Stability of Diagnostic Assessment for Autism Spectrum Disorder between 18 and 36 Months in a High-Risk Cohort

Lonnie Zwaigenbaum, Susan E. Bryson, Jessica Brian, Isabel M. Smith, Wendy Roberts, Peter Szatmari, Caroline Roncadin, Nancy Garon, and Tracy Vaillancourt

Children with autism spectrum disorder (ASD) are diagnosed, on average, around the age of 4 years. However, previous research has shown that the diagnosis can be made as early as 2 years, and that if the child is seen a year or more later, it is highly likely that the diagnosis will be confirmed. In this study, to examine whether diagnoses made as early as 18 months of age are also “stable,” we followed a group of younger siblings of children with ASD (who are known to be at higher risk). We also examined whether the age of ASD diagnosis within this high-risk group was related to the severity of children’s ASD symptoms or developmental delays. Participants (n = 381) were seen at three ages: 18 months, 24 months, and 3 years. ASD symptoms, general development, and adaptive functioning were assessed at each time point. Twenty-three children were diagnosed with ASD at 18 months and a total of 61 at 24 months. Of these diagnoses, 19/23 (82.6%) and 56/61 (91.8%), respectively, were confirmed independently at 3 years. However, 45 children were diagnosed with ASD at 3 years who had not been identified at earlier visits. Children diagnosed at 18 months, in comparison to those diagnosed at 24 months, had less advanced language and adaptive skills at 18 months. Children not diagnosed with ASD until 3 years, compared with those diagnosed earlier, had more advanced language and adaptive skills, and milder ASD symptoms. Autism Res 2015, 00: 000–000. © 2015 International Society for Autism Research, Wiley Periodicals, Inc.

Keywords: autism; early diagnosis; early identification; infants; longitudinal study

Introduction

Recent U.S. data suggest that the average age of autism spectrum disorder (ASD) diagnosis is about 4 years [CDC, 2014] but there is considerable interest in moving the diagnosis to younger ages. Parents of children with ASD often identify concerns by the age of 12–18 months [De Giacomo & Fombonne, 1998; Rogers & DiLalla, 1990; Wimpory et al., 2000] but frequently experience false reassurance from health care providers [Ryan & Salisbury, 2012] and/or frustrating waiting periods following referral for specialized assessments [Wiggins, Baio, & Rice, 2006], both of which contribute to delays in diagnosis. Accordingly, professional practice parameters [Baird, Douglas, & Murphy, 2011; Johnson, Meyers, & Council on Children with Disabilities, 2007; Volkmar et al., 2014] include recommendations for earlier detection and diagnosis by increasing knowledge of early signs of ASD, implementing formal surveillance and/or screening, and by expediting referral pathways. For example, the American Academy of Pediatrics recommends that all children be screened for ASD at 18 and 24 months [Johnson et al., 2007]. There is growing evidence that ASD screening leads to earlier detection of ASD compared with general surveillance by community pediatricians [Pandey et al., 2008; Robins, 2008; Robins et al., 2014]. Furthermore, with increasing awareness of the early signs, some children are being referred and diagnosed with ASD at 18 months or younger [Chawarska et al., 2007]. These efforts are creating opportunities to initiate ASD-specific interventions earlier than ever before (for a review, see Beaudoin, Sebire & Couture, 2014). However, important questions remain about the stability of diagnoses made under the age of 24 months [Camarata, 2014].

In at least 19 published studies, children received an initial diagnostic assessment for possible ASD prior to age 3 and then were reassessed at least 1 year later [Charman et al., 2005; Chawarska et al., 2007, 2009; Corsello et al., 2013; Cox et al., 1999; Eaves et al., 2004; Gillberg et al, 1990; Guthrie et al., 2013; Itzchak & Zachor, 2009; Kleinman et al, 2008; Lord, 1995; Lord...
but also negative predictive value of an initial classification of non-ASD. This is particularly critical among children with recognized ASD symptoms that are sub-threshold for diagnosis at initial assessment.

Prospective studies of high-risk (HR) infant siblings have generated new insights about early signs and developmental trajectories associated with ASD [Zwaigenbaum et al., 2007, 2013]. These cohorts are potentially useful in the study of stability of early diagnoses, as ascertainment does not depend on clinical referral, which may be biased by symptom severity or other extraneous factors. Moreover, among siblings there is a “built-in” contrast group of toddlers who are also at increased risk of ASD (and who may exhibit early behavioral signs) but do not ultimately receive a clinical diagnosis [Georgiades et al., 2013].

Our main aims in this article are to evaluate stability of early assessment for ASD in our HR longitudinal cohort, and to examine factors related to diagnostic misclassification relative to independent assessment at 3 years. Analyses address the following questions: (1) what are the sensitivity and specificity of clinical ASD diagnoses established at 18 and 24 months relative to independent gold-standard diagnoses established at 3 years; and (2) how do clinical features (language, cognitive and adaptive skills, and ASD symptoms) vary by timing of initial clinical diagnosis?

**Methods**

**Participants**

The sample was comprised of 381 infant siblings of children with ASD (hereafter, HR infants), recruited at age 6–12 months from families attending one of five Canadian autism diagnostic and treatment centers: the Glenrose Rehabilitation Hospital in Edmonton, McMaster Children’s Hospital in Hamilton, The Hospital for Sick Children and Holland Bloorview Kids Rehabilitation Hospital in Toronto and the IWK Health Centre in Halifax, and by additional efforts of community clinicians in the surrounding regions. This study was approved by the research ethics boards of the five institutions, and all families gave written informed consent. All HR infants followed to age three were included.

