Non-ASD Outcomes at 36 Months in Siblings at Familial Risk for Autism Spectrum Disorder (ASD): A Baby Siblings Research Consortium (BSRC) Study

Tony Charman[1*], Greg S. Young[2], Jessica Brian[3], Alice Carter[4], Leslie J. Carver[5], Katarzyna Chawarska[6], Suzanne Curtin[7], Karen Dobkins[5], Mayada Elsabbagh[8], Stelios Georgiades[9], Irva Hertz-Picciotto[2], Ted Hutman[10], Jana M. Iverson[11], Emily J. Jones[12], Rebecca Landa[13], Suzanne Macari[5], Daniel S. Messinger[14], Charles A. Nelson[15,16,17], Sally Ozonoff[2], Celine Saulnier[18], Wendy L. Stone[19], Helen Tager-Flusberg[20], Sara Jane Webb[18,21], Nurit Yirmiya[22], , Lonnie Zwaigenbaum[23]

1 King's College London

2 University of California, Davis

3 University of Toronto

4 University of Massachusetts, Boston

5 University of California, San Diego

6 Yale University School of Medicine

7 University of Calgary

8 McGill University

9 McMaster University

10 University of California, Los Angeles

11 University of Pittsburgh

12 Birkbeck College London

13 Kennedy Krieger Institute and John Hopkins School of Medicine

14 University of Miami

15 Harvard Medical School

16 Harvard Graduate School of Education

17 Boston Children's Hospital

18 Emory University School of Medicine

19 University of Washington

20 Boston University

21 Seattle Children's Research Institute

22 Hebrew University of Jerusalem

23 University of Alberta

* Correspondence to:
Tony Charman
tony.charman@kcl.ac.uk
Department of Psychology PO 77
Institute of Psychiatry, Psychology & Neuroscience
King's College London
De Crespigny Park
London
SE5 8AF
UK

Word count text only = 5,460

Number of text pages (main text) = 18

Number of Tables = 5

Number of Figures = None

Plus Supplementary Materials (online only) = 4 Tables

Lay Abstract

This study characterised developmental outcomes of a large sample of siblings at familial high-risk of autism spectrum disorder (ASD), who themselves did not have ASD (n = 859), and low-risk controls with no family history of ASD (n = 473). We characterised outcomes at age 3 years using a developmental assessment of language and learning and an observational measure of ASD symptoms and, where available, parent interviews about ASD behaviours and adaptive functioning. Around one-in-ten high-risk siblings had mild-to-moderate levels of developmental delay a rate significantly higher than the low-risk controls although the groups did not differ in the proportion of toddlers with mild-to-moderate language delay. High-risk siblings were also more likely to have higher levels of observer-rated and parent-reported levels of ASD symptoms and lower adaptive functioning. Males were also more likely to show higher levels of ASD symptoms and lower levels of developmental ability and adaptive behaviour than females across most measures but not mild-to-moderate language delay. Lower maternal education was also associated with lower developmental and adaptive behaviour outcomes. We discuss these findings as evidence for early emerging characteristics related to the 'broader autism phenotype' previously described in older family members of individuals with ASD. There is a need for ongoing clinical monitoring of high-risk siblings who do not show clear signs of ASD by age 3 years, as well as continued follow-up into school age to determine their developmental and behavioural outcomes.

Word count = 236

Scientific Abstract

This study characterised developmental outcomes of a large sample of siblings at familial high-risk of autism spectrum disorder (ASD), who themselves did not have ASD (n = 859), and low-risk controls with no family history of ASD (n = 473). We characterised developmental outcomes at age 3 years using the Mullen Scales of Early Learning, the Autism Diagnostic Observation Schedule (ADOS) and, where available, the Autism Diagnostic Interview – Revised (ADI-R) and adaptive functioning on the Vineland Adaptive Behavior Scales. Around 11% of high-risk siblings had mild-to-moderate levels of developmental delay a rate significantly higher than the low-risk controls, although the groups did not differ in the proportion of toddlers with mild-to-moderate language delay. Thirty percent of high-risk siblings had elevated scores on the ADOS, double the rate seen in the low risk controls. High-risk siblings also had higher parent reported levels of ASD symptoms on the ADI-R and lower adaptive functioning on the Vineland. Males were also more likely to show higher levels of ASD symptoms and lower levels of developmental ability and adaptive behaviour than females across most measures but not mild-to-moderate language delay. Lower maternal education was also associated with lower developmental and adaptive behaviour outcomes. We discuss these findings as evidence for early emerging characteristics related to the 'broader autism phenotype' (BAP) previously described in older family members of individuals with ASD. There is a need for ongoing clinical monitoring of high-risk siblings who do not show clear signs of ASD by age 3 years, as well as continued follow-up into school age to determine their developmental and behavioural outcomes.

Word count = 263

Key Words: autism spectrum disorder, broader autism phenotype, developmental outcomes, high risk siblings, adaptive functioning

Non-ASD Outcomes at 36 Months in Siblings at Familial High-Risk for Autism Spectrum Disorders (ASD): A Baby Siblings Research Consortium (BSRC) Study

Research on infant siblings at familial high-risk (HR) of autism spectrum disorder (ASD) has established that close to 20% of HR siblings have ASD themselves by the age of 36 months. Using pooled data from collaborating sites in the Baby Siblings Research Consortium (BSRC), recurrence was 18.7% from a sample of *n* = 664 HR siblings (Ozonoff et al., 2011) and 19.5% in an expanded cohort of *n* = 1,241 HR siblings (Messinger et al., 2015). A previous report from the BSRC on HR siblings (n = 507) who did not have an ASD outcome at 36 months used latent class analysis to subgroup these children (Messinger et al., 2013). Groups were formed based on scores on a symptom measure (Autism Diagnostic Observation Schedule (ADOS); Lord et al., 2000) and a standardised developmental assessment (Mullen Scales of Early Learning (MSEL); Mullen, 1995). Twenty one percent of non-ASD HR siblings were classified in groups with higher ASD severity scores and/or lower levels of developmental function, a profile found in only 7% of low-risk (LR) controls without a family history of ASD (Messinger et al., 2013).

Understanding more about non-ASD outcomes in HR siblings at an early age would allow us to study the early emergence of the broader autism phenotype (BAP) – subclinical traits or characteristics that are present at an elevated rate in families containing individuals with autism (Bolton et al., 1994; Folstein & Rutter, 1977; Pickles et al., 2000; Piven et al., 1997). The term 'BAP' has been used in different ways in the literature, sometimes including only subclinical features closely aligned to the core diagnostic features of ASD (e.g., social communication and pragmatic language difficulties, behavioural rigidity) and other times referring to a broader range of characteristics that are elevated in family members and associated with, but

not ‘core’ to, the definition of ASD (e.g., co-occurring psychiatric disorders, intellectual disability) (see Sucksmith, Roth & Hoekstra, 2011; for a review).

