RESEARCH ARTICLE

Parent and Clinician Agreement regarding Early Behavioral Signs in 12- and 18-Month-Old Infants at-Risk of Autism Spectrum Disorder

Lori-Ann R. Sacrey, Lonnie Zwaigenbaum, Susan Bryson, Jessica Brian, Isabel M. Smith, Wendy Roberts, Peter Szatmari, Tracy Vaillancourt, Caroline Roncadin, and Nancy Garon

Parent and clinician agreement regarding early behavioral signs of Autism Spectrum Disorder (ASD) in children from a high-risk cohort (siblings of children diagnosed with ASD, $n = 188$) was examined. Infants were assessed prospectively at 12 and 18 months of age using the clinician administered Autism Observational Scale for Infants (AOSI) and the Autism Parent Screen for Infants (APSI) and underwent a blind independent diagnostic assessment for ASD at 36 months of age. Direct comparison of parent and clinician ratings showed poor agreement on all early behavioral signs, with parent-reported symptoms being better able to differentiate between children with and without ASD at both 12 and 18 months of age compared to clinician observations during a brief office visit. The results suggest that parents may detect some clinically informative behaviors based on their day-to-day observations more readily than do clinicians during brief clinical assessments, a result that needs to be replicated in a non-sibling cohort. ***Autism Res*** *2018, 0: 000–000.* © 2018 International Society for Autism Research, Wiley Periodicals, Inc.

Lay Summary: Parents of children at high-risk of autism spectrum disorder (ASD; have an older sibling with ASD) and clinicians were compared on their reporting of 19 early signs of autism. Direct comparison of parent and clinician ratings showed poor agreement on all early behavioral signs, with parent-reported symptoms being better able to differentiate between children with and without ASD at both 12 and 18 months of age compared to clinician observations during a brief office visit. This suggests that parents may have important information regarding early development of their high-risk child.

Keywords: autism; autism spectrum disorder; rater agreement; behavioral signs; assessment; infant sibling

Introduction

Identifying early signs of Autism Spectrum Disorder (ASD) in children is crucial to ensure timely access to needed services, and thus has the potential to improve functional outcomes [Dawson et al., 2010; Perry et al., 2008; Sallows & Graupner, 2005]. ASD frequently goes undiagnosed until 4 years of age or later [Daniels & Mandell, 2013], yet parents often report that they were concerned about their children's development before the second birthday [Chakrabarti & Fombonne, 2005]. Understanding early development in children who will later be diagnosed with ASD is important to facilitate diagnosis and entry into early intervention, and indeed to inform the development of feasible, cost-effective interventions that target early emergent signs of the disorder [Brian, Bryson, & Zwaigenbaum, 2015].

Prospective studies of children who are at a heightened risk for ASD (e.g., younger siblings of children with ASD) have generated important advances in identifying early behavioral signs of the disorder [for a review, see Jones, Gliga, Bedford, Charman, & Johnson, 2014]. This methodological approach can be further strengthened by incorporating multiple respondents (e.g., clinician and parent) to enrich data collection on early behavioral development. Incorporating multiple raters (e.g., clinician and parent) can increase the comprehensiveness of data collected, by providing multiple perspectives regarding the same behavior [Achenbach,

From the Department of Pediatrics, University of Alberta, Edmonton, Alberta (L.R.S., L.Z.); Autism Research Centre, Glenrose Rehabilitation Hospital, Edmonton, Alberta (L.R.S., L.Z.); Dalhousie University/IWK Health Centre, Halifax, Nova Scotia (S.B., I.M.S.); Bloorview Research Institute, Toronto, Ontario (J.B.); University of Toronto, Toronto, Ontario (J.B., P.S., C.R.); ISAND, Toronto, Ontario (W.R.); The Hospital for Sick Children, Toronto, Ontario (P.S.); Centre for Addiction and Mental Health, Toronto, Ontario (P.S.); University of Ottawa, Ottawa, Ontario (T.V.); McMaster Children's Hospital/Hamilton Health Sciences, Hamilton, Ontario (C.R.); Mount Allison University, Sackville, New Brunswick (N.G.)

Received October 10, 2017; accepted for publication December 27, 2017

Address for correspondence and reprints: Lori-Ann R. Sacrey, Ph.D., Autism Research Centre (E209) 10230 - 111 Avenue, Glenrose Rehabilitation Hospital, Edmonton, Alberta T5G 0B7 Canada. E-mail: sacrey@ualberta.ca

Published online 00 Month 2018 in Wiley Online Library (wileyonlinelibrary.com)
DOI: 10.1002/aur.1920

Table 1. Clinical Characteristics of High-Risk (HR) Sample

	HR-ASD	HR-N	Statistics	
Characteristics	Mean (SD)	Mean (SD)	*t*	*P*
Age				
12-month visit	12.35 (0.43)	12.35 (0.39)	.037	.97
18-month visit	18.44 (.55)	18.40 (.41)	.83	.41
36-month visit	39.46 (3.82)	39.99 (3.76)	.89	.38
AOSI				
12-month total	7.27 (4.84)	5.06 (3.32)	3.20	.002
18-months total	8.32 (4.18)	5.12 (3.89)	5.02	.001
APSI				
12-month total	10.87 (7.45)	5.97 (5.12)	4.33	.001
18-month total	8.48 (8.33)	3.73 (3.67)	4.09	.001
MSEL Standard Scores				
Visual reception	102.90 (27.85)	116.49 (16.87)	3.49	.001
Expressive language	93.35 (15.36)	105.03 (16.07)	4.71	.001
Receptive language	93.26 (18.39)	104.62 (14.77)	4.51	.001
Fine motor	88.03 (18.83)	101.03 (20.39)	4.15	.001
Early learning composite	94.70 (20.34)	109.29 (17.09)	5.01	.001
VABS Standard Scores				
Communication	89.95 (13.58)	102.54 (12.95)	5.83	.001
Daily living skills	80.74 (13.14)	93.29 (11.80)	6.20	.001
Socialization	80.13 (11.70)	93.89 (12.31)	6.82	.001
Motor	85.46 (14.95)	95.67 (12.72)	4.14	.001
Adaptive behavior composite	80.15 (12.76)	95.14 (12.78)	7.06	.001
ADOS Severity Scores				
Social affect	6.18 (1.88)	2.75 (1.83)	11.94	.001
Restricted/repetitive behavior	7.40 (2.04)	4.68 (2.51)	7.92	.001
Overall	6.35 (1.88)	2.53 (1.71)	13.86	.001
ADI-R				
Total score	22.83 (9.97)	6.02 (4.65)	12.24	.001