Diagnosis of ASD in the older sibling (i.e., proband) was confirmed by a clinical assessment using DSM-IV-TR criteria. Neither the infant siblings nor the probands had identifiable neurological or genetic conditions, or severe sensory or motor impairments. All infants were born at 36–42 weeks gestation and had a birth weight greater than 2500 g.

**Study Design**

In this prospective cohort study, participants’ diagnostic status as well as developmental and ASD symptom...
levels were assessed from age 18 months to 3 years. At 18 months, the children’s language skills were assessed by parental report using the McArthur Child Development Inventory—Words and Gestures. At 24 months and 3 years, the participants’ language and nonverbal skills were assessed using the Mullen Scales of Early Learning (MSEL). Adaptive skills were assessed at all ages using the Vineland adaptive behavior scales. The Autism Diagnostic Observation Schedule (ADOS) was used to evaluate ASD-related symptoms at each time point; the Autism Diagnostic Interview-Revised (ADI-R) was also used at the 3-years assessment.

**Measures**

**The MacArthur Communicative Development Inventories—Words and Gestures.** The MacArthur Communicative Development Inventories - Words and Gestures (CDI-WG) is a highly reliable and well-validated parent-report measure of communicative development [Fenson et al., 1993]. The infant scale is standardized from 8 to 16 months, but has also been used to map the developmental trajectories of older children with ASD [Charman, Drew, Baird, & Baird, 2003]. The CDI-WG subscales provide quantitative indices of expressive language (word production), receptive language (word comprehension, phrase comprehension) and gestural communication (early gestures; i.e., first communicative gestures such as pointing and late gestures; i.e., actions with objects and imitation of adult actions).

**The Mullen Scales of Early Learning.** The Mullen Scales of Early Learning (MSEL) consists of five scales, four of which (visual reception, receptive language, expressive language, and fine motor) assess nonverbal cognitive and language ability, while the fifth scale measures gross motor development (from 0 to 29 months only) [Mullen, 1995]. 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].

**Vineland Adaptive Behavior Scales.** The Vineland Adaptive Behavior Scales (VABS) is a semistructured parent interview designed to assess adaptive behavior across four subdomains—communication, daily living, socialization, and motor skills (limited to children younger than 30 months), outlined by typical developmental milestones that are anchored to specific ages [Sparrow, Cicchetti, & Balla, 1984]. The scale has excellent reliability and concurrent validity, and is sensitive to impairments experienced by children with ASD [Carter et al., 1998].

**Autism Diagnostic Observation Schedule.** The Autism Diagnostic Observation Schedule (ADOS) uses standardized activities and “presses” to elicit communication, social interaction, imaginative use of play materials, and repetitive behaviors [Lord, Rutter, DiLavore, & Risi, 2002]. Inter-rater reliability of the ADOS is excellent [Lord et al., 2002]. The scoring algorithm 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 nonechoed 3-word phrases, Module 3 = child with fluent language; and Module 4 = adolescent or adult with fluent language). Modules 1–3 were used in this study; only Module 1 was administered at the 18-month assessment. To allow comparability across modules used at 24 months and 3 years, the ADOS severity metric was calculated [Gotham, Pickles, & Lord, 2009].

**Autism Diagnostic Interview-Revised.** The Autism Diagnostic Interview - Revised (ADI-R) is an investigator directed interview that generates quantitative estimates of symptoms related to social development, verbal and nonverbal communication skills, and the presence of repetitive, stereotyped interests and behaviors required to make an ICD-10 or DSM-IV-TR diagnosis of autism [Lord et al., 1994]. The questions are designed to distinguish qualitative impairments from developmental delays. The ADI-R discriminates well between autism and other forms of developmental disability, and inter-rater reliability is excellent [Lord et al., 1994].

**Diagnostic Procedures**

All ASD diagnoses were established by a clinical psychologist, developmental pediatrician or child psychiatrist with at least 10 years assessment experience leading multidisciplinary assessment teams. “Clinical best estimate (CBE)” diagnostic assessments at 18 and 24 months were based on developmental history (including a structured parent concerns interview), the ADOS and other developmental assessments described above. If the child met clinical best-estimate criteria for ASD prior to the 36-month assessment, the diagnosis was communicated to the parents, and the child was referred to community intervention services. At 36–42 months of age (hereafter, “3 years”), and blind to previous assessment results, an independent diagnostic evaluation of each participant was conducted using the ADI-R, ADOS, and DSM-IV-TR. Families were asked not to reveal group status (sibling or control) or previous diagnoses until the clinician had completed their
child’s evaluation. A clinical diagnosis was assigned using DSM-IV-TR criteria, based on the best judgment of the expert clinician, taking into account all available information from the ADI-R, ADOS, and concurrent developmental assessment using the MSEL and VABS.