Groups using the prospective HR sibling design have reported non-ASD developmental outcomes (see Szatmari et al., 2016; for a review). Elevated rates of sub-clinical ASD symptoms (characterised as the BAP), symptoms of emergent attention deficit hyperactivity disorder (ADHD), and lower language and developmental abilities were found in around one quarter of HR siblings who do not have ASD at 36 months of age in a sample from two BSRC sites (Ozonoff et al., 2014). Another report from an overlapping cohort found elevated rates of pragmatic language difficulties in HR siblings who do not have ASD at 36 months; in most cases these children did not have more general language impairments (Miller et al., 2015).

The current study utilises an expanded BSRC HR sample (relative to Messinger et al., 2013) and reports on the outcomes at 36 months of age in HR infants who do not have ASD (n = 859) compared to LR controls, also without ASD (n = 473). In contrast to the statistically derived classification reported by Messinger et al. (2013), we adopt a more clinical framework by reporting outcomes in terms of children who have mild-to-moderate global developmental and/or language delays and those who exhibit elevated scores on the ADOS but who were not categorised as having ASD. The current approach therefore allows us to indicate non-ASD developmental outcomes for individual children. Where available, additional information on autism symptoms (Autism Diagnostic Interview–Revised (ADI-R); Lord et al., 1994) and adaptive functioning (Vineland Adaptive Behavior Scales (Vineland); Sparrow et al., 1984, 2005) is also used to characterise outcomes. Together, this information allows us to report rates of non-ASD developmental

difficulties (mild-to-moderate developmental and language delay) as well as sub-clinical levels of ASD symptoms and adaptive behaviour relevant to the emergent BAP in these HR non-ASD siblings.

Methods

Participants

In line with previous BSRC reports (Chawarska et al., 2014; Messinger et al., 2013; Ozonoff et al., 2011), ASD case definition was a consensus best estimate (CBE) diagnosis of ASD (using DSM-IV (APA, 2000), DSM-5 (APA, 2013) or ICD-10 (WHO, 1993)) and scoring at or above the ASD threshold on the ADOS (calibrated severity score (CSS) ≥4; Gotham et al., 2009). Of the 2,099 infants/toddlers included in this dataset, 620 were low-risk (LR) infants (no first degree relative with ASD), of whom 3 (0.48%) met ASD criteria at age 3 years, and 1,479 were high-risk (HR) infants (at least 1 older sibling with ASD), of whom 275 (18.59%) met ASD criteria. The aim of the paper is to describe 36 month outcomes in HR siblings who do not have ASD, so the 278 cases meeting CBE ASD criteria (from both the HR and LR groups) were removed from the sample. This left 1,204 HR siblings and 617 LR infants. The primary outcome measures are the ADOS-CSS and the MSEL, so children missing either or both were excluded (2 sites did not provide ADOS data and a different 2 sites did not provide MSEL data; from other sites ADOS and MSEL data were incomplete), leaving n = 859 HR siblings and n = 473 LR controls (total n = 1,332) from 9 sites. Vineland (n = 895) and ADI-R (n = 600) data were available on a sub-set of sample.

Measures

The *Mullen Scales of Early Learning* (*MSEL*; Mullen, 1995) is a standardised developmental assessment for children aged between birth and 68 months. It yields a global development quotient, the Early Learning Composite (ELC), with a mean of 100 and a standard deviation of 15. We report T-scores (mean of 50, SD 10) averaged across the two verbal (Expressive language, Receptive language) and two non-verbal (Fine motor, Visual reception) subscales.

The *Autism Diagnostic Observation Schedule* (*ADOS*; Lord et al., 2000) is a play-based, observer-rated assessment of ASD symptoms. Different modules are used depending on the child's language ability and, at the 36 month assessment, 141 children completed Module 1 (no words or single words only) and 1,191 children completed Module 2 (phrase speech). The ADOS-CSS ranges from 1 to 10, with the threshold for an ASD diagnosis being 4 or greater (Gotham et al., 2009).

The *Autism Diagnostic Interview – Revised* (*ADI-R*; Lord et al., 1994) is an informant-based, examiner-rated interview of ASD symptoms. It yields domain scores covering social ('reciprocal social interaction'), communication ('communication and language') and repetitive ('restricted and repetitive, stereotyped interests and behaviours') symptoms.

The *Vineland Adaptive Behavior Scales* (Sparrow et al., 1984, 2005) is an informant interview on everyday adaptive functioning. It yields Socialization, Communication, Daily Living Skills and Motor domain scores and an overall Adaptive Behavior Composite score (ABC), all having a mean of 100 and a standard deviation of 15. Approximately half the sample had Vineland scores from the first edition (Sparrow et al., 1984; $n = 423$) and half from the second edition (Sparrow et al., 2005, $n = 472$).

Characterising developmental outcomes

We characterised atypical developmental outcomes in terms of: (i) elevated ASD symptom expression as assessed by the ADOS and ADI-R; and (ii) below average general developmental and language abilities as assessed by the MSEL and adaptive function as assessed by the Vineland. We first defined mild-to-moderate Developmental Delay as a Mullen ELC >1 SD below the mean (i.e., below 85) and then mild-to-moderate Language Delay as expressive language (EL) and/or receptive language (RL) >1SD below the mean (i.e., T-score below 40), so the two subgroups were mutually exclusive. We defined 'elevated' ASD symptoms using an ADOS CSS threshold of ≥3 (where 3 is one point below the ASD diagnostic threshold) to include subthreshold levels of ASD behaviours (see Chawarska et al., 2014). For the ADI-R we used 'sub-clinical-threshold' cut points of ≥8 for the Social, ≥6 for the Communication and ≥2 for the Stereotyped, Repetitive and Rigid Behaviour (RRB) domains, respectively. For the Vineland we defined mild-to-moderate adaptive behaviour delay as a standardised score >1 SD below the mean (i.e., below 85) on each domain score and the ABC.

Statistical analysis

Proportions of HR siblings and LR controls falling into each categorical outcome group were analysed by multinomial logistic regression (for MSEL-defined nominal outcomes) and logistic regression (for the remaining dichotomous outcomes) with relative risk ratios (RRR), odds ratios (OR), and 95% confidence intervals being reported, respectively. Background variables on which the HR and LR groups were unmatched were entered first and retained in the models when they were significantly associated to outcomes (see below). Risk group (HR vs. LR) and sex (male vs. female) were then entered into the models. The interaction between sex and risk group was then entered and retained if significant. Finally, BSRC site was entered

into the models. Following our 'clinically defined outcomes' approach we report the proportion of male and females HR and LR children falling into each outcome group in the main Tables but we also present the continuous scores on the measures in Supplementary Tables 1 to 4 to aid comparison with previous literature.