Abbreviations: ADI-R – Autism Diagnostic Interview-Revised; ADOS – Autism Diagnostic Observation Schedule; AOSI – Autism Observation Scale for Infants; APSI – Autism Parent Screen for Infants; HR-ASD – High-risk infant with autism spectrum disorder; HR-N – High-risk infant without autism spectrum disorder; MSEL – Mullen Scales of Early Learning; VABS – Vineland Adaptive Behavior Scales; *t*: independent samples *t* statistic, *P*: *P* value.

McConaughy, & Howell, 1987; Holmbeck, Li, Schurman, Friedman, & Coakley, 2002; Schniering, Hudson, & Rapee, 2000]. Parents provide information based on first hand knowledge of the child's behavior across varied contexts, whereas clinicians generally have greater knowledge of the normative behavior of children [Klein, Lavigne, & Seshadri, 2010], but a briefer period of observation, typically in a single context (e.g., clinic or lab environment). Developmental diagnoses generally correspond with parent-reported concerns about their child's development [Glascoe, 2003; Glascoe, Foster, & Wolraich, 1997; Sacrey et al., 2015; Ozonoff et al., 2009]. However, no direct comparison of parent and clinician reports of early behavioral signs of ASD has been reported.

The purpose of this study was to examine parent and clinician agreement regarding ratings of behavior related to early symptoms of ASD in a high-risk infant cohort (younger siblings of children diagnosed with ASD). Parents completed the Autism Parent Screen for Infants [APSI; Bryson, Zwaigenbaum, Brian, & Roberts, 2006; Sacrey et al., 2016] and clinicians completed the Autism Observation Scale for Infants [AOSI; Bryson, Zwaigenbaum, McDermott, Rombough, & Brian, 2008] at 12 and 18 months of age. The APSI is modelled in content on the AOSI, the two measures sharing 19 items for potential comparison.

Methods

Participants

Infant siblings of children with ASD (high-risk infants; HR) were recruited between the ages of 6 and 12 months from families participating in longitudinal research at one of four multidisciplinary ASD diagnostic and treatment centers in Canada: Halifax, Toronto, Hamilton, and Edmonton. The research ethics board at each institution approved this study and all families gave written informed consent at enrollment. All participants were born between 36 and 42 weeks gestation and had a birth weight greater than 2,500 grams. Diagnosis of ASD in the older sibling (i.e., proband) was confirmed by a clinical assessment or a review of diagnostic records, using DSM-IV-TR criteria. Neither the high-risk (HR) infant siblings nor the probands had identifiable neurological or genetic conditions, or severe sensory or motor impairments.

Children from our larger HR cohort were included in this study if (a) they had undergone a 3-year diagnostic assessment and (b) data on both an APSI and an AOSI were available at 12 and/or 18 months of age. Of the 402 HR children with 3-year follow-up, 188 HR children had complete APSI and AOSI data at either 12 or 18 months, or both. Those that did not have complete date at both assessments at either age were excluded from further analyses. The excluded children either had their completed 18-month assessment prior to inclusion of the APSI in the larger longitudinal study protocol ($n = 178$), or did not have a completed APSI returned by the parents at either time-point ($n = 36$), for a participation rate of 85%. Table 1 presents detailed participant characteristics.

Measures

Early behavioral measures of ASD signs were obtained at 12–13 months (hereafter, "12-month assessment") and 18–19 months of age (hereafter, "18-month assessment"). Parent reports of early signs were measured by the APSI and clinician ratings of early signs were measured using the AOSI. Behavioral, developmental, and diagnostic outcomes were evaluated at 36–42 months

(hereafter, "3 year-assessment") using the Mullen Scales of Early Learning (MSEL), the Vineland Adaptive Behavior Scales (VABS), the Autism Diagnostic Observation Schedule (ADOS), and the Autism Diagnostic Interview-Revised (ADI-R).

Measuring early signs of ASD at 12 and 18 months. The *Autism Parent Screen for Infants* [APSI; Bryson et al., 2006] is a 26-item forced-choice parent-report questionnaire with content and format similar to those of the AOSI [Bryson et al., 2008]. It covers a wide range of pre-diagnostic behavioral symptoms, including atypical eye contact, visual tracking, responding to name, imitation, language, social development, joint attention, gestures, play, visual examination of objects, and emotional regulation. For example, to the question, "Does your child use gestures, such as waving goodbye, nodding his/her head, or blowing a kiss?" response choices are "definitely," "possibly," or "no," which are scored "0," "1," and "2," respectively. The APSI is designed to monitor putative signs of ASD in infants aged 6–24 months and takes approximately 10–15 min for the primary caregiver to complete (~85% mother).

