Using the Reliable Change Index Statistic to Analyze Developmental Changes in Infants at Risk for Autism Spectrum Disorder

Rachel Barber, B.Sc.

Applied Disability Studies

Submitted in partial fulfillment
of the requirements for the degree of

Masters of Arts

Centre of Applied Disability Studies, Brock University
St. Catharines, Ontario

Barber, Rachel A. Using the Reliable Change Index Statistic to Analyze Developmental Changes in Infants at Risk for Autism Spectrum Disorder. Master of Arts (Applied Disability Studies), August 2012, 61 pp., 10 tables, 5 illustrations, references, 38 titles.

Longitudinal studies of the development of autism spectrum disorders (ASD) provide an understanding of which variables may be important predictors of an ASD. The objective of the current study is to apply the reliable change index (RCI) statistic to examine whether the Parent Observation of Early Markers Scale (POEMS) is sensitive to developmental change, and whether these changes can be quantified along a child's developmental trajectory. Ninety-six children with older siblings with autism were followed from 1-36 months of age. Group-based RCI analysis confirms that the POEMS is capable of detecting significant changes within pre-defined diagnostic groups. Within-subject analysis suggests that ongoing monitoring of a child at-risk for an ASD requires interpretation of both significant intervals identified by the RCI statistic, as well as the presence of repeated high (i.e., >70) scores. This study provides preliminary evidence for a reasonably sensitive and specific means by which individual change can be clinically monitored via parent report.

Key Words: autism, parent report, reliable change, longitudinal, siblings

Acknowledgements

I would like to thank Dr. Maurice Feldman for introducing me to behaviour analysis and for mentoring me through research opportunities that are so positively influential on the lives of individuals with disabilities. I would also like to thank Dr. Jan Frijters for introducing me to the exciting world of statistics in social sciences – you inspire me. I also thank Dr. Rebecca Ward for your influence, both academically and personally.

I must recognize all of the families that agreed to participate in these studies over the past number of years. Your contribution to bettering our understanding of the developmental course of autism is invaluable.

I would also like to thank my family for their sustained support of my education and career. I would not have had the strength to persevere to see this project to its completion without you.

TABLE OF CONTENTS

List of Tables

Page

List of Illustrations

Page

CHAPTER 1

Introduction

Early Signs of ASD

Sibling recurrence rate of the broad spectrum of autism spectrum disorders (ASD) has recently been estimated at 18.7%, which represents almost 200 times the risk of the general population (Ozonoff et al., 2011). Genetically, concordance in monozygotic (MZ) and dizygotic (DZ) twins is 60% and 0% for autistic disorder (AD), respectively, and 92% and 10% for broader autism-related impairments (Bailey et al., 1995). Given the increased risk among the sibling population, prospective longitudinal studies investigating early signs of ASD are able to delineate two groups of study: 'high-risk' infant siblings of children with autism, and 'low-risk' infant siblings of children with typical development (Landa and Garrett-Mayer, 2006). Use of the same instruments to assess both groups longitudinally enables researchers to compare developmental differences and identify at what age these differences become obvious (Gamliel, Yirmiya, & Sigman, 2007).

Zwaigenbaum et al. (2009) summarized the early characteristics found through prospective research to be associated with a developmental trajectory towards an ASD. They suggested six domains in which delays or impairments have consistently shown to be predictive of a later ASD diagnosis: visual, motor, play, social-communication, language, and general cognitive development (p.1384). A more recent examination of at-risk infants (Feldman et al., 2012) corroborated the findings that early social-communication deficits appear to be predictive of a later ASD diagnosis. They also found that some emotional differences, such as intolerance for waiting and an irritable

mood, were characteristics more likely to be observed in children with later ASD diagnoses than those with no diagnosis at three years of age. Zwaigenbaum et al. (2009) recommend that "active surveillance for early signs" (p.1386) be an integral part of the developmental monitoring of children at risk for an ASD; however, based on the findings from Feldman et al. (2012), a broader ASD screening tool that includes behaviours typically seen in children with ASD may be more effective in detecting characteristics that are not necessarily core features of an ASD, but are still developmentally related to a later ASD diagnosis.

Longitudinal Analysis of Development

Prediagnostic symptoms have been identified through both retrospective and prospective studies. Using parent reports (Chawarska et al., 2007) and videotapes of the child's behaviours (Wetherby et al., 2004), retrospective analyses have identified early signs of an ASD emerging prior to the child's second birthday. Supporting these findings, prospective studies have been able to follow at-risk infants longitudinally in an attempt to more accurately detect at what point in early development these symptoms tend to emerge. Using a novel direct observation measure, Zwaignebaum et al. (2005) assessed 65 high-risk infants and 23 low-risk infants every 6 months, up to 24 months of age. They questioned whether specific behavioural risk markers were able to predict diagnostic classification at 24 months. Through cross-sectional ANOVA analyses, they determined that observable differences were noted between the two groups within the first year of life. Bryson et al. (2007) later reported the individual case trajectories of those children diagnosed with an ASD. Taken together, these studies provided not only evidence to support early continuous monitoring of child development, but also the need

for measures to identify idiosyncratic differences that may be used to predict later diagnostic status.

