎性発達障害の現在	子どものこころと脳の発達	1(1)	9-18	2010
杉山登志郎	子ども虐待への医療機関を核とした子どもと親へのケア	月刊福祉	93(11)	32-35	2010
松本慶太, 杉山登志郎	被虐待児の治療	精神科	17(1)	36-41	2010
杉山登志郎	高機能広汎性発達障害	日本医事新報	4500	70-75	2010
杉山登志郎, 山村淳一	ADHD と子ども虐待	精神科治療学	25(6)	803-808	2010
杉山登志郎	ADHD に対するアトモキセチンの臨床	脳 21	13(2)	186-191	2010

研究成果の刊行に関する一覧表(13 / 13)

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
杉山登志郎	いじめ・不登校と高機能広汎性発達障害	こころの科学	151	64-69	2010
鈴木善統，杉山登志郎	自閉症・アスペルガー障害の早期診断と鑑別診断	小児内科	42(3)	431-434	2010
宇野洋太、内山登紀夫	医療機関における ASD 診療の課題と展望	Monthly Book Medical Rehabilitation	125	38-42	2010
蜂矢百合子、吉田友子	自閉症スペクトラムの子の「察する」力　理解と支援（特集　人の気持ちがわからない子）」	児童心理	64(10)	825-830	2010
吉田友子	自閉症スペクトラムの子どもへの医学心理学教育(告知)	児童青年精神医学とその近接領域	51(3)	281-289	2010
吉田友子	【螺旋　共生社会への歩み】自閉症スペクトラムの子どもへの医学心理学教育	児童青年精神医学とその近接領域	5(3)	281-289	2010
KurodaMiho、WakabayashiAkio、UchiyamaTokio、KamioYoko、YoshidaYuko	Which is Better, Visual or Auditory Information, for Adults with Autism Spectrum Disorders to Read the Mind?	国立精神・神経センター精神保健研究所年報	22	82-83	2010
田ノ岡 志保 ， 辻井 正次	子どもたちの「できること」を伸ばす--発達障害のある子どものスキル・トレーニング実践双方向コミュニケーションを学ぶ	こころの科学	151	128-134	2010
纐纈 えみ ， 辻井 正次	子どもたちの「できること」を伸ばす--発達障害のある子どものスキル・トレーニング実践(7)日常生活快適化スキルを学ぶ	こころの科学	152	107-112	2010
細溝 さやか ， 辻井 正次	子どもたちの「できること」を伸ばす--発達障害のある子どものスキル・トレーニング実践(8)こだわりの調整の仕方を知る--視点を移すスキル	こころの科学	153	108-113	2010

Ⅳ. 研究成果の刊行物・別刷

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

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

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

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 $\geq$80 %, and excellent if they were $\geq$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

both the consensus clinical diagnosis and the algorithm diagnosis. The caregivers were then provided with feedback, including a best-estimate diagnosis.

Results

Reliability Study

No single diagnostic algorithm item showed a weighted kappa (Kw) of lower than 0.6 (see Appendix Table 1 in supplementary materials). Two items showed Kw values at the level of "good" in terms of clinical significance (0.74 for item 39, "Verbal rituals", and 0.69 for item 58, "Inappropriate facial expression"), but the remaining 40 out of 42 diagnostic algorithm items showed Kw values of 0.75 or higher, indicating a level of excellent clinical significance.

All domains and subdomains showed ICC values of 0.75 or higher, indicating an excellent level (Table 3). ICC values were again calculated separately for three age bands (<5:0 years, 5:0–9:11 years, and 10–19 years). Among individuals below 5 years of age, all domains and subdomains had ICC values of ≥0.75 (excellent). For individuals between 5:0 and 9:11 years, all domains and all but one subdomain had ICC values of ≥0.75 (excellent); one exception was subdomain C3, "Stereotyped and repetitive motor mannerisms", which showed an ICC value of 0.73 (good). For those individuals 10 years of age and older, all domains and all but two subdomains showed ICC values of ≥0.75 (excellent); the exceptions were 0.69 for subdomain B2 (V), "Relative failure to initiate or sustain conversational interchange", and 0.62 for subdomain C4, "Preoccupations with part of objects or non-functional elements of material", which had ICC values over 0.6, but below 0.75 (good).