The *Autism Observation Scale for Infants* [AOSI; Bryson et al., 2008] is a semi-structured direct observational measure designed to identify early behavioral markers of ASD in infants/toddlers (e.g., atypicalities in social communication, engagement, affect sharing, attention, and behavioral regulation). The AOSI is designed for use with 6- to 18-month-olds and administration takes approximately 15–20 min. Each behavior is rated on a scale from 0 to 2 or 3, where 0 implies typical function, and higher values indicate increasing atypicality. The AOSI has excellent inter-rater reliability (0.93 for total score at 12 months), fair-to-good test-retest reliability at 12 months [0.61 for total score: Bryson et al., 2008], and good predictive validity at 12 months for its original 16 items [Zwaigenbaum et al., 2016].

The APSI and AOSI share 19 items that measure putative signs of ASD: visual tracking, visual fixation, responding to name, reacting to a change in facial expression, anticipation during a social interaction, imitation, vocalizing back and forth, eye contact, social smiling, coordinating actions and eye gaze, reactivity, showing interest and pleasure, transitioning, difficulty using their hands, repetitive motor behaviors, unusual sensory behaviors, focusing attention, insistence on having the same toy, and sharing interest with others.

Developmental and diagnostic assessments at 3 years. The *Mullen Scales of Early Learning* [MSEL; Mullen, 1995] consists of five scales, four of which (Visual Reception, Receptive Language, Expressive Language, and Fine Motor) assess nonverbal, cognitive, and language abilities, while the fifth scale measures gross motor development (from 0 to 29 months only). An Early Learning Composite is calculated based on scores from the first four scales for children aged 0–69 months. Inter-rater and test-retest reliability are excellent [Mullen, 1995].

The *Vineland Adaptive Behavior Scales* [VABS; Sparrow, Balla, & Cicchetti, 1984] is a semi-structured parent interview designed to assess adaptive behavior across four subdomains—Communication, Daily Living, Socialization, and Motor skills (the last domain limited to children younger than 30 months), outlined by typical developmental milestones that are anchored to specific ages. The scale has excellent reliability and concurrent validity, and is sensitive to impairments experienced by children with ASD [Volkmar, Carter, Sparrow, & Cicchetti, 1993; Carter et al., 1998].

The *Autism Diagnostic Observation Schedule* [ADOS; Lord et al., 2000] includes standardized activities and "presses," which are used to elicit communication, social interaction, imaginative use of play materials, and repetitive behavior [Lord et al., 1989]. Inter-rater reliability for the ADOS is excellent [Lord et al., 2000]. The scoring algorithm was recently revised to optimize discrimination of ASD from other developmental disabilities and is organized into two domains, Social Affect (including Communication and Social items), and Restricted Repetitive Behaviors [Gotham, Risi, Pickles, & Lord, 2007]. The ADOS consists of four modules, each of which is appropriate for individuals of differing language levels (Module 1 = minimal or no language, Module 2 = regular use of non-echoed 3-word phrases, Module 3 = child with fluent language; and Module 4 = adolescent or adult with fluent language), the first three of which were used to assess participants in this study. To optimize comparability across modules (and thus, across language levels), we used the ADOS severity metric [Gotham, Pickles, & Lord, 2009].

The *Autism Diagnostic Interview-Revised* [ADI-R; Lord, Rutter, & LeCouteur, 1994] is an investigator-directed interview that elicits information regarding social development, verbal and nonverbal communication skills, and the presence of repetitive, stereotyped interests and behavior required to make an ICD-10 or DSM-IV-TR diagnosis of ASD. The questions are designed to distinguish qualitative impairments from developmental delays. The ADI-R discriminates well between ASD and other forms of developmental disability, and inter-rater reliability is excellent [Lord et al., 1994].

Diagnostic Procedure

At 3 years of age, each participant underwent an independent diagnostic evaluation, conducted by an expert clinician blind to results from previous AOSI and APSI

findings. ASD diagnoses were assigned using DSM-IV-TR criteria, based on the best judgment of the clinician (developmental pediatrician, child psychiatrist, or clinical psychologist, all with at least 10 years of diagnostic experience), taking into account information from the ADI-R and ADOS, as well as concurrent developmental assessment using the MSEL and VABS.

Statistical Analysis

Group membership was determined from the 3-year outcome assessments: high-risk infants diagnosed with ASD (HR-ASD) and high-risk infants not diagnosed with ASD (HR-N). Clinical characteristics of the groups were compared using independent *t*-tests with Group (HR-N, HR-ASD) as the independent variable and scores on the various assessments as the dependent variables.

Comparisons for the 19 items that are shared between the AOSI and APSI were completed at 12 and 18 months of age using three analytic approaches. First, parent and clinician agreement was determined using intraclass correlations (ICC) with a two-way mixed model evaluating absolute agreement on item-level scores. Cicchetti's [1994] guidelines classify ICC scores of less than 0.40 as "poor," between 0.40 and 0.59 as "fair," between 0.60 and 0.74 as "good," and between 0.75 and 1.00 as "excellent." Second, group differences were examined using independent *t*-tests for each measure to determine which items differentiated HR infants with and without 3-year ASD diagnoses. Third, relative risk was calculated for total score on the APSI and AOSI at 12 and 18 months of age using the receiver operator characteristics (ROC) cut-off scores derived from the larger HR cohort [Sacrey et al., 2016; Zwaigenbaum et al., in prep, respectively]. All statistical analyses were completed using the Statistical Package for the Social Sciences (SPSS v. 23) with $P < .05$ as statistically significant.

Regression analyses using the PROCESS macro developed by Hayes [2013] were completed to determine any moderating effect of family demographics on the relation between AOSI and APSI total scores at 12 and 18 months and diagnostic outcomes at 36 months. Family demographics included in the analyses were participant's birth order, number of children in the family, father's and mother's age at participant's birth, and family socioeconomic status

Results

Completers versus Non-Completers

An independent samples *t*-test was completed to determine if autism symptomology differed between the high-risk children who were included in this analysis (completers) versus those who were not (non-completers). There were no group differences for ADOS severity score ($t = 1.70$, $P = .091$), or the ADI-R total ($t = 0.67$, $P = .50$) at 36 months.