Results

Sample characteristics are shown in Table 1. The LR and HR groups were comparable on background variables with the exception of age-first-seen, which was higher in the HR siblings compared to the LR controls ($F(1, 1{,}330) = 10.25$, $p < .01$), and maternal education, which was higher in the LR controls than in the HR siblings (χ^2 (1, N = 1,194) = 13.78, $p < .001$). Both variables were entered first into the statistical models and retained where significant.

----- Table 1 about here -----

Rates of Developmental Delay and Language Delay

Table 2 shows the rates of the HR and LR children falling into the MSEL-defined mild-to-moderate Developmental Delay and Language Delay groups. A multinomial logistic regression indicated that both (lower) maternal education and (higher) age-first-seen were significantly associated with Developmental Delay and Language Delay outcomes (likelihood ratio (LR) χ^2 (2, n = 1,194) = 49.25, $p < .001$). The relative risk ratio (RRR) of being in the Developmental Delay group (vs. the No Developmental/ No Language Delay group) was 0.59 (95% CIs: 0.48, 0.73, $p < .001$) across the 5-point maternal education scale and 0.60 (0.47, 0.78, $p < .001$) of being in the Language Delay group. The RRR of being in the Developmental Delay group was 1.08 (1.02, 1.13, $p < .01$) for each month of age and 1.09 (1.03, 1.16, $p < .01$) of

being in the Language Delay group[1]. With both maternal education and age-first-seen retained (LR χ^2 (6, n = 1,194) = 83.43, p < .001), risk group and sex were added to the model that remained significant (LR χ^2 (8, n = 1,194) = 99.02, p < .001). The RRR of being in the Developmental Delay group for the HR vs. the LR group was 2.84 (1.62, 5.01, p < .001) and 4.01 (2.41, 6.68, p < .001) for males vs. females. Lower maternal education continued to be associated with a higher risk of being in the Developmental Delay group (p < .001) as did higher age-first-seen (p < .05). Neither risk group (p = .45) nor sex (p = .11) was associated with being in the Language Delay group but lower maternal education (p < .001) and higher age-first-seen (p < .01) were. The interaction of sex and risk group was not significant with being in the Developmental Delay (p = .58) or Language Delay (p = .38) groups. BSRC site was entered last and was not associated with being in either the Developmental Delay group (p = .49) or the Language Delay group (p = .21).

----- Table 2 about here -----

Elevated levels of ASD symptoms

The proportion of children with an elevated ADOS-CSS (≥ 3) was higher in the HR group (n = 250/859, 29.10%) compared to the LR group (n = 80/393, 16.91%) – see Table 3. A logistic regression (Likelihood ratio (LR) χ^2 (2, N = 1,1194) = 2.08, p = .35) indicated that neither maternal education (p = .77) nor age-first-seen (p = .16) was associated with an elevated ADOS CSS and both were dropped from the model. When risk group and sex were entered the overall model was significant ((LR) χ^2 (2, N = 1,332) = 34.42, p < .001). There was a main effect of risk group for being in the elevated ADOS CSS group for HR vs. LR status ((OR) = 2.05 (1.54, 2.72), p < .001)

[1] When infants first seen above 12 months of age were excluded age-first-seen was no longer associated with mild-to-moderate Developmental Delay or mild-to-moderate Language Delay.

and a main effect effect of sex for males vs. females ((OR) = 1.47 (1.14, 1.90), $p <$ .01). The interaction of sex by risk group was not significant ($p = .95$). There was also a main effect of BSRC site ($p < .05$).

----- Table 3 about here -----

Table 4 shows elevated ADI-R scores by risk group and sex. For the ADI-R Social domain, a logistic regression ((LR) χ^2 (2, $N = 512$) = 4.59, $p = .10$) indicated that neither maternal education (p = .07) nor age-first-seen (p = .18) was associated with an elevated score. When risk group and sex were entered the overall model was significant ((LR) χ^2 (2, $N = 600$) = 12.88, $p < .01$). There was a main effect of risk group for HR vs. LR ((OR) = 4.37 (1.30, 14.67), $p < .05$) and a main effect of sex for males vs. females ((OR) = 2.51 (1.12, 5.62), $p < .05$). Neither the interaction of sex by risk group ($p = .78$) nor BSRC site was significant ($p = .22$).

For the ADI-R Communication domain although the overall model including age-first-seen and maternal education was not significant ((LR) χ^2 (2, $N = 513$) = 4.35, $p = .11$), maternal education was significantly associated with an elevated score ((OR) = 0.68 (0.47, 0.99), $p < .05$) and so was retained in the model but age-first-seen was not (p = .42). When risk group and male sex were entered into the model the overall model was significant ((LR) χ^2 (3, $N = 513$) = 18.19, $p < .001$) there was a main effect of risk group for HR vs. LR ((OR) = 4.19 (1.23, 14.29), $p < .05$) and a main effect of sex for males vs. females ((OR) = 3.00 (1.29, 6.99), $p < .05$) but the effect of maternal education was not longer significant ($p = .10$). Neither the interaction of sex by risk group ($p = .75$) nor BSRC site was significant ($p = .21$).

For the ADI-R RRB domain, a logistic regression ((LR) χ^2 (2, $N = 511$) = 3.26, $p = .20$) indicated that neither maternal education (p = .07) nor age-first-seen (p

= .83) was associated with an elevated score. When risk group and male sex were entered into the model the overall model was significant ((LR) χ^2 (2, N = 598) = 16.52, p < .001) there was a main effect of risk group for HR vs. LR ((OR) = 2.40 (1.47, 3.93), p < .001) but the main effect of sex failed to reach significance (p = .07) and there was no interaction between risk group and sex (p = .12) nor an effect of BSRC site (p = .66).

----- Table 4 about here -----

Table 5 shows the subgroups with lower Vineland scores by risk group and sex. For the Vineland Communication domain the overall model including age-first-seen and maternal education was significant ((LR) χ^2 (2, N = 722) = 18.37, p < .001) and maternal education was significantly associated with an elevated score ((OR) = 0.53 (0.40, 0.72), p < .001) and so was retained in the model but age-first-seen was not (p = .11). When risk group and male sex were entered into the model the overall model was significant ((LR) χ^2 (3, N = 722) = 38.20, p < .001) there was a main effect of risk group for HR vs. LR ((OR) = 7.91 (2.41, 26.03), p < .01) and a main effect of maternal education ((OR) = 0.58 (0.42, 0.79), p < .01) but no effect of sex (p = .18). Neither the interaction of sex by risk group (p = .39) nor BSRC site was significant (p = .75).