Analytic Approach

First, we described the HR sample, comparing those with and without ASD (as determined at age 3) with respect to sex, family socioeconomic status (SES; indexed by the Hollingshead Four Factor Index), and age of assessment at each visit. Second, we estimated the sensitivity and specificity, as well as the positive and negative predictive value, of ASD diagnoses established at 18 and 24 months, relative to independent clinical assessments completed at 3 years. We also examined the concordance between the ADOS and CBE at each time point. Finally, we compared children with confirmed ASD diagnoses at age 3 stratified on the basis of timing of initial diagnosis; that is, at 18 months, 24 months, and 3 years. These groups were compared on developmental measures (MacArthur communicative development inventories—words and gestures [MCDI-WG], MSEL subscales), adaptive skills (VABS subscales) and ASD symptoms (ADOS, ADI-R) using 1-way ANOVA with post hoc pairwise comparison by Tukey HSD. Chi-squared tests were used to compare the groups with respect to sex distribution and language level (as indexed by the ADI-R, item 30). Comparisons were conducted based on 3-year assessment data to assess how heterogeneity at this time point related to initial timing of ASD diagnoses, and at ages 18 and 24 months, to assess how early clinical features might have influenced initial clinical judgment.

Results

Participant Characteristics

Participants included 381 HR siblings (215 boys and 166 girls), of whom 103 (27.0%) were diagnosed with ASD at age 3. Data were unavailable at 18 months in two participants, and at 24 months in six participants. Consistent with previous reports of this cohort [Zwaigenbaum et al., 2012], the sex ratio was higher in HR siblings diagnosed with ASD (2.2:1) than in those who were not (1.1:1) ($\chi^2 = 8.97; P = 0.003$). The two HR groups did not differ with respect to parental SES or mean age at the three assessment points (see Table 1).

Agreement between Diagnostic Classification at 18 and 24 Months, Compared with 3 Years

Table 2 summarizes ASD diagnostic status at 18 months, 24 months, and 3 years in this sample of 381 HR siblings. At 18 months, of the children diagnosed with ASD, 19 of 23 had confirmed diagnoses at age 3 (i.e., positive predictive value = 0.83). However, only 19 of 103 children diagnosed with ASD at age 3 were classified as such at 18 months, a sensitivity of only 0.18. Four siblings shifted from ASD to non-ASD between 18 months and 3 years (3 between 18 and 24 months, and 1 between 24 months and 3 years), which was 1.4% of the total number of HR siblings not diagnosed with ASD at age 3 (i.e., specificity = 0.99). As well, 90 of 356 siblings classified as non-ASD at 18 months were subsequently diagnosed at 3 years (i.e., negative predictive value = 0.75). Similar calculations from Table 2 indicate that an ASD diagnosis at 24 months was predictive of 3-year ASD diagnoses, with 0.54 sensitivity and 0.98 specificity, and a positive predictive value of 0.92 and negative predictive value of 0.85. Five of 61 children diagnosed with ASD at the 24-month assessment did not have confirmed diagnoses at age 3. This included 1 child also classified as ASD at age 18 months, so in total, 8 of 66 children (12.1%) who received a clinical diagnosis at 18 and/or 24 months did not meet ASD criteria at age 3. All of these children at age 3 had evidence of elevated ASD symptoms (based on ADOS social affect or restricted repetitive behaviors subscale.

Table 1. Participant Characteristics

<table>
<thead>
<tr>
<th></th>
<th>ASD</th>
<th>Non-ASD</th>
<th>Chi-squared</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex (M:F)</strong></td>
<td>71:32 (2.2:1)</td>
<td>144:134 (1.1:1)</td>
<td>8.97</td>
<td>0.003</td>
</tr>
<tr>
<td><strong>Parental SES</strong></td>
<td>46.9 (13.1)</td>
<td>48.0 (12.8)</td>
<td>0.48</td>
<td>0.49</td>
</tr>
<tr>
<td><strong>Age at 18-month visit</strong></td>
<td>18.4 (0.6)</td>
<td>18.4 (0.5)</td>
<td>1.1</td>
<td>0.30</td>
</tr>
<tr>
<td><strong>Age at 24-month visit</strong></td>
<td>24.7 (1.2)</td>
<td>24.5 (0.8)</td>
<td>2.3</td>
<td>0.10</td>
</tr>
<tr>
<td><strong>Age at 3-year visit</strong></td>
<td>38.7 (3.3)</td>
<td>39.0 (3.2)</td>
<td>0.54</td>
<td>0.46</td>
</tr>
</tbody>
</table>

- Hollingshead index
- Age in months
severity score of 4 or higher; 7 of 8) and/or developmental delay (based on scoring at least 1 SD below the mean on at least one MSEL subscale; 3 of 8).

Concordance between ADOS and Clinical Best Estimate at Each Time Point

At 18 months, 153 of 379 (40.4%) participants scored above the ASD cut-point on the ADOS algorithm; 21 of 153 (13.7%) received a CBE diagnosis of ASD at that time, as did 2 of 226 (0.9%) of those scoring below the ADOS algorithm cut-point. Overall agreement was kappa = 0.15 (95% CI = 0.09–0.21). At 24 months, 156 of 375 (41.6%) participants scored above the ASD cut-point on the ADOS algorithm; 56 of 156 (35.9%) received a CBE diagnosis (CBE) of ASD, as did 2 of 219 (0.9%) of those scoring below the ADOS algorithm cut-point. Overall agreement was kappa = 0.39 (95% CI = 0.31–0.47). At 3 years, 158 of 381 (41.5%) participants scored above the ASD cut-point on the ADOS algorithm; 94 of 158 (59.1%) received a CBE diagnosis (CBE) of ASD, as did 9 of 223 (4.0%) of those scoring below the ADOS algorithm cut-point. Overall agreement was kappa = 0.58 (95% CI = 0.50–0.66).

