

# Development and Initial Validation of a Parent Report Measure of the Behavioral Development of Infants at Risk for Autism Spectrum Disorders

Maurice A. Feldman · Rebecca A. Ward · Danielle Savona ·  
Kaleigh Regehr · Kevin Parker · Melissa Hudson ·  
Henderika Penning · Jeanette J. A. Holden

Published online: 9 March 2011  
© Springer Science+Business Media, LLC 2011

**Abstract** We developed and evaluated a new parent report instrument—Parent Observation of Early Markers Scale (POEMS)—to monitor the behavioral development of infants at risk for autism spectrum disorder (ASD) because they have older affected siblings. Parents of 108 at-risk infants (74 males, 34 females) completed the POEMS from child age 1–24 months. The POEMS had acceptable psychometric properties and promising predictive validity.

Most concerning items were social and communication deficits, and intolerance to waiting. Results provide preliminary evidence that prospective parent report measures can help to detect early ASD symptoms in infants at biological risk. We invite researchers to join us in multi-center studies of the POEMS.

**Keywords** Early identification · At-risk infants

---

M. A. Feldman (✉) · R. A. Ward · D. Savona · K. Regehr  
Centre for Applied Disability Studies, Brock University,  
500 Glenridge Ave., St., Catharines, ON L2S 3A1, Canada  
e-mail: mfeldman@brocku.ca

R. A. Ward  
e-mail: bward@brocku.ca

D. Savona  
e-mail: danielle.savona@gmail.com

K. Regehr  
e-mail: kaleighdawn81@gmail.com

K. Parker  
Department of Psychology, Queen's University, Kingston,  
ON K7L 3N6, Canada  
e-mail: Kevin.parker@queensu.ca

M. Hudson · J. J. A. Holden  
Autism Research Program, Departments of Psychiatry  
and Physiology, Queen's University, 191 Portsmouth  
Ave., Kingston, ON K7M 8A6, Canada  
e-mail: melissa.hudson@queensu.ca

J. J. A. Holden  
e-mail: holdenj@queensu.ca

H. Penning  
Equity Office, Queen's University, Kingston,  
ON K7L 3N6, Canada  
e-mail: heidi.penning@queensu.ca

## Introduction

### Need for Early ASD Screening

Children with possible Autism Spectrum Disorders (ASDs) are not typically diagnosed before 4 years of age (Kleinman et al. 2008). The American Academy of Pediatrics (2006) recommends universal screening for ASDs beginning at 18-months of age, but even earlier diagnosis may be possible (Kleinman et al. 2008; Stone et al. 2008). Earlier identification is crucial to ensure that the right families gain timely access to needed services, particularly to early behavioral interventions that substantially improve functioning in many young children with ASDs (Dawson et al. 2009; Lovaas 1987; Perry et al. 2008; Sallows and Graupner 2005).

### Early ASD Screeners

Below we review several broadband and ASD-specific screeners that have been evaluated in detecting early signs of ASD in the first 2 years of life in the general population or in at-risk children (e.g., children referred for developmental problems; infants with siblings with ASD).

### Broadband Screeners

Broadband screeners are designed to detect a wide range of developmental problems, including communication and social deficits that are key features of ASD. The 10-item Parents' Evaluation of Developmental Status (PEDS: Glascoe 2006) is a brief parent report of concerns of their child's development that can be used to identify at-risk young children (starting at about 18 months). Glascoe et al. (2007) found that 34% of 427, 18–59 month children who were considered at-risk for developmental delay based on the PEDS were identified at risk for ASD based on the Modified Checklist for Autism in Toddlers (M-CHAT: Robins, Fein, Barton, & Green, 2001). The authors identified specific patterns of scores that varied given the child's age that reduced over-referrals and maintained acceptable sensitivity. Pinto-Martin et al. (2008), on the other hand, found that the PEDS had very low sensitivity based on the M-CHAT, although their results have been questioned (Glascoe and Squires 2009). Note that neither study actually diagnosed ASD, but instead compared the PEDS to another screener that may not accurately predict ASD diagnosis (see below).

The Infant–Toddler Checklist (ITC), part of the Communication and Symbolic Behavior Scales Developmental Profile (CSBS DP: Wetherby and Prizant 2002) asks parents 25 questions about possible child communication delays. In a prospective study of a general population sample of 5,385 children less than 24 months of age, the ITC correctly identified 93% of children who developed ASD (Wetherby et al. 2008). However, the ITC did not discriminate ASD from other communication delays unless the child score was less than the tenth percentile on the social composite (Wetherby et al. 2008).

### ASD-Specific Screeners

Several early detection instruments have been developed specifically to find children in the general population who may develop ASD. The Checklist for Autism in Toddlers (CHAT: Baird et al. 2000; Baron-Cohen et al. 2000) combines parent report with health care professional observations to screen for ASD at 18 months of age. A follow-up study of 16,235 screened children when they were 7 years old showed that the CHAT identified only 33 of 94 cases (Baird et al. 2000) giving it questionable sensitivity (Wetherby et al. 2008).

The 23 item parent report M-CHAT initially showed promise (Robins et al. 2001), but subsequent studies have tempered the utility of the M-CHAT as an accurate ASD screener for the general population when used on its own. Kleinman et al. (2008) found that the follow-up telephone interview Robins et al. employed to review failed items

with the parents was necessary to decrease false positives and increase predictive validity with a low risk sample. A follow-up study by Pandey et al. (2008) suggested that the M-CHAT has high sensitivity, but for low risk children under 24-months it had low positive predictive value and did not differentiate children with ASD, language delays or global delays.

The two-stage Early Screening of Autistic Traits Questionnaire (ESAT: Dietz et al. 2006) for infants around 14 months old consists of four prescreen items completed by the infant's physician and a follow-up 14-item screening by a trained psychologist during a home visit. From a population sample of 31,724 Dutch infants between 14 and 15 months of age, 69% of children who were positive on the 4-item prescreen received the 14 item follow-up. From these 255 children, 18 were diagnosed with ASD (by some of the authors). With the large dropout rate, need for a trained professional for the second screen, and its low specificity and predictive power, the value of the ESAT as an ASD-specific screener is questionable.