Clinical Characteristics

Based on the 3-year diagnostic assessments, two groups were identified for comparison: (a) HR infant siblings who received a diagnosis of ASD ("HR-ASD"; $n = 59$; 44 boys and 15 girls), and (b) HR infant siblings who did not receive a diagnosis of ASD ("HR-N"; $n = 129$; 67 boys and 62 girls). There was a significant sex difference ($\chi^2 = 8.22$, $P = .006$), with a higher ratio of boys than girls in the HR-ASD (2.92 boys: 1 girl) group compared to the HR-N group 1.08 boys: 1 girl). The groups did not differ on exact age at the 12-month ($t(184) = .037$, $P = .97$), 18-month ($t(171) = .83$, $P = .41$) or 36-month assessments ($t(188) = .89$, $P = .38$).

Descriptive data on developmental and behavioral features are summarized for the two groups in Table 1. The groups differed on AOSI and APSI scores at 12 and 18 months, with the HR-ASD group having higher scores than the HR-N group on both. Standard scores on all MSEL and VABS subscales differed significantly between the groups, with the HR-ASD group having lower scores. Groups also differed for ADOS severity scores and ADI-R total, with the HR-ASD having higher scores.

Rater Agreement

The two HR groups (both combined and separately) were compared on the 19 shared items at 12 and 18 months. As shown in Table 2, ICCs at both time-points indicated poor agreement between parents and clinicians. At 12 months, ICCs for all HR siblings combined ranged from .001 to .23 (all "poor"), with significant (i.e., non-zero) agreement for six items: responding to name, anticipation during a social interaction, imitation, eye contact, and difficulty using hands (P's $< .05$). ICCs for the HR-ASD group alone ranged −.02 to .30 (all "poor"), with significant agreement for three items: visual fixation, responding to name, and difficulty using hands (P's $< .05$). ICCs for the HR-N group ranged from −.004 to .19 (all "poor"), with significant agreement for two items, responding to name and anticipation during a social interaction (P's $< .05$).

At 18 months, ICCs for all HR siblings combined ranged from −.01 to .21 (all "poor"), with significant agreement for seven items: responding to name, reacting to a change in facial expression, vocalizing back and forth, eye contact, difficulty using their hands, repetitive motor behavior, and unusual sensory behavior (P's $< .05$). ICC's for the HR-ASD group ranged .01 to .23 (all "poor"), with significant agreement for two items, repetitive motor behavior and unusual sensory

Table 2. Intraclass Correlation Coefficients for Individual Items on the AOSI and APSI

	12 Months						18 Months					
Question item	All HR	95%CI	HR-ASD	95%CI	HR-N	95%CI	All HR	95%CI	HR-ASD	95%CI	HR-N	95%CI
Visual tracking	−.06	−.19, .07	−.06	−.29, .18	−.05	−.21, .11	−.03	−.16, .10	−.02	−.25, .22	−.04	−.19, .11
Visually fixate	.09	−.05, .21	.26[b]	.04, .47	−.02	−.18, .14	−.01	−.14, .12	.01	−.23, .25	−.04	−.19, .11
Respond to name	.23[c]	.11, .35	.22[a]	.01, .42	.14[a]	−.01, .29	.21[c]	.08, .33	.19	−.05, .41	.04	−.12, .19
React to facial change	.08	−.05, .20	.11	−.12, .34	.05	−.09, .19	.18[b]	.04, .31	.14	−.13, .39	.19[b]	−.03, .34
Anticipate social interaction	.15[b]	.02, .28	.09	−.15, .32	.19[b]	.03, .33	.04	−.09, .17	.03	−.21, .26	.04	−.12, .19
Imitate	.11[a]	−.02, .24	.08	−.15, .30	.13	−.03, .27	.09	−.04, .21	−.03	−.27, .20	.15[a]	−.01, .30
Vocalize back and forth	.02	−.07, .12	−.02	−.20, .18	.03	−.08, .15	.12[a]	−.01, .24	.19	−.05, .40	.02	−.13, .18
Eye contact	.13[b]	.01, .25	.13	−.08, .34	.08	−.06, .22	.14[a]	.01, .26	.13	−.11, .36	.07	−.09, .22
Social smile	.04	−.05, .14	.11	−.07, .30	−.02	−.11, .09	.09	−.04, .21	.10	−.14, .33	.05	−.10, .20
Coordinating actions/eyes	.03	−.10, .16	.006	−.23, .24	−.004	−.16, .15	.02	−.11, .15	.07	−.17, .30	−.03	−.18, .12
Reactivity	.001	−.13, .13	−.03	−.25, .20	.02	−.13, .18	.03	−.10, .16	.08	−.16, .31	.01	−.14, .17
Shows interest and pleasure	.08	−.03, .20	.12	−.09, .33	.05	−.09, .19	.02	−.11, .15	−.08	−.31, .16	.02	−.13, .17
Transitions	.10	−.03, .22	.09	−.14, .31	.10	−.05, .25	.07	−.06, .20	.07	−.17, .30	.05	−.11, .20
Difficulty using hands	.18[b]	.05, .30	.30[b]	.07, .50	.08	−.08, .24	.14[a]	.01, .26	−.02	−.25, .22	.44[c]	.31, .56
Repetitive motor behaviors	.05	−.06, .17	.14	−.07, .35	−.02	−.16, .12	.20[c]	.08, .32	.23[a]	−.01, .44	.17[b]	.01, .31
Unusual sensory behaviors	.02	−.11, .15	.05	−.20, .28	−.06	−.21, .10	−.03	.03, .28	−.02	−.02, .43	−.04	−.14, .16
Focusing attention	.004	−.11, .12	.03	−.17, .23	−.04	−.18, .10	−.01	−.08, .17	.01	−.21, .26	−.04	−.19, .11
Insistence on toy	.12[b]	−.01, .25	.08	−.17, .32	.13	−.03, .28	.21[c]	−.11, .14	.19	−.34, .14	.04	−.07, .23
Share interest	.08	−.04, .21	.07	−.16, .29	.06	−.08, .21	.18[b]	−.08, .17	.14	−.11, .36	.19[b]	−.19, .12