As an exploratory analysis, we examined rates of 3-year ASD diagnoses among children who did not receive CBE diagnoses at 18 and 24 months in relation to ADOS classification at these earlier assessments. Among children not diagnosed with ASD at 18 months, 54 of 132 (41.9%) “ADOS-positive” (above ASD cut-point) children, and 26 of 224 (11.6%) of “ADOS-negative” children, were diagnosed with ASD at 3 years (chi-squared = 40.9; P < 0.001). Among children not diagnosed with ASD at 24 months, 27 of 100 (27.0%) “ADOS-positive” children, and 20 of 217 (9.2%) of “ADOS-negative” children were diagnosed with ASD at 3 years (chi-squared = 17.1; P < 0.001). Thus, among children who were not clinically diagnosed at the earlier time-points, there appeared to be a gradient of risk of 3-year ASD diagnoses related to ADOS status. Notably, among children with a CBE diagnosis of ASD at 18 months, only one of 23 were concurrently “ADOS-negative”; that child had a CBE diagnosis of ASD at 3 years. Similarly, among children with a CBE diagnosis of ASD at 24 months, only 2 of 61 were concurrently “ADOS-negative,” both of whom had a CBE diagnosis of ASD at 3 years. Thus, it was rare for children to be clinically diagnosed with ASD at 18 or 24 months if they were subthreshold on the ADOS, but in those instances, the diagnosis was stable.

Clinical Variation within ASD Group by Timing of Diagnosis

Among the 103 children with confirmed ASD diagnoses at age three, 19 were diagnosed at 18 months. An additional 37 children were diagnosed at 24 months, and an additional 47 children at the 3-year assessment. Clinical variation among children with ASD was assessed relative to initial age of diagnosis, with separate analyses conducted for 18-month, 24-month, and 3-year assessment data (see Table 3). No differences were found among these three groups with respect to sex distribution, site, family SES, or maternal education (sex and family SES in Table 3; other data available on request).

Clinical features at 18 months. Group differences were detected on all subscales of the MCDI-WG and VABS at 18 months. Post hoc testing indicated that children with ASD initially diagnosed at 18 months had less advanced language skills compared with those initially diagnosed at 24 months or at 3 years, as reported by parents on the MCDI (phrases understood, vocabulary production and early gestures subscales). Post hoc testing also indicated a significant gradient across the 3 age groupings with respect to communication and socialization domains of adaptive functioning on the 18-month VABS, with the most severe delays in the children diagnosed at 18 months. Daily living skills were also less advanced in the 18-month diagnosed group compared with the group diagnosed at 3 years. Group differences were also seen on the social-affect and restricted repetitive behaviors domains of the ADOS (Module 1) at 18 months, with post hoc tests indicating that children diagnosed at 18 or 24 months differed from later diagnosed children, but not from each other (algorithm rather than “severity” scores were used as all were assessed using Module 1).

Clinical features at 24 months. There were group differences at 24 months on all MSEL and VABS subscales, as well ASD symptoms (indexed by the ADOS severity metric). In contrast to group comparisons at

---

Table 2. Diagnostic Classification of HR Siblings at 18-month, 24-month, and 3-year Assessments

<table>
<thead>
<tr>
<th></th>
<th>18 months</th>
<th>24 months</th>
<th>3 years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td>= ASD</td>
<td>17</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>No</td>
<td>39</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>= ASD</td>
<td>2</td>
<td>45</td>
<td>12.1</td>
</tr>
<tr>
<td>Yes</td>
<td>1</td>
<td>1</td>
<td>0.3</td>
</tr>
<tr>
<td>No</td>
<td>3</td>
<td>4</td>
<td>1.1</td>
</tr>
<tr>
<td>= ASD</td>
<td>3</td>
<td>3</td>
<td>0.8</td>
</tr>
<tr>
<td>No</td>
<td>9</td>
<td>262</td>
<td>70.2</td>
</tr>
</tbody>
</table>

Total 373

Note: excludes 6 HR siblings classified as non-ASD at 18 months and ASD at 3 years but in whom 24-month data were not available, as well as 2 HR siblings for whom 18-month data were unavailable.