Landa and Garrett-Mayer (2006) assessed 87 children longitudinally at the ages of 6, 14, and 24 months using the standardized Mullen Scales of Early Learning (MSEL; Mullen, 1995). They compared three distinct groups based on the diagnosis of the child at the 24 month assessment: ASD, Language Delayed (LD), and Unaffected. Using ANOVA analyses, they found no statistically significant differences at 6 months on any of the MSEL domains. By 14 months of age, however, statistically significant differences emerged between the ASD and Unaffected groups on all MSEL domains except Visual Reception; significant differences continued at 24 months of age. Building upon the cross-sectional comparisons, they further explored the data using longitudinal regression models, and found that the period between 14 and 24 months of age "was one of particular vulnerability for toddlers with ASD" (p. 635). That is, a slowing in developmental trajectory was observed for the ASD group between 14 and 24 months of age when compared to the other two groups. The longitudinal modeling provided an estimation of the mean trajectories differentiating those with and without ASD, and those with language delays.

Gamliel, Yirmiya, Jaffe, Manor, and Sigman (2009) also used longitudinal growth curve analysis to compare language and cognition scores of siblings of children with autism from 4 to 54 months. Like Landa and Garrett-Mayer (2006), Gamliel et al. (2009) assessed children at 4, 14, and 24 months of age, but they also continued to monitor these children at 36, 54 months, and 7 years of age. Using the 7 year assessment results, they identified 37 siblings with an ASD diagnosis (SIBS-A) and 47 with typical development

(SIBS-TD), and compared the trajectories of standardized scores on both language and cognitive measures. Scores on standardized language and cognitive tests were used as variables for the growth curve analysis, with the outcome variable being diagnostic status. They summarized that the SIBS-A group had both significantly lower language scores after 14 months and a slower rate of development compared to the SIBS-TD group. Cognitive scores for the SIBS-A group were also significantly lower after 14 months; however, the trajectories were parallel, suggesting a similar rate of development over time.

Ozonoff et al. (2010) utilized both prospective observations and retrospective parent report to compare the social-communicative behaviours of 25 children with autism to those of a matched sample of 25 children with typical development. Participants were blindly evaluated at 6, 12, 18, 24, and 36 months of age; videos were coded for frequency of gaze to faces, social smiles, and directed vocalizations. Best-fit regression models were then developed for each group using the aforementioned coded results as well as examiner ratings of social engagement. Results suggested that social differences do not start to emerge until after 6 months of age. They also used the MSEL (Mullen, 1995) to assess cognitive changes across time, and unlike Gamliel et al. (2009), they found that the rate of growth was slower in children with ASD on all domains compared to those with typical development. Interestingly, although approximately 86% of the children with ASD demonstrated declines in social behaviours through the prospective observations, 83% of the retrospective parent reports did not identify any loss of skills during the same period of time. This study corroborates evidence that differences emerge between the

ages of 6-18 months, but it also underscores the importance of prospective monitoring for early signs of ASD.

Based on conclusions from the above studies, children with ASD are likely to first show developmental differences compared to peers with typical development in the areas of cognition, language, and social-communication sometime between the ages of 6 months and 18 months. One of the limitations of the previously published studies is the lapse of time between assessments, and the resultant inability to more specifically identify at what point in time changes in development occur that eventually differentiate diagnostic groups.

In response to this limitation, Lord, Luyster, Guthrie, and Pickles (2012) increased the frequency of assessment within a narrower period of time in an attempt to monitor change as it occurred within individual participants. Children were evaluated with the observational diagnostic test, Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) approximately every two months from 18 to 36 months of age. Longitudinal ADOS algorithm scores were used with latent class modeling and each child was assigned to one of four classes based on best-fit probability analyses: Severe Persistent, Worsening, Improving, and Nonspectrum. The authors concluded that "trajectories of growth for core features of autism… were quantifiable and predictable in children as young as 15-18 months of age" (p. 9); however, there was great variability in the trajectories of individual children. They also found that having two assessments in three months significantly improved the prediction of a trajectory class. This result confirms the need for more frequent, repeated assessments of children, particularly of those at risk for an ASD.

Along with the increased density of assessments comes an increased necessity for clinician involvement. For families in more remote geographical areas with minimal access to specialized developmental practitioners or assessment procedures, this requisition for multiple clinical visits is highly impractical. In an effort to increase the availability and frequency of developmental monitoring, Feldman et al. (2012) developed the Parent Observation of Early Markers Scale (POEMS). It is described as a checklist for prospective parent monitoring of "possible early symptoms and associated behaviours of an ASD" (p.15). This measure enables parents to repeatedly monitor and track their child's development within natural settings, and without requiring a clinician to complete a formal observation. This original POEMS study followed 108 at-risk children to 24months of age. By the conclusion of the study, 9 children had been diagnosed with an ASD by 36 months. Between-group analyses were conducted, comparing the differences in overall POEMS scores as well as the number of elevated items between the diagnosed group and the non-diagnosed group at 9, 12, 18, and 24 months of age. The diagnosed group had significantly higher POEMS scores and more elevated items at each age. There was also a statistically significant age by diagnosis interaction, suggesting that the difference in scores increased as the child aged. Differences between the two groups began to emerge at 9 months of age.