The Pervasive Developmental Disorders Screening Test-II (PDDST-II: Siegel 2004) is a parent report instrument that has been used to detect ASD in young children. Stage 1 is a general screener for pediatricians. Stage 2 looks at developmental delay, and Stage 3 focuses on ASD. Based on clinical impression, Stage 1 ASD sensitivity was .92 and specificity was .91, but there was no confirmation of ASD diagnosis (Siegel 2004). Screening clinic practitioners have questioned the clinical utility of the PDDST-II (McQuistin and Zieren 2006).

### ASD-Specific Second-Order Screeners Using Infants at Biological Risk

Several researchers have tested early screeners for undiagnosed infants who have siblings with ASD. ASDs are highly inheritable: concordance in monozygotic twins is 60% for Autistic Disorder (AD) and higher for PDD-NOS and the broader phenotype of language delays, social skill deficits and ritualistic and repetitive behaviors (Filipek et al. 1999). Non-twin siblings have a 5–8% risk of AD and 20% risk for the broader phenotype (American Psychiatric Association 2000; Bolton et al. 1994; Wolff 2004). As it is expected to see a higher proportion of infant siblings of affected children to develop an ASD than infants in the general population, a smaller validation sample size may be used than for a population screener. Importantly, these are the very infants who should be screened given their known biological risk.

Prospective research studies have found pre-diagnostic symptoms within the first 2 years of life in infants with siblings with ASD (Bryson et al. 2007; Garon et al. 2008; Iverson and Wozniak 2007; Landa and Garrett-Mayer

2006; Landa et al. 2007; Nadig et al. 2007; Ozonoff et al. 2008a, b; 2010; Yoder et al. 2009; Zwaigenbaum et al. 2005). Across these studies, early markers included communication, social and behavior problems typically seen in older children with ASDs. Early signs detected include deficits in eye contact and tracking, responding to name, imitation, language, social development, joint attention, gestures, play, visual examination of objects, emotional regulation and positive and negative affect.

Stone and colleagues tested the appropriateness of using the Screening Tool for Autism in Two-Year-Olds (STAT) with 71 at-risk infants (59 infant sibs and 12 for whom there were concerns about ASD) between 12 and 23 months old (Stone et al. 2008). The STAT is a 12-item, 20 min interactive test conducted by a trained professional and measures “play (two items), requesting (four items), directing attention (four items), and motor imitation (four items)” (Stone et al. 2008, p. 562). ASD diagnoses were made after 24 months of age by psychologists (it was not stated if they were blind to earlier STAT and other test scores) using the ADOS (Lord et al. 2000) and the DSM-IV-TR (American Psychiatric Association 2000) criteria. With adjustments to the cut-off scores, Stone et al. (2008) found that the STAT had reasonable sensitivity (.93) and specificity (.83) for at-risk infants  $\geq 14$  months of age. A higher proportion of false positives was obtained in the 12–13 month group.

The Autism Observation Scale for Infants (AOSI: Bryson et al. 2008) was designed to track early signs of ASD in 6–18 month old infants with older affected siblings. It uses a set of structured play activities to elicit 18 behaviors related to “Visual Tracking, Disengagement of Attention, Orientation to Name, Reciprocal Social Smiling, Differential Response to Facial Emotion, Social Anticipation and Imitation” (Bryson et al. 2008, p. 733). The researchers administered and coded the AOSI. Total score inter-rater reliability (using intra-class correlations) was acceptable ( $>.90$ ) at 12 ( $n = 34$ ) and 18 months of age ( $n = 26$ ), but below .75 at 6 months ( $n = 32$ ); 2-week total score test-retest at 12 months was .61 ( $n = 20$ ) (Bryson et al. 2008). In a prospective study, the AOSI showed potential to distinguish high from low risk infants as early as 12-months of age (Zwaigenbaum et al. 2005), but more studies are needed.

In summary, few broadband and ASD specific instruments show promise detecting ASD in children under 24 months. Two brief broadband parent-report screeners (PEDS, ITC) may accurately screen for likely ASD, but more work is needed to confirm initial findings. Two ASD specific instruments, the STAT and AOSI, both require direct testing of the child by trained professionals, have limited research and neither has detected ASD in the first year of life.

## Need for Parent-Report Instrument to Monitor Many Early Signs of ASD

In order to detect very early signs, possibly before 12 months of age, it would be helpful to more repeatedly monitor, under natural conditions, a range of infant behaviors that may be related to incipient ASD. Brian et al. (2008) highlight the importance of appraising specific behaviors in infants beyond social-communicative ones that often are seen in older children with ASD. Parents can play an important role in assessing multiple infant behaviors in a cost-effective manner. Developmental diagnoses correspond with parent reported concerns of child developmental problems (Glascoe 2000; Glascoe et al. 1997). Siegel et al. (1986) found high reliability between parent observations in the home and a diagnostic play session on specific autistic behaviors in children with ASDs.

Thus, there is a need for a parent-report instrument that examines a wide range of possible early signs of ASD, including core features as well as other behaviors seen in young children with ASD. To address this gap, we designed the Parent Observation of Early Markers Scale (POEMS) as a checklist that parents can use to *prospectively* monitor 61 specific behaviors that may be possible early symptoms and associated behaviors of an ASD in their 1–24 month old infants. If the prospective use of the POEMS is valid, and differentiates at-risk infants who were, and were not, subsequently diagnosed with an ASD, then autistic symptoms may be identifiable by parents earlier, which may lead to potential preventative interventions (Dawson 2008).