Significance at a=.05; b = .01; c = .001

Abbreviations: AOSI – Autism Observation Scale for Infants; APSI – Autism Parent Screen for Infants; HR – High-risk; HR-ASD – High-risk infant with autism spectrum disorder; HR-N – High-risk infant without autism spectrum disorder; 95% CI – 95% Confidence Intervals.

behavior (P's < .05). ICCs for the HR-N group ranged from −.04 to 44 (all "poor"), with significant agreement for four items: reacting to a change in facial expression, imitation, difficulty using their hands, and repetitive motor behavior (P's < .05).

Group Comparisons on Individual Items

Item-level responses were compared using independent *t*-tests on 12- and 18-month data. As shown in Table 3, at 12 months, three items were informative in predicting diagnostic outcomes on both the AOSI and APSI: responding to name, eye contact, and hand-eye coordination (P's < .05). Nine items were informative on the APSI only: visual fixation, anticipating a social interaction, back-and-forth vocalizations, social smiling, reactivity, repetitive motor behavior, unusual sensory behavior, focusing attention, and sharing interests with others (P's < .05). Seven items were not informative on either assessment: visual tracking, reacting to change in facial expression, imitation, showing interest and pleasure, transitions, difficulty using hands, and insistence on same object (P's > .05). No items were informative on the AOSI only.

At 18 months, seven items were informative in predicting diagnostic outcomes on both the AOSI and APSI. These were responding to name, vocalizing back and forth, eye contact, social smiling, showing interest and pleasure, unusual sensory behaviors, and focusing attention (P's < .05). Eight items were informative on the APSI only: visual fixation, anticipation during a social interaction, imitation, reactivity, transitioning, difficulty using their hands, repetitive motor behavior, and insistence on the same toy (P's < .05), and one item, reacting to a change in facial expression, was informative on the AOSI only (P < .01). Three items were not informative on either assessment: visually tracking, coordinating eyes and hands during action, and sharing interest with others (P's > .05).

Relative Risk of ASD Outcomes

Relative risk of an ASD diagnosis was compared on the APSI and AOSI using the receiver operator characteristics (ROC) cut-off scores that maximized sensitivity and specificity from the larger HR cohort [Sacrey et al., 2016; Zwaigenbaum et al., in prep, respectively]. At 12 months, an AOSI cut-off score of "7" (with scores of 7 and above indicating "positive for ASD") resulted in a relative risk ratio of 1.58 (95% CI = 1.17–2.14; $z = 2.96$, $P = .003$). At 18 months, an AOSI cut-off score of "6" (with scores of 6 and above indicating "positive for ASD") resulted in a relative risk of 1.82 (95% CI = 1.43–2.32; $z = 4.89$; $P = .001$).

At 12 months, an APSI cut-off score of "10" (with scores of 10 and above indicating "positive for ASD") resulted in a relative risk ratio of 3.61 (95% CI = 2.36–5.52; $z = 5.93$, $P = .001$). At 18 months, an APSI cut-off score of "9" (with scores of 9 and above indicating

Table 3. Group Differences for Individual Items on the AOSI and APSI

	12 Months						18 Months					
	APSI			AOSI			APSI			AOSI		
Question item	*t*	*P*	*d*	*t*	*P*	*d*	*t*	*P*	*d*	*t*	*P*	*d*
Visual tracking	1.55	.12	.23	1.12	.24	.16	1.55	.12	.15	1.12	.24	.09
Visually fixate	3.29	.002[b]	.51	0.30	.77	.03	3.29	.002[b]	.42	0.30	.77	.10
Respond to name	3.73	.001[c]	.60	3.57	.001[c]	.54	3.73	.001[c]	.75	3.57	.001[c]	.61
React to facial change	1.45	.15	.22	1.57	.12	.26	1.45	.15	.15	1.57	.12	.44
Anticipate social interaction	2.38	.02[a]	.38	1.54	.13	.25	2.38	.02[a]	.35	1.54	.13	.14
Imitate	1.52	.13	.22	0.60	.55	.09	1.52	.13	.32	0.60	.55	.05
Vocalize back and forth	2.81	.006[b]	.42	0.99	.32	.14	2.81	.006[b]	.31	0.99	.32	.43
Eye contact	3.29	.001[c]	.52	2.42	.02[a]	.36	3.29	.001[c]	.47	2.42	.02[a]	.78
Social smile	2.96	.004[b]	.51	1.76	.08	.24	2.96	.004[b]	.35	1.76	.08	.39
Coordinating actions/eyes	2.08	.04[a]	.31	2.13	.04[a]	.32	2.08	.04[a]	.22	2.13	.04[a]	.21
Reactivity	3.12	.002[b]	.48	0.14	.89	.02	3.12	.002[b]	.46	0.14	.89	.05
Shows interest and pleasure	1.59	.12	.24	1.70	.09	.25	1.59	.12	.50	1.70	.09	.39
Transitions	1.11	.27	.16	0.89	.38	.14	1.11	.27	.34	0.89	.38	.21
Difficulty using hands	1.62	.11	.25	0.02	.98	.00	1.62	.11	.41	0.02	.98	.27
Repetitive motor behaviors	2.36	.02[a]	.39	1.40	.16	.22	2.36	.02[a]	.36	1.40	.16	.21
Unusual sensory behaviors	2.13	.03[a]	.35	1.89	.06	.28	2.13	.03[a]	.51	1.89	.06	.44
Focusing attention	2.35	.02[a]	.36	1.80	.08	.27	2.35	.02[a]	.39	1.80	.08	.57
Insistence on toy	1.68	.10	.25	1.47	.15	.24	1.68	.10	.33	1.47	.15	.18
Share interest	1.99	.05[a]	.31	1.77	.08	.28	1.99	.05[a]	.22	1.77	.08	.29