Some of the limitations of the previously published longitudinal literature are the small group sizes used for analysis and the lack of follow-up data. Fountain, Winter, and Bearman, (2012) responded to these limitations by using annual records for 6975 children aged 2 to 14 years with a diagnosis of autism. They used group-based trajectory modeling to "identify and describe subgroups within the data that have similar

developmental trajectories" (p. 3) and found 6 distinct models that best describe each of the social, communication, and repetitive behaviour development of children with ASD. One of the difficulties in interpreting these results within a practical framework is that individual variation is masked by group means and group modeling. For instance, it was noted within this study that a given child could have a different trajectory path for each of the three domains measured.

Within the literature, there is a mixture between cross-sectional analyses and latent class regression modeling. Regardless of methodology, the majority of studies relied on group comparisons at specific times to inform conclusions about group variation. Group based modeling allows for the classification of global paths of development for groups of children with similar trajectories; however, intra-individual changes between assessment A and assessment B may not necessarily follow the same mean patterns, and thus may not be identified as clinically significant 'red flags' during typical child development monitoring. In fact, all of the studies report obvious heterogeneity amongst the developmental trajectories of children with autism. Lord et al. (2012) attempted to explore the variation in individual trajectories demonstrated the predictive capacity of four different regression models. These monitoring methods, however, require substantial clinical support that may not be available on a wide-spread population level. Thus, the need becomes apparent for a reasonably simple method by which clinically significant change can be identified in practice on an individual level.

Reliable Change Index (RCI)

The reliable change index (RCI) is a statistic used to determine whether a change in an individual's (or group's) score is statistically significant based on the

measurement's test-retest reliability (Jacobson & Truax, 1991). It provides information regarding the likelihood that a change in test scores "results from 'true' or reliable change or results from chance" (Hensel, Angermeyer, & Riedel-Heller, 2007, p.1360). Jacobson, Follette, and Revenstorf (1984) first proposed a mathematical means to report statistically reliable improvement in psychotherapy research. Christensen and Mendoza (1986) suggested an alteration to the original formula to account for estimates of the standard error of difference, which was accepted and adopted by Jacobson and Truax (1991) in their publication of the reliable change index. Although a number of variations of the 'JT method' have been proposed over the years (Hsu, 1989; Speer, 1992), discussions and analyses have concluded that the classical reliable change index is the most appropriate statistic to use when population-based information is unavailable and when practice effects are unlikely to affect assessment scores through repeated testing (Atkins, Bedics, McGlinchey, & Beauchaine, 2005; Maassen, 2004).

Purpose and Hypotheses

All of the longitudinal studies discussed above used cross-sectional comparisons and group-based modeling to identify variables that predict an ASD diagnosis. Changes on these variables are reported as occurring between 10-18 months of age; however, questions still remain as how to measure reliably significant change from one assessment to the next. With studies emerging that provide a better understanding of what variables may be important predictors of an ASD, questions about when and how much individual change must occur to be considered clinically significant should be explored. These intra-individual analyses can then be applied to screening procedures during ongoing developmental monitoring. The purpose of the current study is to apply the RCI statistic to examine whether the POEMS is sensitive to developmental change, and whether these changes can be identified at specific time intervals within a child's developmental trajectory. It is hypothesized that children with an ASD diagnosis will be more likely than those without a diagnosis to have significant reliable increases in POEMS scores early in their development. It is also hypothesized that groups of trajectories will be identified based on the RCI intervals and stability of scores on the POEMS measure.

CHAPTER 2

Methods

Participants

Longitudinal data for 96 children were obtained through the Feldman et al. (2012) study. As described in the Feldman et al. (2012), families were recruited through the website, www.AutismResearch.ca, through presentations to family-based organizations such as Autism Ontario and Autism Society Canada, and via the distribution of pamphlets to physician offices, developmental clinics, and the aforementioned presentations. Initially, 239 families registered online, but upon further inquiry, a number of families were deemed ineligible because they did not meet the eligibility criteria of either having a 0-24 month biological sibling of a child with ASD or having an older child with a confirmed diagnosis of ASD. Other families declined participation due to the time commitment required for the multi-year study. In total, 118 families consented to participate in the study; however, 17 families decided not to continue with the assessment process for unknown reasons. The final group of participants consisted of 108 infants from 103 families. This included one set of identical and one set of fraternal twins. Three additional infants were born to included families while their older infant was already enrolled in the study. None of the siblings had biological or birth conditions that are known to be associated with potential developmental concerns (e.g., Down syndrome, low birth weight, epilepsy) and none of the probands had Fragile X or other syndromes related to ASD.