## Method

### Participants

To participate in this study, families had at least one biological child with an independent diagnosis of ASD—Autistic Disorder (AD), PDD-NOS, Asperger syndrome (AS) or High Functioning Autism (HFA)—and a younger biological sibling between 1 and 24 months of age. We recruited families through a website ([www.AutismResearch.ca](http://www.AutismResearch.ca)) from across North America. To further promote the study, we gave presentations to family organizations (e.g., Autism Ontario, Autism Society Canada) and distributed pamphlets at said organizations, physician offices and developmental clinics.

We recruited 239 families who initially expressed interest in participating by registering online. On further inquiry, a significant number of families (exact numbers not available) were not eligible because they did not have either a  $<24$  month biological sibling of a child with ASD

or an older child with confirmed ASD. Other families decided not to participate because they felt they did not have the time to make a multi-year commitment. The remaining 118 families consented to participate in this study, but 17 of these families declined for unknown reasons after signing the consent form. The final tally of actual participants was 108 eligible infants from 103 families. There were one set of identical twins and one set of non-identical twins. Three additional infants from the same families were born while the older infant was still in the study. Table 1 provides descriptive information on the participating children and families. None of the infant siblings had known biological, birth or medical conditions associated with potential developmental problems (e.g., Down syndrome, low birth weight, epilepsy). None of the older siblings had Fragile X or other syndromes related to ASD.

## Measures

### *Experimental Measure: Parent Observation of Early Markers Scale (POEMS)*

We designed the POEMS based on a review of autism assessment instruments, including the ADI-R (Lord et al. 1994), ADOS-Generic (Lord et al. 2000) and the Childhood Autism Rating Scale (CARS, Schopler et al. 1988). We created 61 items covering problem areas that would be appropriate for infants and toddlers, aged 1–24 months. Some of the items were related to the core deficits of ASDs—problems in social and communicative development; restricted interests; and ritualistic, repetitive non-functional behaviors. Other POEMS items dealt with behavioral, emotional and other problems commonly seen in young children with ASD—e.g., intolerance to transitions, waiting, new foods, loud noises; and problems with

sleep, toileting, emotional regulation, mood, attention, visual tracking; motor agility and movement. The items were grouped by topic—e.g., feeding, response to parent, response to environment, communication; we explicitly did not present subscales.

When we used the POEMS with the parents, we referred to it with the more generic name, “Parent Observation Checklist” because we did not want the parents to only observe what they believed to be early markers of ASDs (i.e., similar infant behaviors to those they saw in their older, now diagnosed child). Parents scored each item based on the child’s behavior in the preceding week. The scoring system was modeled after the CARS: each item was rated on a four-point scale, where 1 is no problem (typical development—described), 2 is mild problem (child’s behavior is not completely typical for his/her age), 3 is moderate problem (child behavior is concerning) and 4 is severe problem (described); 1/2 scores were allowed. Descriptions were provided for the 1 and 4 anchors.

We provide two examples of POEMS items: For the item Shifts Attention To Person a score of 1 signifies *shifts attention from object/toy to person’s face easily*, whereas a score of 4 reads *has great difficulty shifting attention from an object/toy to a face*. For the item Waiting, a score of 1 is described as *tolerates brief wait before needs can be met; remains calm but expectant while waiting*; a score of 4 states *cannot tolerate any wait to have needs met; easily frustrated; quick to cry or tantrum if needs are not met immediately*. We strongly encouraged the parents to base their scores on actual observations of their children and test any item with the child if they were unsure of the child’s response or the child had not had the opportunity to exhibit the behavior. Parents were told to give a score of “not applicable” (NA) to any item that was too developmentally advanced given the chronological age of the child (this decision was discussed with the interviewer). For analysis,

**Table 1** Characteristics of children and families

Variable	
Mean age of infant at start of study (months) ( <i>SD</i> )	8.06 (5.00) (range: 1.03–22.23 months)
Percentage male infants	69%
mean age of affected siblings at start of study (months) ( <i>SD</i> )	68.45 (32.40)
Percentage male affected siblings	85%
Mean age of mothers at start of study (years) ( <i>SD</i> )	35.21 (4.02)
Percentage mothers with at least a college/university degree	82%
Percentage mothers employed other than or in addition to homemaker	59%
Mean age of fathers at start of study (years) ( <i>SD</i> )	37.90 (4.95)
Percentage fathers with at least a college/university degree	70%
Percentage fathers employed other than or in addition to homemaker	96%
Range of annual family income (lowest—highest income range)	\$5 K to > \$95 K
Percentage of two-parent families	97%

we converted NAs to 1's (no problem) so that for every child the minimum POEMS score was 61 and maximum score was 244. The primary caregiver of the infant (usually the biological mother) completed the POEMS via mail, online or by phone interview, with at least 1 month between administrations. The cumulative number of POEMS was: 247 up to child age 9 months, 396 up to 12 months, 671 up to 18 months and 902 up to 24 months (mean of 8.35 POEMS per child up to 24 months).

*The Ages and Stages Questionnaire (ASQ: Bricker and Squires 1999)*

Filipek et al. (1999) recommended the use of the ASQ 2nd edition in the American Academy of Neurology and the Child Neurology Society (AANCNS) Consensus Panel report. The ASQ is a validated parental interview that monitors infant development. Parents completed the ASQ by mail, and results were discussed with parents through a follow-up telephone interview. The ASQ has six subscales: I. Communication, II. Gross motor, III. Fine motor, IV. Problem solving, V. Personal-social, and VI. Overall. In this study, we used the 12-month ASQ to test the POEMS' discriminant validity. A subsequent study will report the longitudinal ASQ findings.

*Autism Diagnostic Instrument-Revised (ADI-R: Lord et al. 1994)*

The ADI-R is an extensive informant interview measure that can be used with verbal and nonverbal children. We used the available algorithm for scoring. We administered the ADI-R over the telephone in accordance with the procedures described and validated in Ward-King et al. (in press).

**Procedure**

*Interview*

We interviewed parents of high-risk infants over the phone on several occasions throughout the study, depending on the availability of the parent. The mean number of interviews per family was 8.05 ( $SD = 4.13$ ). The three interviewers all had considerable experience working with families who had children with ASDs. One interviewer was a psychologist with 20 years experience in assessment and intervention of ASDs; the second interviewer had a B.A. in psychology and had been a therapist in an intensive behavioral intervention program; the third interviewer had a B.Sc. in Biology and worked for several years in psychiatric neurogenetics research.