Significance at a = .05; b = .01; c = .001

Abbreviations: AOSI – Autism Observation Scale for Infants; APSI – Autism Parent Screen for Infants; *t*: independent samples *t* statistic, *P*: *P* value, *d*: Cohen's *d* effect size.

"positive for ASD") resulted in a relative risk of 1.82 (95% CI = 2.34–7.33; $z = 4.89$; $P = .001$).

Combining APSI and AOSI Assessments

Logistic regression and receiver operator curve (ROC) analyses were used to determine the "value added" by combining the AOSI and APSI compared to the APSI alone. Logistic regression was used to calculate the probabilities of identifying group membership (HR-ASD versus HR-N) when combining assessments. The outcome probabilities of the combined assessments were then compared to the APSI alone using ROC analyses at both 12 and 18 months.

At 12 months, there was no "value added" by combining AOSI and APSI ratings. The area under the curve (AUC) for APSI alone was .73 (95% CI .64–.81; $P < .001$). For the combined APSI and AOSI, the AUC remained at .73 (95% CI .64–.81; $P < .001$). At 18 months, however, combining the assessments added value. The AUC for the APSI alone was .69 (95% CI .60–.78; $P < .001$) and combining APSI and AOSI increased the AUC to .77 (95% CI .69–.84; $P < .001$).

Family Demographics as Moderators

Regression analysis was used to determine if family demographics moderated the relation between total scores on the AOSI and APSI at 12 and 18 months and diagnostic outcomes at 36 months. Results indicated that father's age at the participant's birth of the participating child was a significant predictor of diagnostic outcome for both the AOSI at 12 months ($b = -.085$, SE = .034, $P < .05$) and 18 ($b = -.10$, SE = .035, $P < .01$), as well as the APSI at 12 ($b = -.086$, SE = .037, $P < .05$) and 18 ($b = -.079$, SE = .035, $p < .05$) months. Birth order of the participant, number of children in the family, mother's age at participant's birth, and family SES were not significant moderators.

Discussion

We examined parent and clinician agreement on early signs of ASD in an HR infant cohort at 12 and 18 months. There were two main findings. First, a group comparison of ratings on the APSI and AOSI indicated that at both 12 and 18 months of age, indicated that parents endorsed a larger number of items that differentiated children with ASD from those without compared to clinicians. AOSI. Second, agreement was poor between parent and clinician rating of the 19 shared items at both 12 and 18 months of age, as assessed by ICC. Overall, the results suggest that prospective parent reports are informative for early signs of ASD by 12 months of age and add key details regarding early expression of ASD not observed during an interactive clinical assessment.

Parents endorsed more items on the APSI that differentiated HR participants who would be diagnosed with ASD, compared to clinician endorsement on the AOSI. At 12 months of age, only three AOSI items, responding to name, eye contact, and coordinated hand and eye movements, distinguished the two groups. Conversely, 12 items were informative on the APSI, including the three behaviors from the AOSI, plus visual fixation, anticipation during a social interaction, back and forth vocalizations, social smiling, reactivity, repetitive motor behavior, unusual sensory behavior, focused attention, and sharing interest. At 18 months, eight items on the AOSI discriminated between the two groups. These were responding the name, reacting to a change in facial expression, back-and-forth vocalizations, eye contact, social smiling, showing interest and pleasure, unusual sensory behaviors, and focused attention. In contrast, 15 APSI items differentiated between the two groups, including the informative behavior from the AOSI (except reacting to a change in facial expression). This is in line with the calculations of relative risk, indicating that elevated APSI scores increased risk of ASD outcomes more so than elevated AOSI scores, at both 12 and 18 months. That parents are able to identify atypical behavior is consistent with studies examining the predictive utility of parents' concerns about their children's development for diagnostic outcomes [Schertz, Odom, Baggett, & Sideris, 2016; Sacrey et al., 2015; Wetherby et al., 2004; Hess & Landa, 2012; Glascoe, 2003; Robins, Fein, Barton, & Green, 2001]. As such, the APSI may ultimately have greater clinical utility, although evaluation in community settings in both high- and low-risk settings is needed.

Direct comparison of early signs of ASD indicated poor agreement between parents and clinicians. This was surprising given the number of reports of broad agreement amongst raters when identifying atypical behavior [Glascoe, 2000; Glascoe et al., 1997; Sacrey et al., 2015; Verhulst & van der Ende, 1991]. Our study differs, however, not only by informants included (parent and clinician, rather than parent and teacher), but also by context (clinic/laboratory versus home and classroom), observation period (gestalt following brief clinical visit versus everyday experiences at home), and direct coding of specific behaviors compared to reporting general concern. Rater incongruence has a long history in psychological research [De Los Reyes & Kazdin, 2005]. A pioneering meta-analysis of parent and teacher ratings of emotional and behavioral problems indicated poor agreement, with values ranging between .20 and .30 [Achenbach et al., 1987]. With respect to ASD, comparisons of psychiatric diagnoses and parent reports during a neuropsychological interview also resulted in poor agreement across diagnoses, ranging from .06 to .18 [Stadnick et al., 2016]. Furthermore, parent ratings of repetitive behavior on the Repetitive Behavior Scale – Revised were found to be weakly associated with the Repetitive and Restrictive Behavior subscale on the ADOS-Toddler Module [rho = .12; Schertz et al., 2016]. The lack of agreement amongst raters demonstrates the importance of collecting information from many informants, as they provide different perspectives on the same behavior across multiple contexts [Achenbach et al., 1987; Klein et al., 2010; Smith, 2007].