For the Vineland Daily Living Skills domain a logistic regression ((LR) χ^2 (2, N = 700) = 3.64, p = .17) indicated that neither maternal education (p = .55) nor age-first-seen (p = .07) was associated with an elevated score and both were dropped from the model. When risk group and male sex were entered into the model the overall model was significant ((LR) χ^2 (2, N = 772) = 26.70, p < .001) there was a main effect of risk group for HR vs. LR ((OR) = 2.37 (1.51, 3.73), p < .001) and a main effect of

sex for males vs. females ((OR) = 1.93 (1.29, 2.89), $p < .01$). The interaction of sex by risk group was not significant ($p = .67$) but BSRC site was ($p < .01$).

For the Vineland Socialization domain the overall model including age-first-seen and maternal education was significant ((LR) χ^2 (2, $N = 824$) = 15.88, $p < .001$) and maternal education was significantly associated with an elevated score ((OR) = 0.63 (0.51, 0.79), $p < .001$) and so was retained in the model but age-first-seen was not (p = .86). When risk group and male sex were entered into the model the overall model was significant ((LR) χ^2 (3, $N = 824$) = 32.95, $p < .001$) there was a main effect of risk group for HR vs. LR ((OR) = 2.50 (1.45, 4.31), $p < .01$) and a main effect of sex for males vs. females ((OR) = 1.63 (1.03, 2.59), $p < .05$) and the effect of maternal education remained significant ((OR) = 0.66 (0.53, 0.83), $p < .001$). The interaction of sex by risk group was not significant ($p = .69$) but BSRC site was ($p < .01$).

For the Vineland Motor domain the overall model including age-first-seen and maternal education was significant ((LR) χ^2 (2, $N = 700$) = 13.57, $p < .01$) and both maternal education ((OR) = 0.79 (0.64, 0.99), $p < .05$) and age-first-seen ((OR) = 1.09 (1.03, 1.15), $p < .01$) were significantly associated with an elevated score and were retained in the model. When risk group and male sex were entered into the model the overall model was significant ((LR) χ^2 (4, $N = 700$) = 32.62, $p < .001$), there was a main effect of risk group for HR vs. LR ((OR) = 2.82 (1.65, 4.82), $p < .01$) and the effect of age-first-seen remained significant ((OR) = 1.08 (1.02, 1.14), $p < .05$) but neither sex ($p = .15$) nor maternal education ($p = .16$). Neither the interaction of sex by risk group ($p = .35$) nor BSRC site was significant ($p = .83$).

For the Vineland ABC the overall model including age-first-seen and maternal education was significant ((LR) χ^2 (2, N = 698) = 57.72, $p < .001$) and both maternal education ((OR) = 0.65 (0.51, 0.82), $p < .001$) and age-first-seen ((OR) = 1.09 (1.03, 1.16), $p < .01$) were significantly associated with an elevated score and were retained in the model. When risk group and male sex were entered into the model the overall model was significant ((LR) χ^2 (4, N = 698) = 57.72, $p < .001$), there was a main effect of risk group for HR vs. LR ((OR) = 7.28 (3.09, 17.15), $p < .001$) and a main effect of sex for males vs. females ((OR) = 1.81 (1.08, 3.06), $p < .05$) and the effects of maternal education ($p < .01$) and age-first-seen remained significant ($p < .05$)[2]. The interaction of sex by risk group was not significant ($p = .82$) but BSRC site was ($p < .05$).

----- Table 5 about here -----

*Comparison of 36 month old children who did and did not receive the ADI-*R *and Vineland*

We tested whether risk status, low MSEL score and elevated ADOS scores were related to which children received the ADI-R and Vineland. A logistic regression including risk status (HR vs. LR), ADOS CSS (≥3 vs. < 3), MSEL ELC (≥85 vs. < 85) and BSRC site indicated that these factors influenced whether ADI-Rs were completed (Likelihood ratio (LR) $\chi^2 = 44.70$, df = 4, $p < .001$). The OR of the ADI-R being completed for HR vs. LR toddlers was 1.31 (1.04 – 1.66), $p < .05$, for elevated vs. non-elevated ADOS CSS 1.34 (1.04 – 1.73), $p < .05$, for BSRC site 0.96 (0.94 – 0.98), $p < .0001$) but was not related to low MSEL score ($p = .10$). The OR of the Vineland being completed for HR vs. LR toddlers was 0.70 (0.55 – 0.88), $p < .01$ and for BSRC site 0.98 (0.96 – 0.99), $p < .01$ but was not related to elevated ADOS

[2] When infants first seen above 12 months of age were excluded age-first-seen was associated only with Vineland Daily Living Skills ($p < .05$) and Socialization ($p < .01$) domains.

CSS (p = .06) or low MSEL score (p = .53).

Discussion

Rates of Developmental Delay, Language Delay and lower adaptive behaviour

Rates of mild-to-moderate Developmental Delay (10.59%) at 36 months of age were approximately 3 times in HR toddlers without an ASD diagnosis than in the LR controls (3.38%). In contrast, rates of mild-to-moderate language delay did not differ between HR toddlers without ASD (6.87%) and LR controls (5.07%). Male sex was also associated with mild-to-moderate developmental delay (but not language delay) and lower maternal education was associated with both.

The increased rate of developmental delay in HR non-ASD siblings is consistent with the previous latent class analysis of a sub-set of the current sample (Messinger et al., 2013). However, in contrast to previous single site studies that have reported increased rate of language delays in HR siblings (Landa et al., 2012; Ozonoff et al., 2014) we did not find elevated rates of language delay in HR non-ASD siblings. Previous studies of older siblings/family members of individuals with an ASD have been inconsistent in findings of language delay and general developmental delay (Bartak, Rutter & Cox, 1975; Fombonne et al., 1997; Lindgren et al., 2009; Pilowsky et al., 2003; Szatmari et al., 1993) and there is not a clear consensus that either should be considered part of the BAP (see, Sucksmith et al., 2011). It might be that the MSEL captures only structural language developmental and that aspects of the pragmatic use of language and communication are impaired in a subgroup of HR siblings (Miller et al., 2015). Males were at greater risk than females of having mild-to-moderate levels of Developmental Delay but not Language Delay outcomes. This is consistent with the group level analysis of mean scores from an overlapping dataset that reported that males had lower Mullen subscale scores, regardless of risk group

and outcome (Messinger et al., 2015).

We defined 'mild-to-moderate' Developmental and Language Delay by MSEL scores >1 SD below the mean. Under a normal standardisation curve this would include ~16.5% of the population, whereas our rates are considerably lower in both the HR and LR groups. However, the current sample had high levels of maternal education (~70% of the HR mothers and ~80% of the LR mothers were educated to college level) and is not representative of the broader population and higher maternal education was associated with lower rates of both Developmental Delay and Language Delay. It is therefore likely that the prevalence of developmental difficulties seen in non-ASD HR siblings more generally may be higher than in our self-selecting research samples.