---
Table 3. Comparison of Children with ASD by Timing of Diagnosis

<table>
<thead>
<tr>
<th></th>
<th>a 18-month visit (n = 19)</th>
<th>b 24-month visit (n = 37)</th>
<th>c 3-year visit (n = 47)</th>
<th>chi-squared</th>
<th>P-value</th>
<th>Post hoc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex (male:female)</td>
<td>14:5</td>
<td>25:12</td>
<td>32:15</td>
<td>0.25</td>
<td>.88</td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td>F-test</td>
<td>P-value</td>
<td></td>
</tr>
<tr>
<td>Family SES (Hollingshead Index)</td>
<td>46.9</td>
<td>13.0</td>
<td>46.0</td>
<td>12.5</td>
<td>0.14</td>
<td>0.87</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td>F-test</td>
<td>P-value</td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td>F-test</td>
<td>P-value</td>
<td></td>
</tr>
<tr>
<td>18-month assessments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MacArthur CDI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phrases understood</td>
<td>9.4 (5.8)</td>
<td>16.0 (6.7)</td>
<td>18.8 (5.5)</td>
<td>15.5</td>
<td>&lt; 0.001</td>
<td>a &lt; b,c</td>
</tr>
<tr>
<td>Vocabulary production</td>
<td>8.1 (17.4)</td>
<td>32.6 (41.8)</td>
<td>36.1 (33.5)</td>
<td>14.9</td>
<td>&lt; 0.001</td>
<td>a &lt; b,c</td>
</tr>
<tr>
<td>Early gestures</td>
<td>8.1 (3.1)</td>
<td>11.0 (3.2)</td>
<td>12.6 (2.4)</td>
<td>6.48</td>
<td>0.003</td>
<td>a &lt; b &lt; c</td>
</tr>
<tr>
<td>Late gestures</td>
<td>14.0 (8.4)</td>
<td>18.2 (8.4)</td>
<td>22.4 (6.1)</td>
<td>2.68</td>
<td>0.075</td>
<td>a &lt; c</td>
</tr>
<tr>
<td>VABS subscales (Std scores)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communication</td>
<td>75.2 (6.0)</td>
<td>84.4 (10.9)</td>
<td>89.7 (7.2)</td>
<td>15.5</td>
<td>&lt; 0.001</td>
<td>a &lt; b &lt; c</td>
</tr>
<tr>
<td>Social</td>
<td>82.9 (9.3)</td>
<td>88.6 (7.5)</td>
<td>94.5 (5.7)</td>
<td>14.9</td>
<td>&lt; 0.001</td>
<td>a &lt; b &lt; c</td>
</tr>
<tr>
<td>Daily living skills</td>
<td>78.0 (5.8)</td>
<td>83.8 (6.6)</td>
<td>87.7 (9.8)</td>
<td>6.48</td>
<td>.003</td>
<td>a &lt; c</td>
</tr>
<tr>
<td>Motor</td>
<td>92.2 (7.2)</td>
<td>94.8 (11.3)</td>
<td>98.5 (8.3)</td>
<td>2.68</td>
<td>0.075</td>
<td></td>
</tr>
<tr>
<td>ADOS-Module 1 subscales</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social affect</td>
<td>12.6 (4.6)</td>
<td>10.8 (3.6)</td>
<td>6.4 (4.6)</td>
<td>18.5</td>
<td>&lt; 0.001</td>
<td>a &lt; b &lt; c</td>
</tr>
<tr>
<td>Restricted repetitive behavior</td>
<td>3.6 (1.9)</td>
<td>3.8 (2.0)</td>
<td>2.5 (1.7)</td>
<td>2.68</td>
<td>0.005</td>
<td>b &lt; c</td>
</tr>
<tr>
<td>24-month assessments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSEL subscales (T scores)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expressive language</td>
<td>34.8 (13.1)</td>
<td>40.8 (13.0)</td>
<td>48.4 (9.6)</td>
<td>9.06</td>
<td>&lt; 0.001</td>
<td>a &lt; c</td>
</tr>
<tr>
<td>Receptive language</td>
<td>32.7 (15.8)</td>
<td>31.2 (13.5)</td>
<td>48.9 (12.4)</td>
<td>18.0</td>
<td>&lt; 0.001</td>
<td>a &lt; b &lt; c</td>
</tr>
<tr>
<td>Visual receptive</td>
<td>42.9 (15.1)</td>
<td>40.9 (14.0)</td>
<td>50.4 (9.7)</td>
<td>5.85</td>
<td>0.04</td>
<td>b &lt; c</td>
</tr>
<tr>
<td>Fine motor</td>
<td>39.9 (14.6)</td>
<td>38.8 (12.4)</td>
<td>46.9 (9.8)</td>
<td>4.76</td>
<td>0.011</td>
<td>b &lt; c</td>
</tr>
<tr>
<td>VABS subscales (std scores)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communication</td>
<td>74.4 (8.9)</td>
<td>82.5 (13.8)</td>
<td>93.5 (8.7)</td>
<td>16.9</td>
<td>&lt; 0.001</td>
<td>a &lt; b &lt; c</td>
</tr>
<tr>
<td>Social</td>
<td>78.3 (9.8)</td>
<td>81.5 (8.5)</td>
<td>90.3 (5.9)</td>
<td>15.2</td>
<td>&lt; 0.001</td>
<td>a &lt; b &lt; c</td>
</tr>
<tr>
<td>Daily living skills</td>
<td>76.6 (8.4)</td>
<td>79.4 (10.0)</td>
<td>88.0 (7.9)</td>
<td>15.1</td>
<td>&lt; 0.001</td>
<td>a &lt; b &lt; c</td>
</tr>
<tr>
<td>ADOS severity scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social affect</td>
<td>5.8 (1.8)</td>
<td>6.5 (1.7)</td>
<td>3.9 (2.2)</td>
<td>16.7</td>
<td>&lt; 0.001</td>
<td>a &lt; b &lt; c</td>
</tr>
<tr>
<td>Restricted repetitive behavior</td>
<td>7.7 (2.4)</td>
<td>8.3 (1.4)</td>
<td>6.3 (2.3)</td>
<td>9.3</td>
<td>&lt; 0.001</td>
<td>b &lt; c</td>
</tr>
<tr>
<td>Overall</td>
<td>6.0 (1.9)</td>
<td>7.3 (1.7)</td>
<td>4.1 (2.3)</td>
<td>22.5</td>
<td>&lt; 0.001</td>
<td>a &lt; b &lt; c</td>
</tr>
<tr>
<td>3-year assessments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSEL subscales (T scores)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expressive language</td>
<td>36.7 (14.9)</td>
<td>39.4 (13.6)</td>
<td>46.9 (9.6)</td>
<td>6.30</td>
<td>0.003</td>
<td>a,b &lt; c</td>
</tr>
<tr>
<td>Receptive language</td>
<td>36.2 (14.7)</td>
<td>39.7 (15.2)</td>
<td>45.6 (11.3)</td>
<td>3.73</td>
<td>0.028</td>
<td>a &lt; c</td>
</tr>
<tr>
<td>Visual receptive</td>
<td>40.6 (22.3)</td>
<td>44.5 (19.2)</td>
<td>52.0 (16.3)</td>
<td>3.16</td>
<td>0.047</td>
<td>a &lt; c</td>
</tr>
<tr>
<td>Fine motor</td>
<td>35.4 (18.7)</td>
<td>38.2 (15.6)</td>
<td>42.8 (11.2)</td>
<td>2.07</td>
<td>0.132</td>
<td></td>
</tr>
<tr>
<td>VABS subscales (std scores)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communication</td>
<td>76.9 (20.5)</td>
<td>87.4 (15.4)</td>
<td>92.9 (11.7)</td>
<td>4.82</td>
<td>0.011</td>
<td>a &lt; c</td>
</tr>
<tr>
<td>Social</td>
<td>70.4 (8.9)</td>
<td>75.9 (10.2)</td>
<td>82.6 (9.5)</td>
<td>7.56</td>
<td>0.001</td>
<td>a &lt; b &lt; c</td>
</tr>
<tr>
<td>Daily living skills</td>
<td>67.1 (10.6)</td>
<td>77.9 (12.1)</td>
<td>81.7 (10.3)</td>
<td>6.91</td>
<td>0.002</td>
<td>a &lt; c</td>
</tr>
<tr>
<td>ADOS severity scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social affect</td>
<td>6.4 (2.7)</td>
<td>6.8 (2.1)</td>
<td>5.9 (1.7)</td>
<td>2.22</td>
<td>0.114</td>
<td></td>
</tr>
<tr>
<td>Restricted repetitive behavior</td>
<td>8.0 (1.5)</td>
<td>8.0 (1.5)</td>
<td>7.3 (2.2)</td>
<td>2.02</td>
<td>0.139</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>6.6 (2.9)</td>
<td>7.0 (2.1)</td>
<td>6.0 (1.5)</td>
<td>2.30</td>
<td>0.106</td>
<td></td>
</tr>
<tr>
<td>ADI-R domain scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communication</td>
<td>11.2 (3.3)</td>
<td>10.6 (4.1)</td>
<td>8.0 (3.8)</td>
<td>6.70</td>
<td>0.002</td>
<td>a,b &lt; c</td>
</tr>
<tr>
<td>Social</td>
<td>14.9 (5.2)</td>
<td>12.6 (6.0)</td>
<td>8.3 (5.0)</td>
<td>12.6</td>
<td>&lt; 0.001</td>
<td>a,b &lt; c</td>
</tr>
<tr>
<td>Repetitive behavior</td>
<td>6.0 (2.7)</td>
<td>4.5 (2.4)</td>
<td>2.7 (2.3)</td>
<td>14.8</td>
<td>&lt; 0.001</td>
<td>a,b &lt; c</td>
</tr>
<tr>
<td>Language level (ADI-R)</td>
<td>13:3:3</td>
<td>29:6:2</td>
<td>45:1:1</td>
<td>10.92</td>
<td>0.027</td>
<td>a,b &lt; c</td>
</tr>
</tbody>
</table>