In order to be included in the current study, family inclusion criteria were defined as having at least one biological child with an independent diagnosis of ASD and a

younger biological sibling between 1-30 months of age at the time of their first assessment. Given the longitudinal nature of this analysis, an additional inclusion criterion was having at least three assessments completed within the study period. In total, a subset of 96 participants from the Feldman et al. (2012) study of 108 infant siblings was included in the current analysis. Table 1 provides demographic information on the participating children and families. *T*-test analyses using the available information on the specified demographic characteristics yielded no significant differences between those families included and those excluded.

Table 1

Demographic characteristics of the 96 participating children and families.

Variable		Percentage of participants for whom data were available
Mean age of infant at first POEMS assessment (months) (*SD*)	8.44 (5.71) (range: 1.03-28.00 months)	100%
Percentage male infants	63%	100%
Mean age of proband at first POEMS assessment (months) (*SD*)	71.00 (35.32)	86%
Mean age of mothers at first POEMS assessment (years) (*SD*)	35.32 (3.94)	75%
Percentage of mothers with at least a college/university degree	88%	75%
Percentage of mothers employed other than or in	65%	75%

addition to homemaker		
Mean age of fathers at first POEMS assessment (years) (*SD*)	37.51 (4.28)	75%
Percentage of fathers with at least a college/university degree	69%	75%
Percentage of fathers employed other than or in addition to homemaker	96%	75%
Range of annual family income (lowest-highest income range)	$5 000 to > $95 000	63%
Percentage of two-parent families	97%	75%

Measure

The overall scores from the Parent Observation of Early Markers Scale (POEMS; Feldman et al., 2012) were used for analysis. Each of the 61 items on this measure is scored by a parent on a Likert-type scale, ranging from 1 (no problem) to 4 (severe problem) with 1/2 scores allowed. Parents were encouraged to score the checklist based on actual observation of their child. If an item was developmentally advanced for their child's age, parents were instructed to score N/A which was then converted to a score of 1 for the data analyses. Tests were completed by the child's primary caregiver via mail, online, or by telephone interview. Psychometric properties were deemed acceptable (Feldman et al., 2012) based on comparisons to the Ages and Stages Questionnaire (ASQ; Bricker & Squires, 1999) and the Autism Diagnostic Interview – Revised (ADI-R; Lord, Rutter, & Le Couteur, 1994).

CHAPTER 3

Analysis 1: Between-Group Comparison of Developmental Changes

Methods & Results

***POEMS Test-Retest Reliability*.** POEMS test-retest reliability was calculated for three month intervals, using the raw data obtained from the 96 participants that met criteria for study inclusion (see Table 2). If a participant had multiple assessments recorded within a given interval, an average score was calculated for the interval. The *N* value represents the number of participants that contributed an assessment value to the specified interval. Given the lack of a comparison interval, test-retest reliability was not calculated for the last interval (36+ months); rather, an average reliability score was imputed.

Table 2

POEMS test-retest reliability calculated for between-group analysis.

Age (mos)	**0-3**	**4-6**	**7-9**	**10-12**	**13-15**	**16-18**	**19-21**	**22-24**	**25-27**	**28-30**	**31-33**	**34-36**	**36+**
N	18	44	45	66	61	62	57	64	52	41	37	38	29
Test-Retest Reliability	0.94	0.93	0.81	0.60	0.92	0.98	0.96	0.99	0.99	0.99	0.96	0.42	0.87*

* Imputed average reliability value.

Descriptive statistics. Individual raw scores were used to calculate values of central tendency. The number of assessments that were completed within each time interval is listed in Table 3 along with the mean total POEMS scores and standard deviation across all participants.

Table 3

Descriptive statistics based on 3-month intervals

	Age (months)												
	0-3	**4-6**	**7-9**	**10-12**	**13-15**	**16-18**	**19-21**	**22-24**	**25-27**	**28-30**	**31-33**	**34-36**	**36+**
Number of assessments	18	44	45	66	61	62	57	64	52	41	37	38	29
Mean POEMS score	67.06	66.71	68.30	71.22	74.55	71.66	69.48	68.87	68.50	68.80	68.81	64.80	68.02
Standard deviation	8.45	9.96	11.86	12.50	19.44	18.80	20.50	19.52	20.28	23.08	24.17	8.31	23.68

Reliable Change Index. A standard error of change (S_E) value was then computed for each interval using the following formula:

where SD_x refers to the standard deviation of test 1 (i.e., the chronologically first test in the comparison), SD_y refers to that of test 2, and *rel* refers to the test-retest reliability value of the chronologically first time interval (Maassen 2004).

Reliable change was calculated by dividing the change in adjacent intervals' mean scores by the standard error of change of the measurement:

where x_1 and x_2 represent the mean total POEMS scores at intervals 1 and 2, respectively. Standard error and RCI scores are listed in Table 4.

Table 4

Standard error and reliable change index values for each diagnostic group across age intervals.