Each interview started with a general update about the child's health, significant family events and changes in the family situation. If the parents had not recently completed the POEMS and ASQ, they were administered by phone in no particular order. To keep the phone interviews as similar as possible to when the parents completed hard copy or email versions of the POEMS, the parents had copies of the instruments to read during the interview and told the interviewer their scores to each item. The interviewer transcribed the parent's scores onto the interviewer's form. The interviewer then answered any queries that parents had about the questionnaires. If the parents had already completed the instruments, the interview lasted about 30 min. If some or all of the instruments needed to be completed over the phone, the interview could last up to 90 min. Throughout the study, we used the consensual AANCNS "level one: routine developmental surveillance" criteria that included language, pointing and other gestural communication deficits starting at 12 months of age and low scores on [autism screeners] at 18 months (Filipek et al. 1999, p. 449) to determine when we should advise the parent to follow up with the child's pediatrician.

*Confirmation of Sibling ASD*

Parents sent in copies of diagnostic reports in their possession for the older child with an ASD. In addition, during the course of the study, a research reliable assessor administered the ADI-R (Lord et al. 1994) by phone interview (Ward-King et al. in press) to confirm 86% of the older siblings' ASD diagnosis. There were no older siblings who failed to meet an ASD diagnosis on the ADI-R; the 14% not tested was due to the unavailability of the parent for the interview.

**Results**

**POEMS Internal Consistency**

We calculated internal consistency for those infants for whom we have POEMS scores at 3, 6, 9, 12, 18 and 24 months of age. Internal consistency was acceptable at each age—Cronbach's *alphas* were .83 at 3 months ( $n = 36$ ), .94 at 6 months ( $n = 41$ ), .90 at 9 months ( $n = 38$ ), .93 at 12 months ( $n = 57$ ), .96 at 18 months ( $n = 51$ ) and .97 at 24 months ( $n = 43$ ). Spearman Brown coefficients were .86 at 3 months, .95 at 6 months, .94 at 9 months, .96 at 12 months, .99 at 18 months and .98 at 24 months.

### POEMS Test–Retest Reliability

Test–retest reliability calculated at one-month intervals between 2 and 23 months showed that the POEMS was stable over the time periods examined. Test–retest reliabilities for the POEMS total score were: 2–3 months old,  $r = .93$  ( $n = 13$ ); 3–4 months,  $r = .91$  ( $n = 14$ ); 4–5 months,  $r = .48$  ( $n = 15$ ); 5–6 months,  $r = .92$  ( $n = 20$ ); 7–8 months,  $r = .98$  ( $n = 20$ ); 9–10 months,  $r = .80$  ( $n = 25$ ); 11–12 months,  $r = .82$  ( $n = 35$ ); 14–15 months,  $r = .83$  ( $n = 26$ ); 16 to 17 months,  $r = .94$  ( $n = 20$ ); 18 to 19 months,  $r = .97$  ( $n = 23$ ); and 22–23 months,  $r = .98$  ( $n = 21$ ).

### POEMS Construct Validity

To examine the POEMS construct validity, we looked at both convergent and discriminant validity between the POEMS and ASQ when the study infants were 12 months old.

#### *Convergent Validity*

We used the ASQ social and communication domains for convergent validity with the POEMS because social and communication deficits are core features of autism. The correlation between the 12 month POEMS total score and ASQ social and communication domain scores were,  $r = -.41$  and  $-.45$ ,  $p$ 's  $< .01$  ( $n$ 's = 43), respectively.

#### *Divergent Validity*

It was difficult finding areas of development not associated with ASDs. Given the instruments we used, we picked the ASQ gross and fine motor scales. Although problems in sensory-motor development have been implicated in persons with ASDs (Baranek 1999), we did not expect that the correlations between the POEMS total score and the ASQ motor domain scores would be as strong as with the ASQ social and communication domains (Ozonoff et al. 2008b). The 12 month POEMS total score was not correlated with the 12 month ASQ gross motor domain,  $r = .09$  ( $n = 43$ ), but the 12 month POEMS total score and ASQ fine motor domain were significantly correlated,  $r = -.32$ ,  $p < .05$  ( $n = 43$ ). A test of differences between correlations showed that the correlation between the POEMS total score and the ASQ gross motor domain was significantly different from the correlation between the POEMS total score and the ASQ social domain,  $t = 2.75$ ,  $df = 42$ ,  $p < .01$ . Likewise, the correlation between the POEMS total score and the ASQ gross motor domain was significantly different from the correlation between the POEMS total score and the ASQ communication domain,  $t = 2.77$ ,  $df = 42$ ,

$p < .01$ ). Thus, relationships between the POEMS and ASQ were stronger with the core features of ASD (social and communication problems) than with gross motor problems. There were no significant differences between the POEMS/ASQ fine motor correlation and the POEMS/ASQ communication and social scales correlations.

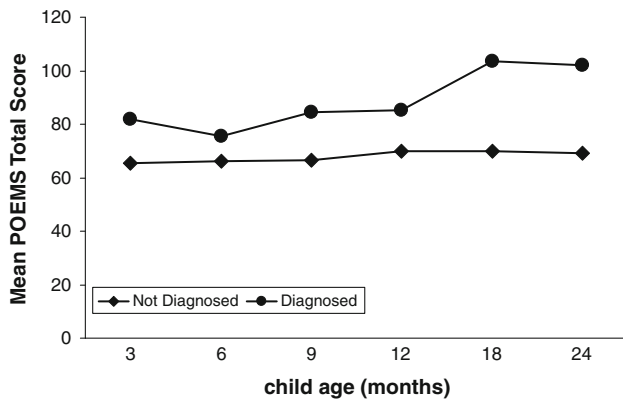
### POEMS Predictive Validity

The *raison-d'être* for developing early detection instruments is that they might be able to discriminate at a young age which children will and will not eventually receive an ASD diagnosis and who may need early intervention. To begin this process, we followed the study children until they were 3 years old. By that age, nine children (6 males, 3 females) had received independent community ASD diagnoses (seven with AD and two with PDD-NOS). We were able to administer the ADI-R by a research-reliable tester to 69% of the infants when they were 3-years old. Unavailability of the parent precluded administration of the ADI-R for the remaining sample. None of the undiagnosed infants scored positive on the ADI-R. We confirmed the community ASD diagnosis with the ADI-R for three infants. Another infant's community diagnosis of PDD-NOS was not confirmed by our administration of the ADI-R. The remaining five infants did not receive an ADI-R from us.