*ADI-R language level categories: 0: functional use of 3-word phrases; 1: some speech (>5 single words); 2: nonverbal (<5 words)
age 18 months, post hoc testing indicated differences between the 18-month diagnosed group and/or the 24-month diagnosed group compared with the 3-year diagnosed group, but no differences between the 18-month and 24-month diagnosed groups on early language, visual reception or fine motor development, as indexed by the MSEL subscales. Similarly, post hoc testing revealed that the 18-month and 24-month diagnosed groups differed from the 3-year diagnosed group with respect to all domains of adaptive functioning indexed by the VABS, as well as ASD symptom severity. Thus, at age 24 months, children diagnosed with ASD at 18 or at 24 months were less advanced in general development and adaptive functioning, and had more severe symptoms than children with ASD with later diagnoses. The 18-month and 24-month groups did not differ from each other on any 24-month measure.

**Clinical features at 3 years.** There were group differences on MSEL language and visual reception (but not fine motor) subscales, all VABS subscales, and on indices of ASD symptoms from both the ADOS and ADI-R. Post hoc testing revealed a pattern of differences similar to that at 24 months, with children diagnosed with ASD at 18 months and those diagnosed at 24 months differing from children diagnosed at 3 years (but not from each other) with respect to expressive language (indexed by the MSEL, as well as language level defined on the ADI-R—phrase speech, single words, less than 5 single words), social adaptive functioning and the three ASD symptom domains of the ADI-R (social, communication, restricted interests, and repetitive behaviors). Children diagnosed with ASD at 18 months also differed from children diagnosed at 3 years with respect to receptive language (on the MSEL) and the communication and daily living skills domains of the VABS. In general, the 24-month group means fell between the 18-month and 3-year group means, even if no significant differences were identified. Notably, no group differences in ASD symptom severity on the ADOS were detected.