		Age (months)												
		0-3	**4-6**	**7-9**	**10-12**	**13-15**	**16-18**	**19-21**	**22-24**	**25-27**	**28-30**	**31-33**	**34-36**	**36+**
All Groups	Standard Error (S_E)	3.08	3.98	7.58	14.60	7.82	4.13	5.57	3.12	2.67	2.93	5.19	19.17	n/a

ASD Diagnosis	Number of assessments	4	6	9	6	13	9	4	7	4	1	2	0	1
	Mean POEMS score	81.3	93.2	87.1	80.7	111.5	114.6	103.3	113.4	104.5	189.0	138.5	N/A	188.0
	Reliable Change Index (RCI)		3.87**	-1.52	-0.85	2.11**	0.40	-2.75*	1.82	-2.84*	31.70**	-17.21*	n/a	n/a
Provisional or Lost ASD Diagnosis	Number of assessments	0	7	6	6	5	3	3	5	3	3	1	6	0
	Mean POEMS score	N/A	66.1	66.5	74.1	76.7	86.3	86.8	70.7	67.5	73.3	82.5	69.1	N/A
	Reliable Change Index (RCI)		n/a	0.09	1.00	0.18	1.23	0.12	-2.90*	-1.02	2.19**	3.12**	-2.59*	n/a
No ASD Diagnosis	Number of assessments	28	67	66	105	100	101	82	101	70	56	59	49	28
	Mean POEMS score	65.3	64.5	65.4	70.1	70.9	68.1	66.3	66.1	66.5	64.4	63.5	63.8	63.7
	Reliable Change Index (RCI)		-0.24	0.20	0.62	0.06	-0.36	-0.43	-0.04	0.11	-0.79	-0.29	0.05	n/a

Note. If the absolute value of the reliable change index (RCI) is greater than 1.96, the difference in scores between the two intervals is statistically significant at a 95% confidence interval. RCI values signifying a significant increase in scores between given intervals are denoted by (**), and those signifying a significant decrease are denoted by (*).

Within the ASD Diagnosis group (*n*=7), there is a significant reliable increase in overall POEMS scores between the intervals of 0-3 months and 4-6 months, and again between 10-12 months and 13-15 months intervals. These are followed by two significant decreases: one between the 16-18 month and 19-21 month intervals, and another between the 22-24 month and 25-26 month intervals. There is then a significant increase between the 25-27 month and 28-30 month intervals, which is followed immediately by a significant decrease into the 31-33 month interval. The later increase

must be interpreted with caution, however, since it represents only one participant's score. Within the Provisional/Lost Diagnosis group (*n*=4), there is a significant reliable decrease between the 19-21 month and 22-25 month intervals. This is followed by two significant reliable increases between the intervals of 25-27 months and 31-33 months and one significant decrease between the intervals of 31-33 months and 34-36 months. Again, it is important to note that only one assessment was completed within the 31-33 months interval, and this score may or may not be representative of the group. Finally, within the No ASD Diagnosis group (*n*=85), there are no intervals with significant reliable increases or decreases.

Discussion

The results of Analysis 1 suggest that the POEMS may be sensitive to developmental changes on a group level. The significant increase between 10-15 months within the ASD Diagnosis group is consistent with both the cross-sectional findings in the Feldman et al. (2012) POEMS study as well as the general finding that behavioural differences tend to emerge between 12-15 months of age (Landa & Garrett-Mayer, 2006; Ozonoff et al., 2010). This pattern also supports findings that parents typically identify behavioural abnormalities between 12-18 months of age (Young, Brewer, & Pattison, 2003; Zwaigenbaum et al., 2009). The subsequent significant decreases within the ASD Diagnosis group could indicate a general reduction in symptom severity around two years of age, or it could be an artifact of the small sample size from which the average scores were obtained.

There is a relatively high, but non-significant increase (*RCI* = 1.23) amongst the Lost/Provisional ASD Diagnosis group within the same developmental window as above.

This group also has later significant decreases that may reflect an 'improved' trajectory, in which autistic characteristics become less severe, resulting in the loss of a diagnosis. Unfortunately, the dates of diagnosis and/or loss of diagnosis were not known, so correlations between diagnostic status and trajectory patterns were not possible. These comparisons, however, may not be completely accurate even if the data were available, given the typical delay between parent identification of early signs and the time of diagnosis (Young et al., 2003).

The lack of reliable changes within the No ASD Diagnosis group suggests that this group trajectory is relatively stable across months. This, coupled with the presence of significant changes in the groups with an ASD diagnosis suggest promising predictive validity of the POEMS. It also corroborates the finding in Feldman, et al. (2012) that the POEMS has reasonable sensitivity and specificity.

CHAPTER 4

Analysis 2: Evaluation of the POEMS Using Within-Subject Analysis

Methods & Results

***Test-Retest Reliability*.** Test-retest reliability was calculated for each month of age between 1-36 months, using the raw data obtained from the 96 participants that met criteria for study inclusion (see Table 5). If a participant had multiple assessments recorded during a given month, an average score was calculated for that month. Assessments that were completed after the age of 36 months were included in the final age category as a consolidated '36+ month' age group. The *N* value represents the number of participants that contributed an assessment value to the specified age month. Given the lack of a comparison interval, test-retest reliability was not calculated for the last interval (36+ months); rather, an average reliability score was imputed. Also included in Table 5 are the average overall POEMS score and standard deviation for each month age group.