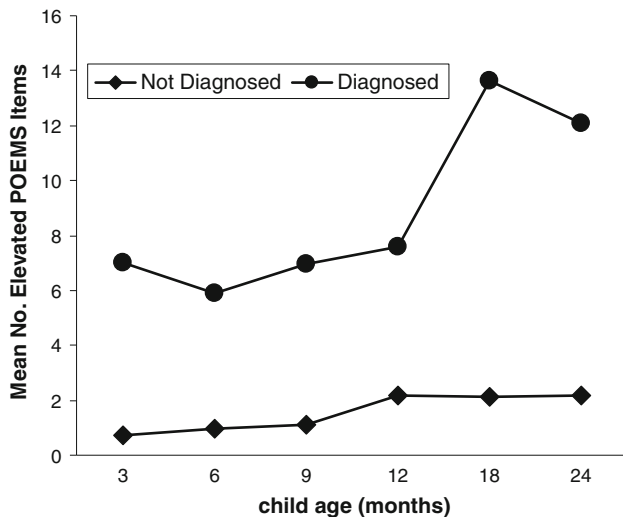
#### *Between-Group Comparisons*

Figures 1 and 2 show that the subsequently diagnosed children had higher mean total POEMS scores and elevated POEMS items (score  $\geq 3$ ) between 3 and 24 months than the remaining 99 children who were not diagnosed by 36 months. While the total scores remained steady for the undiagnosed group (overall mean = 65.18,  $SD = 3.74$ ), they increased over age in the diagnosed group (overall mean = 92.22,  $SD = 28.82$ ). The mean number of elevated items across ages was .78 ( $SD = .81$ ) in the undiagnosed group and 8.86 ( $SD = 10.12$ ) in the diagnosed group.

We performed a mixed model (diagnosis between, ages within) ANOVA, using Type III sums of squares for children with POEMS total score data from 9, 12, 18 and 24 months. We started this analysis at 9 months because there were only two and four to-be-diagnosed children who had POEMS less than 3 and 6 months of age, respectively. There were seven to-be-diagnosed children and 63 children not diagnosed by age 36 months included in this ANOVA. POEMS scores were significantly higher in the diagnosed group,  $F(1, 68) = 52.23$ ,  $p < .001$ ,  $\eta^2 = .66$ ) and for older children  $F(3, 204) = 6.98$ ,  $p < .001$ ,  $\eta^2 = .30$ ). The Diagnosis by Age interaction was significant,  $F(3,$



**Fig. 1** Mean total POEMS scores at 3, 6, 9, 12, 18 and 24 months of age for the nine infants diagnosed and the 99 infants not diagnosed with an ASD by 36 months. POEMS minimum total score = 61; maximum total score = 244. *N*'s—3 months: 2 diagnosed, 22 not diagnosed; 6 months: 4 diagnosed, 50 not diagnosed; 9 months: 7 diagnosed, 63 not diagnosed; 12 months: 7 diagnosed, 81 not diagnosed; 18 months: 9 diagnosed, 90 not diagnosed; 24 months: 9 diagnosed, 99 not diagnosed



**Fig. 2** Mean number of POEMS elevated items at 3, 6, 9, 12, 18 and 24 months of age for the nine infants diagnosed and the 99 infants not diagnosed with an ASD by 36 months. POEMS elevated scores = 3, 3.5, 4 on a scale of 1–4 (1 = no problem, 4 = severe problem). *N*'s—3 months: 2 diagnosed, 22 not diagnosed; 6 months: 4 diagnosed, 50 not diagnosed; 9 months: 7 diagnosed, 63 not diagnosed; 12 months: 7 diagnosed, 81 not diagnosed; 18 months: 9 diagnosed, 90 not diagnosed; 24 months: 9 diagnosed, 99 not diagnosed

204) = 8.06,  $p < .001$ ,  $eta = .33$ ) reflecting the separation in mean POEMS total scores between the diagnosed and undiagnosed groups as the children aged (see Fig. 1). The Age, and Diagnosis by Age effect sizes are considered medium in magnitude, and the between-groups Diagnosis effect size is considered very large ( $eta$  can be understood as roughly equivalent to a  $r$  statistic). We found similar results on the mean number of elevated POEMS items. The

diagnosed group had significantly more elevated items than the undiagnosed group,  $F(1, 68) = 69.38$ ,  $p < .001$ ,  $eta = .74$ ), as did older children,  $F(3, 204) = 11.01$ ,  $p < .001$ ,  $eta = .37$ ). The significant Diagnosis by Age interaction,  $F(3, 204) = 4.57$ ,  $p < .01$ ,  $eta = .25$ ) reflected the divergence in mean POEMS elevated scores between the groups across age (see Fig. 2). Post-hoc comparisons showed that the diagnosed group had significantly higher mean total scores and number of elevated items than the undiagnosed group at 9, 12, 18 and 24 months respectively,  $p < .05$ .