Our study of parent and clinician agreement has several strengths. We have a large sample of HR infants with and without ASD, a prospective design, and parent and clinician reports collected from complementary contemporaneous assessments at two ages. Our study is not without limitations, however. First, by definition, parents of HR infants already had at least one child with a diagnosis of ASD. It may be that these parents are more aware of the types of early behavior that are associated with ASD, resulting in higher vigilance for developmental and behavioral differences. Second, our analyses are based on a subsample of children who had both assessments completed at 12 and/or 18 months. The impact of this reduced sample is reflected in a lower number of significant item-level comparisons, particularly for the AOSI. Indeed, our focus on categorical predictive relationships (i.e., whether items reached statistical significance with respect to relationship to ASD outcomes) rather than strength of association may have influenced level of agreement. Third, parents received feedback concerning their child's performance at each visit, which may have affected how they rated the APSI at subsequent visits. Despite these limitations, parents of children who received a diagnosis of ASD at age 3 reported higher rates of atypical behavior on the APSI than parents of children who did not receive a diagnosis, suggesting that parental reports can complement clinician observations of early behavioral indicators of ASD, particularly for high-risk infants under 18 months of age (a need highlighted by the American Academy of Pediatrics, Johnson et al., 2007). Future examination of the utility of the APSI may benefit from comparisons of agreement on the APSI between parents and other caregivers (e.g., day care providers)

Although we are rapidly gaining knowledge about the early emergence of ASD [Zwaigenbaum, Bryson, & Garon, 2013; Jones et al., 2014], the average age of diagnosis is still around 4 years [Daniels & Mandell, 2013]. Parental reports of early symptoms of ASD suggest that, in a high-risk context, parents are able to provide meaningful information regarding diagnostic outcomes as early as six months of age [Sacrey et al., 2015; Sacrey et al., 2016]. Importantly, parents provide clinically meaningful information from their day-to day observations that can complement behavior observed in a clinical setting. It is therefore more important than

ever for health care providers to ask parents about their children's development and carefully listen to their concerns.

Acknowledgment

The authors would like to thank all the families who participated in their research as well as the research coordinators at each site. This study was supported by Canadian Institutes of Health Research and Kids Brain Health Network (formerly NeuroDevNet).

References

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. Psychological Bulletin, 101, 213–232.

Brian, J. A., Bryson, S. E., & Zwaigenbaum, L. (2015). Autism spectrum disorder in infancy: Developmental considerations in treatment targets. Current Opinions in Neurology, 28, 117–123.

Brian, J. A., Smith, I. M., Zwaigenbaum, L., Roberts, W., & Bryson, S. E. (2015). The Social ABCs caregiver-mediated intervention for toddlers with autism spectrum disorder: Feasibility, acceptability, and evidence of promise from a multisite study. Autism Research, 9, 899–912.

Bryson, S. E., Zwaigenbaum, L., Brian, J., & Roberts, W. (2006). Autism Parent Screen for Infants (questionnaire).

Bryson, S. E., Zwaigenbaum, L., McDermott, C., Rombough, V., & Brian, J. (2008). The Autism Observation Scale for Infants: Scale development and reliability data. Journal of Autism and Developmental Disorders, 38, 731–738.

Carter, A. S., Volkmar, F. R., Sparrow, S. S., Wang, J. J., Lord, C., Dawson, G., ... Schopler, E. (1998). The Vineland Adaptive Behavior Scales: Supplementary norms for individuals with autism. Journal of Autism Developmental Disorders, 28, 287–302.

Chakrabarti, S., & Fombonne, E. (2005). Pervasive developmental disorders in preschool children: Confirmation of high prevalence. American Journal of Psychiatry, 162, 1133–1141.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychological Assessment, 6, 284–290.

Daniels, A. M., & Mandell, D. S. (2013). Explaining differences in age at autism spectrum disorder diagnosis: A critical review. Autism, 18, 583–597.

Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., ... Varley, J. (2010). Randomized, controlled trial of an intervention for toddlers with autism: the early start Denver model. Pediatrics, 125, e17–e23. doi:10.1542/peds.2009-0958.

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. Psychological Bulletin, 131, 483–509.

Glascoe, F. P. (2000). Evidence-based approach to developmental and behavioural surveillance using parents' concerns. Child: Care, Health and Development, 26, 137–149.

Glascoe, F. P. (2003). Parents' evaluation of developmental status: How well do parents' concerns identify children with behavioral and emotional problems? Clinical Pediatrics, 42, 133–138.

Glascoe, F. P., Foster, M., & Wolraich, M. L. (1997). An economic analysis of developmental detection methods. Pediatrics, 830–837.

Gotham, K., Pickles, A., & Lord, C. (2009). Standardizing ADOS severity scores for a measure of severity in Autism Spectrum Disorders. Journal of Autism and Developmental Disorders, 39, 693–705.

Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The autism diagnostic observation schedule: Revised algorithms for improved diagnostic validity. Journal of Autism and Developmental Disorders, 37, 613–627.