**Discussion**

The findings of this study of HR infants may have important implications for the early assessment and diagnosis of ASD in other samples and contexts. First, ASD diagnoses established at 18 months and at 24 months were associated with high positive predictive value (82.6% and 91.8%, respectively), although almost 50% of children diagnosed with ASD at age 3 were not identified as such at previous assessments. Second, children not diagnosed with ASD until age 3 had more advanced language and general adaptive skills and less severe ASD symptoms (by parent-report on the ADI-R; ADOS scores did not differ) relative to children who were diagnosed at an earlier age. Children diagnosed at 18 months differed from those diagnosed at subsequent ages, including those diagnosed at 24 months, on the basis of more severe delays in language and adaptive functioning, particularly in social and communication domains. Children diagnosed at 18 months had similar levels of ASD symptoms to those not diagnosed until 24 months at all three time points. Interestingly, although the agreement between ADOS classification and CBE was quite low at 18 and 24 months, in part because of the majority of children scoring above the ASD cut-point on the ADOS did not receive a CBE diagnosis at that time, “ADOS-positive” children among the nondiagnosed group were at higher (albeit still <50%) risk of ASD diagnosis by CBE at age 3 than those with lower ADOS scores.

This study is unique in two respects. First, we examined diagnostic stability over three time points, starting at 18 months. To our knowledge, only one previous study examined diagnostic stability from such an early age [Guthrie et al., 2013]. That study also used a smaller and less diverse sample of toddlers who screened positive for communication delays known to be predictive of ASD. Other studies examining stability of early ASD diagnoses have included children 18 months of age and younger, but have not reported on this group separately. Indeed, to our knowledge, this study is the first to identify clinical features (i.e., language and adaptive skill delays) that distinguish children diagnosed with ASD at 18 months vs. 24 months or later. Second, we examined diagnostic stability in a HR cohort (siblings of children with ASD), rather than a clinically referred sample. The two sampling strategies (that is, HR vs. clinical referral) likely provide complementary information. The potential advantage of samples recruited on the basis of familial risk is that there may be better representation of children with milder impairments than those referred due to clinical concerns at 18–24 months. For example, children with milder symptoms and less impaired language (e.g., those diagnosed with Asperger disorder during the DSM-IV era) are generally referred much later than other children on the autism spectrum [Howlin & Asgharian, 1999]. The American academy of pediatrics [Johnson et al., 2007] and other professional bodies have specifically targeted siblings of children with ASD for screening and diagnostic evaluation starting at 18 months, so age-specific validation of diagnostic procedures in this HR group is a priority.

The positive predictive values of 18-month and 24-month clinical diagnoses established in our HR cohort, relative to 3-year assessments, are comparable to the high stability rates that have been reported in follow-up studies of clinically referred 2-year-olds [Charman et al., 2005; Gillberg et al., 1990; Kleinman et al., 2008, Lord et al., 2006; Stone et al., 1999; Turner & Stone, 2007]. However, young children with ASD who are missed by early detection and screening efforts, and thus, not referred for diagnostic assessment, are excluded from
estimates of early diagnostic stability. For example, Guthrie et al. (2013) reported 100% positive predictive value for early ASD diagnoses (mean age = 19.1 months) relative to follow-up assessments at age 3 years in a group of toddlers who had screened positive on the ITC and follow-up SORF. However, this group had a verbal developmental quotient (DQ; i.e., age equivalent on the MSEL divided by chronological age) of 61.7 at the first time point. Thus, children without early language delays (which in our study, were generally not classified with ASD until the 3-year assessment) were not represented. Guthrie et al. (2013) also reported that only 3 of 26 (11%) toddlers who initially did not meet criteria for ASD at their initial assessment were ultimately diagnosed, all of whom were among the 14 toddlers whose initial presentation was ambiguous (and classification “deferred”). However, this later-diagnosed group also had significant language delays, with a mean verbal DQ of 75 at the initial assessment. As well, a recent population-based screening study in Japan reported that the modified checklist for autism in toddlers detected 47.4% of 18-month-olds who were subsequently diagnosed with ASD, either at follow-up to a positive screen at age 2, or at a routine 3-year health check-up (all diagnoses were confirmed at ages 3–5 years). Children with ASD who were missed by the 18-month screen (i.e., false negatives) were less likely than screen-positive children with ASD (i.e., true positives) to be developmentally delayed when assessed at age 3–5 years [Kamio et al., 2014]. Thus, children with ASD without developmental delays may be systematically missed at 18–24 months, by current screening strategies, and even if they are referred, may be missed even at the time of diagnostic assessment.