We examined the most frequently reported elevated POEMS items that differentiated the to-be-diagnosed from the non-diagnosed group up to 9, 12, 18 and 24 months of age. At 9 months of age, the most frequently reported elevated items in the diagnosed group were: interest in faces (45% of the diagnosed group), shifts attention to person (40%), mood (35%), response to name (35%) and waiting (35%). At 12 months they were: interest in faces (33%), waiting (33%), shifts attention to person (30%) and imitates sounds or words (27%). At 18 months they were: waiting (43%), imitates sounds or words (36%), and coordinates point and gaze (30%). At 24 months they were: imitates sounds or words (41%), waiting (41%), imitates actions (30%), coordinates point and gaze (29%), points in response to questions (29%) and communicates with words (29%). As can be seen, concerns followed a developmental progression; for the most part, the diagnosed group showed unique social and communication markers that may be incipient signs of ASD. On the other hand, the undiagnosed infants did not show as many elevations for any POEMS items at any age. The median percentage of undiagnosed infants who were elevated on the same items as the diagnosed group (reported above) across all ages was zero; the highest percentage of undiagnosed infants who had elevated scores on the above items was 16% for waiting at 12 months.

*Sensitivity, Specificity and Positive Predictive Value*

We selected a relatively low POEMS total score cut-off score of 70 because it was about midway between the means in the full sample at each age examined. We found overall sensitivity to be .74. Sensitivity at 3, 6, 9, 12, 18, and 24 months of age was .50, .25, .57, .71, .89 and 1.00, respectively. Overall specificity was .73. Specificity at 3, 6, 9, 12, 18, and 24 months of age was .87, .82, .84, .68, .65 and .70, respectively. Overall positive predictive value (PPV) was .21. PPV at 3, 6, 9, 12, 18, and 24 months of age was .25, .10, .29, .16, .20, .24, respectively. The above calculations were influenced by the disproportionate number of non-diagnosed to subsequently diagnosed children. Using a matched sample of nine undiagnosed children

would have yielded unstable coefficients. The values likely would be higher in larger and more similar sized samples of diagnosed and non-diagnosed children.

## Discussion

In this preliminary study, the POEMS demonstrates acceptable reliability and validity for future research. Importantly, the POEMS shows promising predictive validity, differentiating at-risk infants at 3–24 months of age who were, versus were not, independently diagnosed with an ASD at 36 months. With a conservative cutoff score, the POEMS shows reasonable sensitivity and specificity. These attributes require confirmation with a larger sample size of diagnosed children and a longer follow-up with expert diagnostic observations.

In this study, significant differences between to-be-diagnosed and undiagnosed at-risk infant groups emerged at 9 months of age (our sample sizes of 2 and 4, were likely too small to show any statistical significance at 3 and 6 months, respectively) These results provide further evidence that early signs may emerge as early as 9 months, if not before, in infants with siblings who have ASD (Ozonoff et al. 2010; Zwaigenbaum et al. 2005). Our approach of using frequent, ongoing parental surveillance of numerous possible markers in natural environments may yield earlier detection than previous studies relying on more formal testing (Ozonoff et al. 2010).

POEMS items that appeared most frequently in the subsequently diagnosed infants included social-communication problems seen in other longitudinal and laboratory studies of infant siblings, such as interest in faces, responding to name, shifting attention and imitation (e.g., Landa et al. 2007; Nadig et al. 2007; Zwaigenbaum et al. 2005). Indeed, the fact that interest in faces differentiated the diagnosed and undiagnosed groups as early as 9 months supports Ozonoff et al.'s (2010) contention that this behavior (which they refer to as “gaze to faces”) is “most sensitive to emerging signs of autism” (p. 265). The POEMS also may be picking up behavioral and emotional problems, such as intolerance for waiting and irritable, unhappy mood that, while not core features of ASD, do tend to occur in young children with an ASD (Brian et al. 2008; Bryson et al. 2007; Garon et al. 2008). Note that the most common problems reported by the parents on the POEMS at different ages occurred in fewer than 50% of the diagnosed children. Thus, there may be different phenotypes and developmental pathways to diagnosis, which may be revealed through indepth case studies (Bryson et al. 2007; Ozonoff et al. 2008).

This study adds evidence that parents may be good reporters of their child's development, especially when

done prospectively (Glascoe 2000, 2005). The prospective predictive ability results of the POEMS starting at child age 9 months suggest that problems with parental retrospective reporting seen in the Ozonoff et al. (2010) study may reflect memory issues rather than parental insensitivity to notice early signs in infants who are subsequently diagnosed with an ASD. The relatively low POEMS scores for the undiagnosed at-risk infants—mean total score was 65 (minimum POEMS score is 61)—suggest that parents who already had children with an ASD were not necessarily anxious about their infants also developing ASD and therefore over-reporting elevated POEMS scores.

We note several limitations of this study. First, our sample of nine at-risk infants who were subsequently diagnosed at 3 years of age is relatively small compared to other prospective studies that used ASD diagnostic assessments (Landa and Garrett-Mayer 2006; Landa et al. 2007; Zwaigenbaum et al. 2005). Following the prospective cohort to older ages may yield more children diagnosed with ASD or showing evidence of the broader phenotype. Like any screener of low base rate conditions, the POEMS was prone to false positive identifications and produced a relatively low positive predictive value even with reasonable overall sensitivity and specificity when the entire sample was used. If we do discover more children with ASD diagnoses, these values may increase in that children who were treated as false positives would be changed to true positives. False negatives also may present a problem, as seen in long-term follow-up studies of the CHAT (Baird et al. 2000). Currently, as single point identification needs to be interpreted with caution, a better approach may be to monitor a child over time looking for evidence of increasing POEMS scores. Another limitation is our reliance on independent community diagnosis and we did not confirm all the children's diagnoses with our own administration of the ADI-R (Lord et al. 1994) or the ADOS (Lord et al. 2000). A third limitation is that we did not obtain inter-rater agreements (despite requesting other family members to occasionally complete the POEMS independently).

Future research will attempt to replicate the above POEMS findings using a larger and broader socio-demographic sample through internet recruitment and online data collection. We invite other researchers to join us in a multi-center study. We will conduct POEMS factor and item analyses, investigate the longitudinal development of these children up to at least 5 years of age and compare POEMS scores between at-risk infants and low risk infants—i.e., no family history of ASD. We will examine the utility and properties of the POEMS as a screening device for all families who would like to monitor the development of their infants.