Table 5

Test-retest reliability for the POEMS.

Age (months)	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
N	8	13	16	19	27	29	21	31	21	44	28	49
Average POEMS	71.3	64.8	67.8	66.4	66.4	68.8	67.8	67.3	68.8	68.6	73.3	74.4
Score (SD)	(13.86)	(4.88)	(8.72)	(11.16)	(11.49)	(14.08)	(13.41)	(8.61)	(10.93)	(7.74)	(12.27)	(21.47)
Test-Retest Reliability	0.15	0.87	0.99	0.99	0.98	0.99	0.90	0.97	0.96	0.64	0.77	0.83

Age (months)	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
N	30	43	25	38	29	31	26	39	20	39	17	47
Average POEMS	76.6	73.8	73.8	71.7	70.4	74.8	69.0	70.2	68.2	67.1	71.5	68.6

Score (SD)	(14.04)	(22.92)	(15.97)	(21.07)	(14.54)	(26.02)	(13.49)	(24.46)	(13.02)	(18.93)	(22.17)	(20.92)
Test-Retest Reliability	0.86	0.93	0.87	0.94	0.95	0.99	1.00	0.99	0.97	0.95	0.99	0.86

Age (months)	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36+**
N	17	18	38	11	12	36	16	13	32	10	12	48
Average POEMS	68.1	70.4	69.2	62.9	62.7	67.0	64.0	77.0	63.8	62.3	63.3	67.2
Score (SD)	(19.69)	(18.59)	(23.25)	(4.23)	(3.04)	(21.35)	(5.21)	(39.63)	(4.71)	(1.81)	(2.47)	(19.58)
Test-Retest Reliability	1.00	0.99	1.00	0.99	0.91	0.67	0.87	0.86	0.77	0.97	1.00	0.893*

* Imputed average reliability value.

***Reliable Change Index*.** For this analysis, the original formula for standard error was used to calculate RCI scores (Jacobson & Truax, 1991):

The RCI statistic was applied to each age month of the 96 participants using the formulas described in Analysis 1. Half of the participants (n = 48) did not have any significant RCI values between assessments. All of the participants in both the ASD and Lost/Provisional ASD diagnostic groups had at least one significant positive or negative RCI value.

In order to further classify the 48 participants with significant RCI intervals, RCI values were examined for patterns. If a participant had a significant RCI value (i.e., an absolute value greater than 1.96) between two assessments, the time between the two assessments was considered ‘increasing’ if the RCI value was positive, or ‘decreasing’ if the RCI value was negative. Classification was then explored in three ways: visual pattern, actual number of increasing/decreasing intervals, and 12 month intervals.

Visual Pattern. Seven groups were identified based on the observable pattern of significant RCI intervals. The results are summarized in Table 6.

Table 6

Results of within-participant analysis using visual patterning as a basis for grouping.

	No ASD Diagnosis Frequency (%)	**ASD Diagnosis Frequency (%)**	**Lost/Provisional ASD Diagnosis Frequency (%)**	**Total**
Decreasers	15 (88)	2 (12)	0	17
Increasers	11 (92)	1 (8)	0	12
Decrease then Increase	4 (67)	1 (17)	1 (17)	6
Increase then Decrease	2 (67)	1 (33)	0	3
Increase, Decrease, then Increase	0	2 (67)	1 (33)	3
Decrease, Increase, then Decrease	4 (100)	0	0	4
Multiple	1 (33)	0	2 (67)	3
Total	37	7	4	48

Decreasers. A ‘Decreasers’ group was classified as only having decreasing intervals. There were 17 participants in this group, of which two were from the ASD diagnostic group. Seven of these participants had decreasing intervals within the first 12 months of age (one with an ASD diagnosis). A group of six participants had decreasing intervals only between 12-24 months of age (one with an ASD diagnosis). Two participants had decreasing intervals only after 24 months, and two other participants had multiple decreasing intervals after 12 months of age.

Increasers. An ‘Increasers’ group was classified as only having increasing intervals. Twelve participants comprised this group, of which one was from the ASD diagnostic group. Seven group members had increasing intervals only prior to 15 months of age, and four members had increasing intervals only after 28 months of age. The participant from the ASD diagnostic group had two increasing intervals between 6-18 months of age.

Decrease then Increase. A ‘Decrease then Increase’ group was identified as having one or more decreasing intervals preceding one or more increasing intervals. Six participants, including one from the ASD diagnostic group and one from the Lost/Provisional ASD diagnostic group, comprised this group. Two participants had both increasing and decreasing intervals after their second birthday. Two others had decreasing intervals prior to 12 months of age and increasing intervals after 24 months (one from the Lost/Provisional diagnostic group). One group member had both increasing and decreasing intervals between 18-24 months of age. The final participant in this group, who was diagnosed with ASD, had both decreasing and increasing intervals prior to 13 months of age.

Increase then Decrease. An ‘Increase then Decrease’ group was defined as having one or more increasing intervals precede one or more decreasing intervals. Three participants made up this group, of which one was from the ASD diagnostic group. All three participants had increasing intervals prior to 9 months of age; decreasing intervals ranged from 7 months of age (participant from ASD group) to 21 months.