In this study, children diagnosed with ASD at 18 months were primarily distinguished from children diagnosed with ASD at 24 months on the basis of severity of language and adaptive skill delays (i.e., impairments in day-to-day functioning), rather than by level of observed ASD symptoms. Children diagnosed initially at 24 months, on average, had ADOS scores consistent with ASD when assessed at 18 months but had milder developmental delays than their earlier-diagnosed peers, which may have influenced clinical decision-making. Given current referral patterns, even “expert clinicians” who have led diagnostic teams for many years may have relatively limited experience in assessing 18-month olds with suspected ASD. If community-wide screening at that early age is implemented, diagnostic teams may require additional training to diagnose children with ASD accurately across the developmental continuum. Children initially diagnosed at 3 years were less developmentally delayed and had less severe ASD symptoms than earlier diagnosed children at each assessment from 18 months to 3 years. Indeed, they generally had mean MSEL and VABS scores at age 2 and 3 years within the average range (with the exception of VABS—daily living skills at age 3). However, these children, as a group, had elevated ASD symptoms as measured by the ADOS; by age 2, mean severity scores were at or near the cut-point associated with the boundary of the ASD spectrum [Gotham, Pickles, & Lord, 2009; Hus, Gotham, & Lord, 2014]. Consistent with the “deferred assessments” reported by Guthrie et al. (2013), there may be a significant subgroup of children with ASD for whom differential diagnosis remains very difficult at 18 and 24 months, and for whom careful follow-up and reassessment is essential. Only 19 of 103 (18%) of all children diagnosed with ASD at age 3 were classified as such at 18 months, despite undergoing formal diagnostic evaluation, including assessment using the ADOS. This has important implications for interpretation of diagnostic assessment following positive ASD screening at 18 months: children who are not regarded as having ASD are by no means “out of the woods,” and should be monitored for potential change in diagnostic status as they get older. Indeed, differential diagnosis may be difficult in some children with suspected ASD even at 24 months [Zwaigenbaum et al., 2009]; in this study, only 56 of 103 (55%) of all children diagnosed with ASD at age 3 received a clinical diagnosis by age 2. Importantly, children with early features of ASD may benefit from intervention aimed at areas of concern before final diagnostic status is confirmed. Kasari et al. (2014), Rogers et al. (2014) and Wetherby et al. (2014) have reported the benefits of parent-mediated interventions for infants and toddlers identified as “at risk” of ASD based on the presence of behavioral indicators; other models for that purpose are being evaluated [Siller et al., 2014]. Access to effective interventions targeting functional deficits identified through ASD-specific surveillance and screening could help ensure that such initiatives benefit children who are identified, despite challenges in early diagnosis highlighted in the current study.

Notably, there was no association between timing of diagnosis (nor absolute rates of diagnosis) with family socioeconomic class (see Table 3) within this HR sample. These findings contrast with analyses of population-based data from the autism and developmental disabilities monitoring network showing lower rates and later mean age of ASD diagnosis of children from families of lower SES [Durkin et al., 2010]. Indeed, the lack of association between diagnostic rates and timing and SES in this study suggests that access to intensified surveillance not dependent on clinical referral could mitigate against under-identification of ASD in lower SES families, at least among HR infants, although this remains to be demonstrated in a community context.
This study has a number of strengths including longitudinal design and independent assessment of diagnosis at 3 years but also some important limitations. First, children with ASD ascertained from HR infant sibling cohorts are likely not representative of children identified from the community. For example, intensified surveillance and comprehensive diagnostic assessment regardless of level of clinical concern as implemented in our study may lead to identification of ASD cases with milder symptoms and less intellectual and language impairment, and thus, a shift in the overall phenotypic distribution may be observed among children receiving a diagnosis. We would suggest that the HR design provides a unique opportunity to assess age of diagnosis relative to early developmental profiles and symptom emergence, independent of clinical referral decisions. However, elevated rates of developmental impairments “beyond ASD” [Messinger et al., 2013] relative to LR samples may complicate differential diagnosis in very young children. As well, there may be other factors related to having an older child with ASD in the family and deciding to participate in a HR cohort study, that influence parental expectations, and in turn, clinical decision-making. Our experience with this HR cohort suggests that the older sibling with ASD may serve as important reference for comparison for parents. When the younger child is perceived to be less impaired than their older sibling, this can generate ambivalence about considering an early diagnosis, which may influence clinical processes surrounding communication of concerns. We are currently embarking on a qualitative study to better understand parental and clinician experience in this context and potential influences on diagnostic decision-making. For many reasons, caution should be exercised in generalizing the findings of this HR cohort study to children with ASD ascertained by clinical referral. Second, different language measures are used at different time points; specifically, the MSEL language subscales were not available at 18 months, so the MacArthur child development inventory reported as an alternate language measure. Although not normed at 18 months, it was used for comparative purposes, as in previous ASD research (e.g., Charman et al., 2003). Finally, we acknowledge that follow-up to 3 years may not provide the complete story on diagnostic stability, as the status of some children may continue to change. Further follow-up of our HR cohort to age 9 years is underway.

Conclusions

The success of early identification and screening initiatives depends on clinicians’ ability to establish valid and stable ASD diagnoses in very young children. Despite accumulating clinical experience with diagnostic assessment of toddlers, establishing an ASD diagnosis in children under 2 years, particularly those without language delays, remains a special challenge. Further efforts to improve diagnostic measures and hone diagnosticians’ clinical acumen are essential to achieving the goal of earlier diagnosis of children across the autistic and developmental spectra. Opportunities to improve outcomes through earlier initiation of intervention, as well as to lessen the burden on families associated with uncertainties about diagnosis, certainly justify these efforts.

Acknowledgments

This research project was funded by the Canadian Institutes of Health Research (grant numbers 62924 and 102665), Autism Speaks Canada and NeuroDevNet. Dr. Zwaigenbaum was supported by the Stollery Children’s Hospital Foundation Chair in Autism Research. Drs. Bryson and Smith were supported by the Jack and Joan Craig Chair in Autism Research, Dr. Szatmari is supported by the Chedoke Health Chair in Child Psychiatry, and Dr. Vaillancourt is supported by a Canada Research Chair in Children’s Mental Health and Violence Protection. We also like to thank the children and families who have participated in this project.

References