In conclusion, this study found that a new behavioral checklist designed for parents to prospectively monitor the behavioral development of infants who have older siblings with an ASD has acceptable psychometric properties. Parents were able to distinguish infants as early as 9 months of age who subsequently were diagnosed with an ASD from those who were not diagnosed with an ASD by 36 months of age. Although more research is needed, the POEMS shows promise as a simple, low-cost monitoring system that parents can use that may result in earlier detection and intervention of remediable developmental and behavioral problems in infants at-risk for an ASD.

**Acknowledgments** This work was supported by CIHR Interdisciplinary Health Research Team grant (RT-43820) to the Autism Spectrum Disorders Canadian-American Research Consortium (ASD-CARC: [www.autismresearch.ca](http://www.autismresearch.ca)) and funding from Ongwanada (J. J. Holden, principal investigator). D. Savona and K. Regehr were funded trainees in the CIHR/Autism Speaks Autism Spectrum Interdisciplinary REsearch (ASPIRE) Strategic Training in Health Research Program (STIHR) (JJAH, principal investigator, [www.AutismTraining.ca](http://www.AutismTraining.ca)). We extend our sincere appreciation to the families for their enthusiastic support of this study, and the students and research assistants (S. Hannah, J. Doiron, F. Altoumah, R. Leung) who helped with this study. Copies of the POEMS and other information concerning this study can be obtained from Maurice Feldman, Ph.D., C.Psych., BCBA-D, Professor and Director, Centre for Applied Disability Studies, Brock University, 500 Glenridge Ave., St. Catharines, ON L2S 3A1. Email: [mfeldman@brocku.ca](mailto:mfeldman@brocku.ca). Phone (905) 688-5550 x 4894.

## References

- American Academy of Pediatrics, Council on Children with Disabilities, Section on Developmental Behavioral Pediatrics, Bright Futures Steering Committee, Medical Home Initiatives for Children with Special Needs Project Advisory Committee. (2006). Identifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. *Pediatrics*, *118*, 405–420.
- American Psychiatric Association (2000). *Diagnostic and statistics manual of mental disorders* (4th ed., text revision DSM-IV-TR). Washington, DC: Author.
- Baird, G., Charman, T., Baron-Cohen, S., Cox, A., Swettenham, J., Wheelwright, S., et al. (2000). A screening instrument for autism at 18 months of age: a 6-year follow-up study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*, 694–702.
- Baranek, G. (1999). Autism during infancy: A retrospective video analysis of sensory-motor and social behaviors at 9–12 months of age. *Journal of Autism and Developmental Disorders*, *29*, 213–224.
- Baron-Cohen, S., Wheelwright, S., Cox, A., Baird, G., Charman, T., Swettenham, J., et al. (2000). The early identification of autism: The checklist for autism in toddlers (CHAT). *Journal of the Royal Society of Medicine*, *93*, 521–525.
- Bolton, P., Macdonald, H., Pickles, A., Rios, P., Goode, S., Crowson, M., et al. (1994). A case-control family history study of autism. *Journal of Child Psychology and Psychiatry*, *35*, 877–900.
- Brian, J., Bryson, S. E., Garon, N., Roberts, W., Smith, I. M., Szatmari, P., et al. (2008). Clinical assessment of autism in high-risk 18-month-olds. *Autism*, *12*, 433–456.
- Bricker, D., & Squires, J. (1999). *The ages & stages questionnaires*. Baltimore, MD: Paul H. Brookes.
- Bryson, S. E., Zwaigenbaum, L., Brian, J., Roberts, W., Szatmari, P., Rombough, V., et al. (2007). A prospective case series of high-risk infants who developed autism. *Journal of Autism and Developmental Disorders*, *37*, 2–24.
- Bryson, S. E., Zwaigenbaum, L., McDermott, C., Rombough, V., & Brian, J. (2008). The Autism observation scale for infants: Scale development and reliability data. *Journal of Autism and Developmental Disorders*, *38*, 731–738.
- Dawson, G. (2008). Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. *Development and Psychopathology*, *20*, 775–803.
- Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., et al. (2009). Randomized, controlled trial of an intervention for toddlers with autism: the early start Denver model. *Pediatrics*. doi:10.1542/peds.2009-0958.
- Dietz, C., Swinkels, S., Daalen, E., Engeland, H., & Buitelaar, J. (2006). Screening for autistic spectrum disorders in children aged 14–15 months. II: Population screening with the Early Screening of Autistic Traits Questionnaire (ESAT). Design and general findings. *Journal of Autism and Developmental Disorders*, *36*, 713–722.
- Filipek, P. A., Accardo, P. J., Baranek, G. T., Cook, E. H., Dawson, G., Gordon, B., et al. (1999). The screening and diagnosis of autistic spectrum disorders. *Journal of Autism and Developmental Disorders*, *29*, 439–484.
- Garon, N., Bryson, S. E., Zwaigenbaum, L., Smith, I. M., Brian, J., Roberts, W., et al. (2008). Temperament and its relationship to autistic symptoms in a high-risk infant sib cohort. *Journal of Abnormal Child Psychology*, *37*, 59–78.
- Glascoe, F. P. (2000). Evidence-based approach to developmental and behavioural surveillance using parents' concerns. *Child: Care, Health and Development*, *26*, 137–149.
- Glascoe, F. P. (2005). Screening for developmental and behavioral problems. *Mental Retardation and Developmental Disabilities Research Reviews*, *11*, 173–179.
- Glascoe, F. P. (2006). *Parents' evaluations of developmental status: A method for detecting and addressing developmental and behavioral problems in children*. Nashville, TN: Ellsworth & Vandermeer Press.
- Glascoe, F. P., Foster, M., & Wolraich, M. L. (1997). An economic analysis of developmental detection methods. *Pediatrics*, *99*(6), 830–837.
- Glascoe, F. P., Macias, M. M., Wegner, L. M., & Robertshaw, N. S. (2007). Can a broadband developmental-behavioral screening test identify children likely to have autism spectrum disorder? *Clinical Pediatrics*, *46*, 801–805.
- Glascoe, F. P., & Squires, J. K. (2009). To the editor. *Journal of Developmental and Behavioral Pediatrics*, *30*, 174.
- Iverson, J. M., & Wozniak, R. H. (2007). Variation in vocal-motor development in infant siblings of children with autism. *Journal of Autism and Developmental Disorders*, *37*, 158–170.
- Kleinman, J. M., Robins, D. L., Ventola, P. E., Pandey, J., Boorstein, H. C., Esser, E. L., et al. (2008). The modified checklist for autism in toddlers: A follow-up study investigating the early detection of autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *38*, 827–839.
- Landa, R., & Garrett-Mayer, E. (2006). Development in infants with autism spectrum disorders: A prospective study. *Journal of Child Psychology and Psychiatry*, *47*(6), 629–638.
- Landa, R. J., Holman, K. C., & Garrett-Mayer, E. (2007). Social and communication development in toddlers with early and later