For the participants for whom scores on the Vineland was also available rates of lower adaptive function were found for HR siblings compared to LR controls on all Vineland subdomains and for the Adaptive Behavior Composite (ABC) and the effect sizes were large with odds ratios of greater than 7 for the Communication domain and ABC. Males were more likely than females to show lower levels of adaptive behaviour in some domains (Daily Living Skills, Socialization, ABC) but not others (Communication, Motor). Lower maternal education was also associated with lower Communication, Socialization and overall adaptive behaviour. Although there was some overlap between toddlers with mild-to-moderate Developmental Delay on the MSEL and mild-to-moderate adaptive delay on the Vineland the two groups were not coincident with only 28/84 (33.33%) of those with Vineland ABC scores <85 also having MSEL ELC scores <85. Thus, some of the risk of lower developmental outcomes for HR siblings (and males) is separate for developmental abilities as measured by the Mullen and adaptive functioning as measured by the Vineland. It

was also the case that there were no significant interactions between risk group and sex indicating that the risk effects were independent.

In the full sample older age-first-seen was also associated with Developmental Delay and Language Delay but when infants first seen above 12 months of age were excluded this was no longer the case, suggesting that later recruited infants might show a bias towards increased parental concern, possibly due to early emerging developmental difficulties in their child.

Elevated levels of ASD symptoms

When a threshold of ≥3 was set for an elevated ADOS CSS (i.e., one point below the ASD threshold), then ~15% of LR controls and ~30% of HR siblings fell into this category. The elevated rate of ASD symptoms in HR siblings who do not have ASD is consistent with the notion of emergent BAP characteristics in the HR toddlers and in line with the pattern found in older family members in previous samples (Constantino et al., 2006; Pickles et al., 2000; Piven et al., 1997). It is also consistent with the report by Georgiades et al. (2013) of even earlier emergent BAP characteristics measured around 12 months of age using the Autism Observation Scale for Infants (AOSI; Bryson et al., 2008). Male sex was also associated with having an elevated ADOS CSS in line with the finding elevated ADOS RRB but not SA score) reported in a continuous analysis of a largely overlapping sample reported by Messinger et al. (2015).

For the participants for whom scores on the ADI-R were also available HR siblings were more likely to show elevated scores than LR controls on all three domains and males more likely than females on the Social and Communication but not the Repetitive Behaviour domain (p = .07). This confirms by parental report the findings from the ADOS observational interaction with an unfamiliar examiner,

indicating that sub-clinical characteristics of the BAP are found in at least a proportion of HR siblings who do not have an ASD by 3 years of age. It is important to note that the HR sibling group mean scores on the ADOS (Table S2) and ADI-R (Table S3) are well below the clinical cut-points for ASD on both measures and that only a relatively small minority fall above the threshold for 'elevated' ASD symptoms, arbitrarily set as 2 points below the clinical threshold on the ADI-R Social and Communication domains and 1 point below the clinical threshold for ADI-R RRB and the ADOS CSS. Thus, it appears that characteristics of the BAP might only present in some but not all HR siblings who do not have an ASD by the age of 3 years.

Why do ~15% of LR controls have elevated ADOS CSS scores? We cannot be sure whether these findings necessarily reflect sub-threshold ASD symptoms or behavioural features such as inattention/over-activity, non-compliance and social anxiety or social inhibition, that can result in elevated ADOS scores in toddlers at this age. However, in the HR sibling group, we have suggested that the higher (double) rates of elevated ADOS scores might represent sub-clinical traits of ASD characteristics or BAP phenomena. Does this suggest that this 'BAP characteristic' might be present in as many as one-in-seven LR controls? It is now widely recognised that ASD traits are broadly distributed in the general population without a clear boundary between individual variation and psychopathology (Constantino, 2011; Robinson et al., 2011). An alternative explanation is that scores on any instrument do not always mean the same thing when participants have been purposively sampled in different ways, as is inherent in the HR siblings vs. LR control design. It might be the case that elevated ADOS scores in the LR controls occur for different reasons, such as social inhibition, shyness, or poor attention.

Without independent measures, using methods such as psychophysiological arousal, we cannot know if such a phenomenon might be operating between the HR siblings and LR controls with elevated ADOS scores, and this should be tested in future studies.

Limitations

The present study has several strengths, including a large sample, use of standard assessment measures across sites and characterisation of the sample to a diagnostic age of 36 months when clinical features are stable (Ozonoff et al., 2015). However, it also has some limitations. The ADI-R and Vineland were only available on a subsample. In addition, risk group, ADOS scores, and site were all associated with whether an ADI-R or a Vineland was administered, so the children for whom these additional measures are available might not be representative of the broader sample and the results may reflect some clinical concern which we have not systematically been able to capture. The high levels of maternal education seen in both the HR and LR groups indicate that the current research samples are not representative of the wider population, so caution needs to be exercised regarding the extent to which the current findings will generalise to HR siblings more generally. Finally we have used a relatively low threshold to indicate 'delay' on the Mullen and the Vineland – scores greater than one SD below the population mean. When more conventional clinical thresholds were set for defining Developmental Delay (>2SD below the mean) and Language Delay (>1.5SD below the mean) the rates of delay in the HR sibling group were 2.79% and 1.40%, respectively.

Conclusions

Levels of mild-to-moderate developmental delays but not language delays are elevated in HR siblings who at 36 months do not have ASD compared to LR controls.

In addition to identifying those HR siblings who go on to have ASD by age 3, paediatricians and other healthcare practitioners should continue to offer surveillance of all young siblings of children with ASD to monitor their developmental progress (Zwaigenbaum et al., 2015). Subclinical levels of ASD symptoms are relatively common in HR siblings with an ASD but are also present, though at a significantly reduced rate, in LR controls. A sub-group of HR siblings also have higher parent-reported levels of ASD symptoms and poorer adaptive behaviour – and this pattern is consistent with descriptions of the BAP in older family members. Because the HR sibling and LR control groups are differently and purposively sampled, we cannot be sure that the reasons underlying *why* a HR child and a LR child receives a CSS score greater than or equal to 3 on the ADOS are the same. In some children, this score may reflect the presence of ASD symptoms, whereas in others it may reflect social inhibition, inattention/overactivity or other traits – all of which have a broad distribution in the general population but also which are known to be present at elevated rates in individuals with ASD and their family members (Bolton et al., 1998; Piven & Palmer, 1999).

What will happen to the ~30% of HR siblings (and ~15% of LR controls) who have somewhat elevated ADOS CSS as they develop into school age? Will these mostly sub-clinical but slightly elevated levels of ASD symptoms increase, remain stable, or decrease, and will there be any evidence that they have a functional impact on these children? We know from HR sibling cohorts (Brian et al., in press; Ozonoff et al., 2015) that a diagnosis of ASD is made at different ages in different children, even by expert clinical research teams. We also know from clinical practice that, in some children, obvious ASD symptomatology and accompanying impairment does not emerge until they are in school, where the social demands exceed their capacity

(*pace* DSM-5). Some of the HR siblings who had elevated ADOS scores at 36 months but did not meet a CBE diagnosis of ASD might meet ASD criteria when reassessed in the school-age years (Brian et al., in press).