Increase, Decrease, then Increase. An ‘Increase, Decrease, then Increase’ pattern was displayed by three participants; two of this group’s members were from the ASD

diagnostic group, and one was from the Lost/Provisional group. One of the participants (ASD group) had early increasing and decreasing intervals (i.e., 9-10 months) and then a late increasing interval (i.e., after 25 months). The other two members had the defining pattern between 12-24 months of age.

Decrease, Increase, then Decrease. All four of the participants in the 'Decrease, Increase, then Decrease' group were from the No ASD diagnostic group. Each of the group members had distinct patterns, not shared by other group members.

Multiple. A 'Multiple' pattern was seen with three participants, where alternating periods of increasing and decreasing RCI intervals were recorded. Two of these, also from the Lost/Provisional diagnostic group, showed an early decreasing interval (i.e., 5-6 months), then an increasing interval between 9-15 months, then a decreasing interval around their second birthday, and finally an increasing interval around 30 months of age. The third group member had multiple alternating increasing and decreasing intervals starting after 18 months of age.

Number of Increasing/Decreasing Intervals. The number of increasing versus decreasing intervals was calculated for each participant. If the number of increasing intervals was greater than the number of decreasing intervals, the participant was considered an 'Increaser'. Alternatively, if the number of decreasing intervals was greater than the number of increasing intervals, the participant was considered a 'Decreaser'. If, however, the number of increasing and decreasing intervals was the same, the participant would be considered 'Variable'. The distribution of participants across groups is listed in Table 7.

Table 7

Results of within-participant analysis using the number of increasing/decreasing intervals as a basis for grouping.

	No ASD Diagnosis Frequency (%)	**ASD Diagnosis Frequency (%)**	**Lost/Provisional ASD Diagnosis Frequency (%)**	**Total**
Increasers	11 (61)	5 (28)	2 (11)	18
Decreasers	23 (92)	2 (8)	0	25
Variable	3 (60)	0	2 (40)	5
Total	37	7	4	48

Increasers. Within this group of 18 participants, five were from the ASD diagnostic group and two were from the Lost/Provisional group. The majority of participants (*n* = 13), including three diagnosed with ASD, had increasing intervals prior to 12 months and/or after 27 months of age. Three participants had increasing intervals between 12-24 months of age; one of these three had an ASD diagnosis and one was from the Lost/Provisional group. Two participants (one from each of the ASD and Lost/Provisional groups) had a decreasing interval around 22 months of age that was both preceded and followed by increasing intervals.

Decreasers. This was the largest group category (n = 25), and it included two from the ASD diagnostic group. The majority of participants (n = 15) had decreasing intervals either prior to 14 months or after 25 months of age; whereas the other 10

participants had decreasing intervals between 12-24 months. Each of the aforementioned patterns included one of the participants with an ASD diagnosis.

Variable. Two of the five participants that comprised this group were from the Lost/Provisional diagnostic group. Four participants, including both of the Lost/Provisional group members, had increasing intervals prior to 12 months and/or after 24 months of age. Three of the participants had decreasing intervals between 12-24 months of age (one from the diagnostic group).

Twelve Month Intervals. Since some of the participants had time lapses greater than six months between assessments, a 12 month interval was decided upon for analysis. This reduced the risk that one significant RCI value would be misrepresented as multiple 'increasing' or 'decreasing' intervals for a given participant.

Each of the three 12 month intervals (0-12 months; 13-24 months; 25-36 months) were then classified as "Increasing" if a positive RCI interval was contained within the 12 months, and "Decreasing" if a negative RCI interval was contained within the 12 months. If a participant had multiple significant RCI values within a given 12 month interval, the strongest value was used to classify the interval as 'increasing' or 'decreasing'. For example, if a participant had an RCI of +2.56 between ages 4-5 months and an RCI of -4.35 between ages 8-9 months, the 0-12 months interval would be considered 'decreasing' since the absolute value of the negative RCI value is greater than that of the positive RCI value. Using these criteria, five groups were defined: Increasers; Decreasers; Increase then Decreasers; Decrease then Increasers; and Increase, Decrease, then Increasers. Group composition is summarized in Table 8.

Table 8

Results of within-participant analysis using 12-month intervals as a basis for grouping.

	No ASD Diagnosis Frequency (%)	**ASD Diagnosis Frequency (%)**	**Lost/Provisional ASD Diagnosis Frequency (%)**	**Total**
Increasers	15 (83)	3 (17)	0	18
Decreasers	18 (90)	2 (10)	0	20
Increase then Decrease	3 (75)	0	1 (25)	4
Decrease then Increase	1 (25)	2 (50)	1 (25)	4
Increase, Decrease, then Increase	0	0	2 (100)	2
Total	37	7	4	48

Increasers. 'Increasers' were defined as those participants with no 'decreasing' intervals and at least one 'increasing' interval. There were 18 participants who classified as 'Increasers', of whom three were in the ASD Diagnosis group. Nine participants had increasing intervals in the first 12 months; five participants had increasing intervals in the middle 12 months; and six participants had increasing intervals in the last 12 months. Two of the participants from the ASD diagnostic group had significant increases within the first trimester, whereas the third ASD group member had increasing intervals in both the first and second trimesters.