Validity Study

Discriminant Validity: Difference in Mean Scores of Items/Subdomains/Domains Across Three Diagnostic Groups

As regards the mean scores for diagnostic algorithm items (Table 4), all items but one showed a clear, significant difference across the three diagnostic groups using one-way ANOVA (AD vs. PDDNOS vs. non-ASD, $p < 0.001$

Table 3 Inter-rater reliability: intraclass correlation coefficients (ICC) of ADI-R domain and subdomain scores across three age bands (N = 51)

Domain/sub-domain code	Item	ICC all subjects [N = 51]	ICC <5:0 years [N = 20]	ICC 5:0–9:11 years [N = 15]	ICC 10–19 years [N = 16]
A	Qualitative abnormalities in reciprocal social interaction	.96	.93	.97	.95
A1	Failure to use nonverbal behaviors to regulate social interaction	.92	.91	.94	.91
A2	Failure to develop peer relationships	.95	.92	.92	.90
A3	Lack of shared enjoyment	.96	.94	.98	.97
A4	Lack of socioemotional reciprocity	.91	.93	.89	.88
B	Qualitative abnormalities in communication	.97	.95	.96	.98
B1	Lack of, or delay in, spoken language and failure to compensate through gesture	.93	.94	.91	.92
B4	Lack of varied spontaneous make-believe or social imitative play	.96	.93	.97	.98
B2(V)	Relative failure to initiate or sustain conversational interchange	.92	.90	.92	.69
B3(V)	Stereotyped, repetitive, or idiosyncratic speech	.92	.96	.95	.77
C	Restricted, repetitive, stereotyped patterns of behaviour	.95	.96	.96	.87
C1	Encompassing preoccupation or circumscribed pattern of interest	.94	.97	.92	.81
C2	Apparently compulsive adherence to non-functional routines or rituals	.86	.85	.90	.81
C3	Stereotyped and repetitive motor mannerisms	.86	.85	.73	.96
C4	Preoccupations with part of objects or non-functional elements of material	.82	.89	.94	.62

Table 4 Discriminant validity: mean scores of diagnostic algorithm items, subdomains, and domains

Items	(1) AD [N = 138]	(2) PDDNOS [N = 89]	(3) Non-ASD [N = 90]	Statistics
A1. Failure to use nonverbal behaviors to regulate social interaction	3.8 (1.7)	2.6 (2.0)	0.2 (0.6)	$F(2, 314) = 138.4$, $p < 0.001$ $1 > 3$: $p < 0.001$ $2 > 3$: $p < 0.001$ $1 > 2$: $p < 0.001$
50. Direct gaze	1.5 (0.9)	1.1 (1.0)	0.0 (0.3)	$F(2, 227) = 61.5$, $p < 0.001$ $1 > 3$: $p < 0.001$ $2 > 3$: $p < 0.001$ $1 > 2$: $p = 0.002$
51. Social smiling	1.9 (1.1)	1.4 (1.2)	0.1 (0.4)	$F(2, 230) = 60.5$, $p < 0.001$ $1 > 3$: $p < 0.001$ $2 > 3$: $p < 0.001$ $1 > 2$: $p = 0.01$
57. Range of facial expressions used to communicate	1.2 (1.0)	0.8 (1.0)	0.0 (0.1)	$F(2, 231) = 34.1$, $p < 0.001$ $1 > 3$: $p < 0.001$ $2 > 3$: $p < 0.001$ $1 > 2$: $p = 0.03$
A2. Failure to develop peer relationships	5.7 (1.9)	4.4 (2.1)	0.7 (1.1)	$F(2, 314) = 226.5$, $p < 0.001$ $1 > 3$: $p < 0.001$ $2 > 3$: $p < 0.001$ $1 > 2$: $p < 0.001$
49. Imaginative play with peers	2.1 (0.9)	1.9 (1.0)	0.2 (0.6)	$F(2, 224) = 95.4$, $p < 0.001$ $1 > 3$: $p < 0.001$ $2 > 3$: $p < 0.001$ $1 > 2$: NS
62. Interest in children	1.9 (1.1)	1.4 (1.1)	0.1 (0.4)	$F(2, 229) = 63.1$, $p < 0.001$ $1 > 3$: $p < 0.001$ $2 > 3$: $p < 0.001$ $1 > 2$: $p = 0.01$
63. Response to approaches of other children	1.3 (0.9)	1.1 (0.8)	0.1 (0.3)	$F(2, 226) = 50.25$, $p < 0.001$ $1 > 3$: $p < 0.001$ $2 > 3$: $p < 0.001$ $1 > 2$: NS
64. Group play with peers	2.2 (0.8)	1.8 (0.9)	0.4 (0.7)	$F(2, 221) = 94.7$, $p < 0.001$ $1 > 3$: $p < 0.001$ $2 > 3$: $p < 0.001$ $1 > 2$: $p < 0.001$
65. Friendships	1.6 (1.1)	1.7 (0.9)	0.2 (0.5)	$F(2, 139) = 19.4$, $p < 0.001$ $1 > 3$: $p < 0.001$ $2 > 3$: $p < 0.001$ $1 > 2$: NS
A3. Lack of shared enjoyment	4.3 (1.7)	3.7 (1.8)	0.7 (1.2)	$F(2, 314) = 146.3$, $p < 0.001$ $1 > 3$: $p < 0.001$ $2 > 3$: $p < 0.001$ $1 > 2$: $p = 0.006$
52. Showing and directing attention	1.5 (1.2)	1.0 (1.1)	0.0 (0.3)	$F(2, 229) = 39.3$, $p < 0.001$ $1 > 3$: $p < 0.001$ $2 > 3$: $p < 0.001$ $1 > 2$: $p = 0.01$