- diagnosis of autism spectrum disorders. *Archives of General Psychiatry*, 64(7), 853–864.
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., et al. (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.
- Lord, C., Rutter, M., & Le Couteur, 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, 659–685.
- Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology*, 55, 3–9.
- McQuistin, A., & Zieren, C. (2006). Clinical experiences with the PDDST-II. *Journal of Autism and Developmental Disorders*, 36, 577–578.
- Nadig, A. S., Ozonoff, S., Young, G. S., Rozga, A., Sigman, M., & Rogers, S. J. (2007). A prospective study of response to name in infants at risk for autism. *Archives of Pediatric and Adolescent Medicine*, 161(4), 378–383.
- Ozonoff, S., Heung, K., Byrd, R., Hansen, R., & Hertz-Picciotto, I. (2008a). The onset of autism: Patterns of symptom emergence in the first years of life. *Autism Research*, 1, 320–328.
- Ozonoff, S., Iosif, A.-M., Baguio, F., Cook, I. C., Hill, M. M., Hutman, T., et al. (2010). A prospective study of the emergence of early behavioral signs of autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, 49(3), 258–268. In press article retrieved from <http://www.jaacap.org> on Feb. 19, 2010.
- Ozonoff, S., Macari, S., Young, G. S., Goldring, S., Thompson, M., & Rogers, S. (2008b). Atypical object exploration at 12 months of age is associated with autism in a prospective sample. *Autism*, 12, 457–472.
- Ozonoff, S., Young, G. S., Goldring, S., Greiss-Hess, L., Herrer, A. M., Steele, J., et al. (2008c). Gross motor development, movement abnormalities, and early identification of autism. *Journal of Autism and Developmental Disorders*, 38, 644–656.
- Pandey, J., Verbalis, A., Robins, D. L., Boorstein, H., Klin, A., Babitz, T., et al. (2008). Screening for autism in older and younger toddlers with the Modified Checklist for Autism in Toddlers. *Autism*, 12(5), 513–535.
- Perry, A., Cummings, A., Dunn Geier, J., Freeman, N. L., Hughes, S., LaRose, L., et al. (2008). Effectiveness of intensive behavioral intervention in a large, community-based program. *Research in Autism Spectrum Disorders*, 2, 621–642.
- Pinto-Martin, J. A., Young, L. M., Mandell, D. S., Pogosyan, L., Giarelli, E., & Levy, S. E. (2008). Screening strategies for autism spectrum disorders in pediatric primary care. *Journal of Developmental and Behavioral Pediatrics*, 29, 345–350.
- Robins, D., Fein, D., Barton, M., & Green, J. (2001). The modified checklist for autism in toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 31(2), 131–144.
- Sallows, G. O., & Graupner, T. D. (2005). Intensive behavioral treatment for children with autism: Four-year outcome and predictors. *American Journal on Mental Retardation*, 110(6), 417–438.
- Schopler, E., Reichler, R. J., & Renner, B. R. (1988). *The childhood autism rating scale*. Los Angeles, CA: Western Psychological Services.
- Siegel, B. (2004). *Pervasive developmental disorders screening test-II (PDDST-II)*. San Antonio, TX: Harcourt Assessment Inc.
- Siegel, B., Anders, T. F., Ciaranello, R. D., Bienenstock, B., & Kraemer, H. C. (1986). Empirically derived subclassification of the autistic syndrome. *Journal of Autism and Developmental Disorders*, 16(3), 275–293.
- Stone, W. L., McMahon, C. R., & Henderson, L. M. (2008). Use of the screening tool for autism in two-year-olds (STAT) for children under 24 months: An exploratory study. *Autism*, 12, 557–573.
- Ward-King, J., Cohen, I. L., Penning, H., & Holden, J. J. A. (in press). Brief report: Telephone administration of the autism diagnostic interview—Revised: Reliability and suitability for use in research. *Journal of Autism and Developmental Disorders*.
- Wetherby, A. M., Brosnan-Maddox, S., Peace, V., & Newton, L. (2008). Validation of the Infant–Toddler Checklist as a broadband screener for autism spectrum disorders from 9 to 24 months of age. *Autism*, 12(5), 487–511.
- Wetherby, A., & Prizant, B. (2002). *Communication and symbolic behavior scales developmental profile* (1st normed ed.). Baltimore, MD: Brookes.
- Wolff, S. (2004). The history of autism. *European Child and Adolescent Psychiatry*, 13(4), 201–208.
- Yoder, P., Stone, W. L., Walden, T., & Malesa, E. (2009). Predicting social impairment and ASD diagnosis in younger siblings of children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 39, 1381–1391.
- Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J., & Szatmari, P. (2005). Behavioral manifestations of autism in the first year of life. *International Journal of Developmental Neuroscience*, 23, 143–152.