Hayes, A. F. (2013). Introduction to mediation, moderation, and conditional process analysis: A regression-based approach. New York, NY: The Guilford Press.

Hess, C. R., & Landa, R. J. (2012). Predictive and concurrent validity of parent concern about young children at risk for autism. Journal of Autism and Developmental Disorders, 42, 575–584.

Holmbeck, G. N., Li, S. T., Schurman, J. V., Friedman, D., & Coakley, R. M. (2002). Collecting and managing multisource and multimethod data in studies of pediatric populations. Journal of Pediatric Psychology, 27, 5–18.

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

Jones, E. J., Gliga, T., Bedford, R., Charman, T., & Johnson, M.H. (2014). Developmental pathways to autism: A review of prospective studies of infants at risk. Neuroscience and Biobehavioral Reviews, 39, 1–33.

Klein, K. B., Lavigne, J. V., & Seshadri, R. (2010). Clinician-assigned and parent-report questionnaire-derived child psychiatric diagnoses: Correlates and consequences of disagreement. American Journal of Orthopsychiatry, 80, 375–385.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., ... Rutter, M. (2000). 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, 30, 205–223.

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., & Schopler, E. (1989). Autism diagnostic observation schedule: A standardized observation of communicative and social behavior. Journal of Autism and Developmental Disorders, 19, 185–212

Lord, C., Rutter, M., & LeCouteur, A. (1994). Autism diagnostic interview-revised: A revised version of time diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. Journal of Autism and Developmental Disorders, 24, 659–685.

Mullen, E. (1995). Mullen Scales of Early Learning. Circle Pines, MN: American Guidance Services.

Ozonoff, S., Young, G. S., Steinfeld, M. B., Hill, M. M., Cook, I., Hutman, T., ... Sigman, M. (2009). How early do parent concerns predict later autism diagnosis? Journal of Development and Behavioral Pediatrics, 30, 367–375.

Perry, A., Cummings, A., Dunn Geier, J., Freeman, N. L., Hughes, S., LaRose, L., ... Williams, J. (2008). Effectiveness of intensive behavioural intervention in a large, community-based program. Research in Autism Spectrum Disorders, 2, 621–642.

Robins, D. L., Fein, D., Barton, M. L., & Green, J. A. (2001). The Modified Checklist for Autism in Toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. Journal of Autism and Developmental Disorders, 31, 131–144.

Sacrey, L. R., Bryson, S., Zwaigenbaum, L., Brian, J., Smith, I. M., Roberts, W., ... Garon, N. (2016). The autism parent screen for infants: Predicting risk of ASD based on parent-reported behaviour observed at 6 to 24 months of age. Autism, 54, 470–478. DOI: https://doi.org/10.1177/1362361316675120

Sacrey, L. R., Zwaigenbaum, L. R., Bryson, S. E., Brian, J., Smith, I. M., Roberts, W., ... Armstrong, V. (2015). Can parents' concerns predict autism spectrum disorder? A prospective study of high-risk siblings from 6 to 36 months of age. Journal of the American Academy of Child & Adolescent Psychiatry, 54, 470–478.

Sallows, G. O., & Graupner, T. D. (2005). Intensive behavioural treatment for children with autism: Four-year outcome and predictors. American Journal on Mental Retardation, 110, 417–438.

Schertz, H. H., Odom, S. L., Baggett, K. M., & Sideris, J. H. (2016). Parent-reported repetitive behavior in toddlers on the autism spectrum. Journal of Autism and Developmental Disorders, 46, 3308–3316.

Schniering, C. A., Hudson, J. L., & Rapee, R. M. (2000). Issues in the assessment and diagnosis of anxiety disorders in children and adolescents. Clinical Psychological Review, 20, 453–478.

Smith, S. R. (2007). Making sense of multiple informants in child and adolescent psychopathology: A guide for clinicians. Journal of Psychoeducation Assessment, 25, 139–149.

Sparrow, S. S., Balla, D., & Cicchetti, D. (1984). Vineland Adaptive Behavior Scales (Survey Form). Circle Pines, MN: American Guidance Services.

Stadnick, N., Chlebowski, C., Baker-Ericzen, M., Dyson, M., Garland, A., & Brookman-Frazee, L. (2016). Psychiatric comorbidity in autism spectrum disorder: Correspondence between mental health clinical report and structured parent interview. Autism, 21, 841–851. doi: 10.1177/136231316654083

Verhulst, F. C., & van der Ende, J. (1991). Assessment of child psychopathology: Relationships between different methods, different informants and clinical judgment of severity. Acta Psychiatrica Scandinavica, 84, 155–159.

Volkmar, F. R., Carter, A., Sparrow, S. S., & Cicchetti, D. V. (1993). Quantifying social development in autism. Journal of the American Academy of Child and Adolescent Psychiatry, 32, 627–632.

Wetherby, A. M., Woods, J., Allen, L., Cleary, J., Dickinson, H., & Lord, C. (2004). Early indicators of Autism Spectrum Disorders in the second year of life. Journal of Autism and Developmental Disorders, 34, 473–493.

Zwaigenbaum, L., Bryson, S. E., Brian, J., Smith, I. M., Roberts, W., Szatmari, P., ...Vaillancourt, T. (2016). Stability of diagnostic assessment for autism spectrum disorder between 18 and 36 months in a high-risk cohort. Autism Research, 9, 790–800.

Zwaigenbaum, L., Bryson, S. E., Brian, J., Smith, I. M., Sacrey, L. R., Armstrong, V., et al. Development of autism symptoms in the first 18 months: Findings from a prospective cohort of high-risk infants [Manuscript in preparation].

Zwaigenbaum, L., Bryson, S., & Garon, N. (2013). Early identification of Autism Spectrum Disorders. Behavioral Brain Research, 251, 133–146.