In line with Ozonoff et al. (2014) and Miller et al. (2015), the fact that non-ASD HR siblings show elevated rates (but largely below clinical thresholds) of ASD symptoms and lower levels everyday adaptive functioning suggests that these are features of 'the emergent BAP'. Follow-up of these children into school age will be necessary to determine whether this pattern might reflect emergent ('late onset') ASD, particularly in those whose general abilities are above the average range, or other features that relate to the BAP concept, such as pragmatic language difficulties or other forms of emotional and behavioural psychopathology (e.g. ADHD, anxiety disorder). This will also allow us to determine whether the early emerging BAP described in the current sample shows continuity with the more established BAP phenomenon described in older children and adults in the literature (Pruett, 2014).

Acknowledgements

Autism Speaks provided support for the creation of the Baby Siblings Research Consortium (BSRC) Database. Data collection, analyses, and manuscript preparation were supported by the National Institutes of Health (Nelson R01DC010290; Messinger R01HD057284; Messinger R01GM105004; Tager-Flusberg R01DC010290; Tager-FlusbergR21DC008637; Chawarska R01MH087554; Chawarska PO1HD003008, Project 1; Iverson R01HD054979; Landa R01MH059630; Ozonoff R01MH068398; Webb P50HD055782; Webb R01HD064820; Hutman K01MH096961; Hutman U54 MH068172; Hutman P50 HD55784; Stone R01HD057284); the Canadian Institutes of Health Research (Bryson, Zwaigenbaum); the UK Medical Research Council (Charman); NeuroDevNet (Bryson, Zwaigenbaum), Autism Speaks Canada (Bryson, Zwaigenbaum); Autism Speaks [U.S.] Pilot Grant (Tager-Flusberg, Iverson); Autistica (Charman); the FAR Fund (Chawarska); and the Simons Foundation (Nelson).

Human subjects oversight and approval was provided by The University of Miami Institutional Review Board (Messinger), University of California Davis Institutional Review Board (Young, Ozonoff, Herz-Picciotto), University of Washington Institutional Review Board (Webb, Stone), Vanderbilt University Institutional Review Board (Stone), IWK Health Centre Research Ethics Board (Bryson), UCSD Human Research Protection Program (Carver, Dobkins), London MREC 06/MRE02/73 (Charman), Yale University School of Medicine Human Investigation Committee (Chawarska), The Conjoint Health Research Ethics Board University of Calgary (Curtin), UCLA Institutional Review Board (Hutman), University of Pittsburgh Institutional Review Board (Iverson), Johns Hopkins

Medicine Institutional Review Board (Landa), Boston Children's Hospital Institutional Review Board (Nelson, Tager-Flusberg, Carter), Health Research Ethics Board for University of Alberta and Alberta Health Services (Zwaigenbaum).

Competing Interest

Dr. Zwaigenbaum receives operating funds but not honoraria from SynapDx. Dr. Charman receives royalties from Guilford Press and Sage Publications. Drs. Brian, Bryson, Carter, Carver, Chawarska, Curtin, Dobkins, Elsabbagh, Georgiades, Hertz-Picciotto, Hutman, Iverson, Jones, Landa, Macari, Messinger, Nelson, Ozonoff, Saulnier, Stone, Tager-Flusberg, Webb, Yirmiya and Young declare that they have no competing interests.

References

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th text revised ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*. Washington DC: American Psychiatric Association.

Bartak, L., Rutter, M., Cox, A. (1975). A comparative study of infantile autism and specific development receptive language disorder. I. The children. *British Journal of Psychiatry, 126 (2)*, 127-145

Bolton, P., Macdonald, H., Pickles, A., Rios, P., Goode, S., Crowson, M., . . . Rutter, M. (1994). A case - control family history study of autism. *Journal of Child Psychology and Psychiatry, 35*(5), 877-900. doi: 10.1111/j.1469-7610.1994.tb02300.x

Bolton, P. F., Pickles, A., Murphy, M., & Rutter, M. (1998). Autism, affective and other psychiatric disorders: patterns of familial aggregation. *Psychological Medicine, 28*(2), 385-395. doi: 10.1017/s0033291797006004

Brian, J., Bryson, S. E., Smith, I. M., Roberts, W., Roncadin, C., Szatmari, P., & Zwaigenbaum, L. (in press). Stability and change in autism spectrum disorder diagnosis from age 3 to middle childhood in a high-risk sibling cohort. *Autism*.

Bryson, S.E., Zwaigenbaum, L., McDermott, C., Rombough, V., & Brian, J. (2007). The Autism Observation Scale for Infants: Scale development and reliability

data. *Journal of Autism and Developmental Disorders, 38*, 731–738.

Chawarska, K., Shic, F., Macari, S., Campbell, D., Brian, J.,… Klin, A. (2014). 18-month predictors of later outcomes in younger siblings of children with autism spectrum disorder: A BSRC study. *Journal of the American Academy of Child and Adolescent Psychiatry, 53*, 1317-1327. doi: 0.1016/j.jaac.2014.09.015

Constantino, J. N. (2011). The quantitative nature of autistic social impairment. *Pediatric Research, 69*(5), 55R-62R. doi: 10.1203/PDR.0b013e318212ec6e

Constantino, J. N., Lajonchere, C., Lutz, M., Gray, T., Abbacchi, A., McKenna, K., . . . Todd, R. D. (2006). Autistic social impairment in the siblings of children with pervasive developmental disorders. *American Journal of Psychiatry, 163*(2), 294-296. doi: 10.1176/appi.ajp.163.2.294

Folstein, S., & Rutter, M. (1977). Infantile autism: a genetic study of 21 twin pairs. *Journal of Child Psychology & Psychiatry, 18*, 297-321

Fombonne, E., Bolton, P., Prior, J., Jordan, H., & Rutter, M. (1997). A family study of autism: Cognitive patterns and levels in parents and siblings. *Journal of Child Psychology and Psychiatry, 38*(6), 667-683. doi: 10.1111/j.1469-7610.1997.tb01694.x

Georgiades, S., Szatmari, P., Zwaigenbaum, L., Bryson, S., Brian, J., … & Garon, N. (2012). A prospective study of autistic-like traits in unaffected siblings of probands with autism spectrum disorder. *JAMA Psychiatry, 70(1),* 42-48.

Gotham, K., Pickles, A., & Lord, C. (2009). Standardizing ADOS scores for a measure of

severity in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*(5), 693-705. doi: 10.1007/s10803-008-0674-3

Landa, R.J., Gross, A. L., Stuart, E. A., & Bauman, M. (2012). Latent class analysis of early developmental trajectory in baby siblings of children with autism. *Journal of Child Psychology and Psychiatry, 53*(9), 986-996. doi: 10.1111/j.1469-7610.2012.02558.x

Lindgren, K. A., Folstein, S. E., Tomblin, J. B., & Tager-Flusberg, H. (2009). Language and reading abilities of children with autism spectrum disorders and specific language impairment and their first-degree relatives. *Autism Research, 2*(1), 22-38. doi: 10.1002/aur.63

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., 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*(3), 205-223. doi: 10.1023/a:1005592401947

Lord, C., Rutter, M., & Lecouteur, A. (1994). Autism Diagnostic Interview-Revised - a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*(5), 659-685. doi: 10.1007/bf02172145

Messinger, D., Young, G. S., Ozonoff, S., Dobkins, K., Carter, A., Zwaigenbaum, L., . . . Sigman, M. (2013). Beyond autism: A Baby Siblings Research Consortium Study of high-risk children at three years of age. *Journal of the American Academy of Child and Adolescent Psychiatry, 52*(3), 300-308. doi: 10.1016/j.jaac.2012.12.011

Messinger, D. S., Young, G. S., Webb, S. J., Ozonoff, S., Bryson, S. E., Carter, A., . . .

Zwaigenbaum, L. (2015). Early sex differences are not autism-specific: A Baby Siblings Research Consortium (BSRC) study. *Molecular Autism, 6*. doi: 10.1186/s13229-015-0027-y

Miller, M., Young, G. S., Hutman, T., Johnson, S., Schwichtenberg, A. J., & Ozonoff, S. (2015). Early pragmatic language difficulties in siblings of children with autism: implications for DSM-5 social communication disorder? *Journal of Child Psychology and Psychiatry, 56*(7), 774-781. doi: 10.1111/jcpp.12342

Mullen, E. M. (1995). *Mullen Scales of Early Learning* (AGS ed.). Circle Pines, MN: American Guidance Service Inc.

Ozonoff, S., Young, G. S., Belding, A., Hill, M., Hill, A., Hutman, T., . . . Iosif, A. (2014). The broader autism phenotype in infancy: When does it emerge? *Journal of the American Academy of Child and Adolescent Psychiatry, 53*(4), 398-407. doi: 10.1016/j.jaac.2013.12.020

Ozonoff, S., Young, G. S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., . . . Stone, W. L. (2011). Recurrence risk for autism spectrum disorders: A Baby Siblings Research Consortium study. *Pediatrics, 128*(3), e488-e495. doi: 10.1542/peds.2010-2825

Ozonoff, S., Young, G. S., Landa, R. J., Brian, J., Bryson, S., Charman, T., . . . Iosif, A. (2015). Diagnostic stability in young children at risk for autism spectrum disorder: a Baby Siblings Research Consortium study. *Journal of Child Psychology and Psychiatry, 56*(9), 988-998. doi: 10.1111/jcpp.12421

Pickles, A., Starr, E., Kazak, S., Bolton, P., Papanikolaou, K., Bailey, A., . . . Rutter, M. (2000). Variable expression of the autism broader phenotype: Findings from extended

pedigrees. *Journal of Child Psychology and Psychiatry, 41*(4), 491-502. doi: 10.1017/s0021963099005557

Pilowsky, T., Yirmiya, N., Shalev, R. S., & Gross-Tsur, V. (2003). Language abilities of siblings of children with autism. *Journal of Child Psychology and Psychiatry, 44*(6), 914-925. doi: 10.1111/1469-7610.00175

Piven, J., & Palmer, P. (1999). Psychiatric disorder and the broad autism phenotype: Evidence from a family study of multiple-incidence autism families. *American Journal of Psychiatry, 156*(4), 557-563.

Piven, J., Palmer, P., Landa, R., Santangelo, S., Jacobi, D., & Childress, D. (1997). Personality and language characteristics in parents from multiple-incidence autism families. *American Journal of Medical Genetics, 74*(4), 398-411. doi: 10.1002/(sici)1096-8628(19970725)74:4<398::aid-ajmg11>3.0.co;2-d

Pruett, J. R. (2014). BAP: Not-quite-autism in infants. *Journal of the American Academy of Child and Adolescent Psychiatry, 53*(4), 392-394. doi: 10.1016/j.jaac.2014.01.011

Robinson, E. B., Koenen, K. C., McCormick, M. C., Munir, K., Hallett, V., Happe, F., . . . Ronald, A.. (2011). Evidence that autistic traits show the same etiology in the general population and at the quantitative extremes (5%, 2.5%, and 1%). *Archives of General Psychiatry, 68*(11), 1113-1121.

Sparrow, S., Cicchetti, D. V., Balla, D.A. (1984). *Vineland Adaptive Behavior Scales*. Circle Pines, MN: American. Guidance Service Inc.

Sparrow, S., Cicchetti, D. V., Balla, D.A. (2005). *Vineland Adaptive Behavior Scales*. 2nd ed. Circle Pines, MN: American. Guidance Service Inc.

Sucksmith, E., Roth, I., & Hoekstra, R. A. (2011). Autistic traits below the clinical threshold: re-examining the broader autism phenotype in the 21st century. *Neuropsychological Review, 21(4)*, 360-389. doi: 10.1007/s11065-011-9183-9

Szatmari, P., Jones, M. B., Tuff, L., Bartolucci, G., Fisman, S., Mahoney, W. (1993). Lack of cognitive impairment in first-degree relatives of children with pervasive developmental disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 32(6)*, 1264-1273

Szatmari, P., Chawarska, K., Dawson, G., Georgiades, S., Landa, R.,…Halladay A. (2016). Prospective longitudinal studies of infant siblings of children with autism: Lessons learned and future directions. *Journal of the American Academy of Child & Adolescent Psychiatry,* 55(3),179-187. doi: 10.1016/j.jaac.2015.12.014

World Health Organisation. (1993). *Mental Disorders: A Glossary and Guide to their Classification in Accordance with the 10th Revision of the International Classification of Disease - Research Diagnostic Criteria: ICD-10*. Geneva: Author.

Zwaigenbaum, L., Bauman, M. L., Stone, W. L., Yirmiya, N., Estes, A., Hansen, R. L., . . . Wetherby, A. (2015). Early identification of autism spectrum disorder: Recommendations for practice and research. *Pediatrics, 136*, S10-S40. doi:

10.1542/peds.2014-3667C

Table 1 – Sample characteristics

	Low Risk (N = 473)	**High Risk (N = 859)**
Sex (% Male)	53.07%	50.06%
Non-Caucasian (%)[a]	18.85%	19.84%
Maternal Education (% college or higher)[b]	82.18%***	72.75%
Age first seen (months; mean (SD))	7.00 (3.67)**	7.76 (4.43)
Age at outcome (months; mean (SD))[c]	37.40 (2.32)	37.53 (2.44)

a n = 733; b n = 1,194; c n = 1,271; ** $p < .01$; *** $p < .001$

Table 2 – Mullen-defined outcomes (Developmental Delay and Language Delay) by risk group and sex

	Low Risk non-ASD			High Risk non-ASD		
	Total (N=473)	**Females (N=222)**	**Males (N=251)**	**Total (N=859)**	**Females (N=429)**	**Males (N=430)**
No Delay	433 (91.54%)	206 (92.97%)	227 (90.44%)	709 (82.54%)	386 (89.98%)	323 (75.12%)
Developmental Delay (ELC < 85)	16 (3.38%)	4 (1.80%)	12 (4.78%)	91 (10.59%)	19 (4.43%)	72 (16.74%)
Language Delay (T-score < 40)	24 (5.07%)	12 (5.41%)	12 (4.78%)	59 (6.87%)	24 (5.59%)	35 (8.14%)

Table 3 – Elevated ADOS cores (Total CSS 3-and-above) by risk group and sex

	Low Risk non-ASD			High Risk non-ASD		
	Total (N=473)	**Females (N=222)**	**Males (N=251)**	**Total (N=859)**	**Females (N=429)**	**Males (N=430)**
ADOS CSS < 3	393	191	202	609	321	288
	(83.09%)	(86.04%)	(80.48%)	(70.90%)	(74.83%)	(66.96%)
ADOS CSS >=3	80	31	49	250	108	142
	(16.91%)	(13.96%)	(19.52%)	(29.10%)	(25.17%)	(33.02%)

Table 4 – ADI-R 'mild impairment' groups by risk group and sex

	Low Risk non-ASD			**High Risk non-ASD**		
	Total (N=188)	**Females (N=89)**	**Males (N=99)**	**Total (N=412)**	**Females (N=215)**	**Males (N=197)**
ADI Social	3 (1.60%)	1 (1.12%)	2 (2.02%)	26 (6.31%)	8 (3.72%)	18 (9.14%)
ADI Communication	3 (1.60%)	1 (1.12%)	2 (2.02%)	33 (8.01%)	10 (4.67%)	23 (11.62%)
ADI RRB	23 (12.30%)	6 (6.74%)	17 (17.35%)	102 (24.82%)	49 (22.90%)	53 (26.90%)

RRB – Rigid and Repetitive Behavior domain

Table 5 – Vineland 'delayed' groups (standard score < 85) by risk group and sex

	Low Risk non-ASD			High Risk non-ASD		
	Total (N=346)[a]	**Females (N=176)**	**Males (N=170)**	**Total (N=549)**	**Females (N=267)**	**Males (N=282)**
Communication	3 (0.98%)	2 (1.29%)	1 (0.67%)	43 (8.79%)	16 (6.84%)	27 (10.59%)
Daily Living Skills	28 (9.52%)	11 (7.38%)	17 (11.72%)	96 (20.08%)	33 (14.35%)	63 (25.40%)
Socialization	19 (5.49%)	8 (4.55%)	11 (6.47%)	79 (14.39%)	28 (10.49%)	51 (18.09%)
Motor	22 (7.48%)	12 (8.05%)	10 (6.90%)	93 (19.46%)	35 (15.15%)	58 (23.48%)
ABC	8 (2.72%)	3 (2.01%)	5 (3.45%)	76 (16.03%)	27 (11.74%)	49 (20.08%)

a Sample size varies across the Vineland subscales from n = 294 for ABC to n = 346 for Socialization domain for the LR group and from n = 474 for ABC to n = 549 for Socialization domain for the HR group
ABC – Adaptive Behavior Composite

Supplementary Table 1 – Mullen Non-Verbal and Verbal T-scores (mean (SD)) scores by risk group and sex

	Low Risk			**High Risk**		
	Total (N=472)	**Females (N=222)**	**Males (N=250)**	**Total (N=858)**	**Females (N=429)**	**Males (N=429)**
Mullen NV T-score	58.56 (10.28)	60.88 (9.53)	56.51 (10.49)	55.08 (11.09)	57.84 (10.21)	52.32 (11.23)
Mullen V T-score	56.35 (8.24)	57.93 (8.46)	54.96 (7.79)	52.37 (8.73)	54.12 (8.30)	50.61 (8.80)

Supplementary Table2 – ADOS CSS (mean (SD)) by risk group and sex

	Low Risk			**High Risk**		
	Total (N=473)	**Females (N=222)**	**Males (N=251)**	**Total (N=859)**	**Females (N=429)**	**Males (N=430)**
ADOS-CSS	1.79 (1.38)	1.69 (1.15)	1.89 (1.55)	2.28 (1.72)	2.12 (1.68)	2.44 (1.75)

Supplementary Table 3– ADI-R mean (SD) domain scores by risk group and sex

	Low Risk non-ASD			**High Risk non-ASD**		
	Total (N=188)	**Females (N=89)**	**Males (N=99)**	**Total (N=412)**	**Females (N=215)**	**Males (N=197)**
ADI Social	1.94 (2.03)	1.49 (1.78)	2.33 (2.17)	2.87 (2.98)	2.64 (2.84)	3.12 (3.12)
ADI Communication	1.83 (1.91)	1.30 (1.52)	2.29 (2.10)	2.82 (3.13)	2.31 (2.90)	3.38 (3.26)
ADI RRB	0.65 (1.08)	0.50 (0.80)	0.80 (1.27)	0.98 (1.59)	0.84 (1.37)	1.14 (1.79)

RRB – Rigid and Repetitive Behavior domain

Supplementary Table 4 – Vineland (mean (SD)) domain scores by risk group and se

	Low Risk non-ASD			High Risk non-ASD		
	Total (N=346)[a]	**Females (N=176)**	**Males (N=170)**	**Total (N=549)**	**Females (N=267)**	**Males (N=282)**
Communication	108.53 (10.84)	110.23 (11.46)	106.77 (9.89)	102.46 (12.90)	104.75 (12.48)	100.35 (12.94)
Daily Living Skills	99.66 (12.12)	101.56 (12.52)	97.71 (11.42)	94.72 (12.68)	97.47 (11.72)	92.16 (13.02)
Socialization	101.32 (11.71)	102.13 (11.74)	100.47 (11.65)	96.77 (12.40)	98.88 (12.18)	94.78 (12.30)
Motor	102.11 (12.59)	103.37 (13.02)	100.81 (12.04)	96.39 (13.49)	98.42 (12.62)	94.48 (14.02)
ABC	103.44 (11.66)	105.28 (12.48)	101.55 (10.45)	96.74 (12.86)	99.50 (12.22)	94.13 (12.86)

a Sample size varies across the Vineland subscales from n = 294 for ABC to n = 346 for Socialization domain for the LR group and from n = 474 for ABC to n = 549 for Socialization domain for the HR group; ABC – Adaptive Behavior Composite