Decreasers. "Decreasers" were defined as those participants with no 'increasing' intervals and at least one 'decreasing' interval. There were 20 participants in this

category, including two participants from the ASD Diagnosis group. Nine participants (one with an ASD diagnosis) had decreasing intervals within the first 12 months. Again, nine participants (one with an ASD diagnosis) had decreasing intervals in the second 12 month interval; and five participants had decreasing RCI values in the third 12 month interval.

Increase then Decrease. Four participants comprised the 'Increase then Decrease' group, where at least one 'increasing' interval preceded at least one 'decreasing' interval; one of the four was from the Lost/Provisional ASD Diagnosis group. Three participants (one from the diagnostic group) had increasing intervals in the first trimester, and one had their second trimester classified as increasing.

Decrease then Increase. The 'Decrease then Increase' group – defined as having at least one 'decreasing' interval that preceded at least one 'increasing' interval – also consisted of four participants. This group contained two participants from the ASD Diagnosis group and one participant from the Lost/Provisional ASD Diagnosis group. Three of the group members (one from each diagnostic group) had their first 12 month interval classified as decreasing; two members (one with and one without an ASD diagnosis) had increasing intervals in the second 12 months. All four participants had an increasing interval around 30 months of age (i.e., the third trimester).

Increase, Decrease, then Increase. Finally, the 'Increase, Decrease, then Increase' group criteria was such that the participant's 0-12 months and 25-36 months intervals were 'increasing', and their 13-24 months interval was decreasing. Two participants met these criteria, and both of them were from the Lost/Provisional ASD Diagnosis group. Both participants had decreasing and increasing intervals in the first

trimester, and one of them had increasing and decreasing intervals in the second trimester. Both group members had an increasing interval around 30 months of age.

Discussion

In this analysis, significant RCI intervals calculated from the overall POEMS scores were used as a basis for qualitative classification of developmental trajectories for children at risk of an ASD diagnosis. There was significant heterogeneity within this population, as expected given the suspected variation in interactions between neurobiological, genetic, epigenetic, developmental, and environmental factors within each child (Fountain et al., 2012).

There were three different interpretations of the RCI data presented. The first, where groupings were established based on a visual analysis of the data patterns, provided a differentiation of seven potential trajectories. Those without an ASD diagnosis were more likely to have only increasing or only decreasing RCI intervals (i.e., 70% of the 37 participants); however, those with a diagnosis at any point in their development were more likely to have some combination of the two (i.e., 73% of the 11 participants).

The second interpretation – using relative numbers of increasing and decreasing significant RCI intervals to classify participants – resulted in three different groups. Those diagnosed with an ASD were more likely to have a greater number of increasing RCI intervals (i.e., 71% of the 7 participants), which is consistent with the findings that developmental symptoms emerge after a number of months of otherwise typical development (Zwaigenbaum et al., 2009). On the other hand, those without a diagnosis were more likely to have a greater number of decreasing RCI intervals (i.e., 62% of the

37 participants), suggesting that these children may have an alternative rate of development that eventually matches that of their same-age peers. This trajectory may serve to explain why this group does not eventually get an ASD diagnosis. Those from the lost/provisional ASD diagnostic group were most likely to have a variable or increasing pattern (i.e., 50% of the 4 participants in each case), which corresponds to the fluctuations expected with a tentative or lost diagnosis.

Finally, the third method of interpretation, where 12-month intervals were classified as increasing or decreasing, outlined 5 different groups. Compared to the two other interpretations, however, this method was the most affected by when assessments occurred across the 36-month period. That is, if a participant had assessments in any of the three 12-month intervals but not others, their classification would be dependent on only the interval(s) in which they were assessed, rather than a truly representative classification of their 3-year trajectory. Although this method was an attempt at consolidating the data across larger intervals, it is likely the least informative given the density of the data set.

The variation in degree of change detected by the RCI analysis can be explained by a number of different factors. There is the possibility that the POEMS was not able to detect changes on a monthly basis to a significant degree. That is, the rate of change within a given participant may have been too gradual over time, and the spikes and plunges detected by the RCI analysis were not observed. Another explanation could be related to the fact that there were large gaps between assessments for the majority of participants. It could be the case where significant changes were missed because they occurred during the void in assessments. The same can be said for the reverse scenario

where a multi-month interval was falsely considered significant because monthly assessments were not completed within the RCI interval. More frequent assessments may have shown a more gradual change rather than a single significant increase or decrease.

In order to account for the discrepancies in assessment intervals, the Jacobson and Truax (1991) formula for standard error was used for this analysis, despite evidence suggesting the formula used in Analysis 1 may be preferred (Maassen, 2004). Standard errors for each month of assessment were determined using the whole sample population in order to maximize generalizability of the results and to minimize sampling error. If the formula discussed in Maassen (2004) were used, it would be necessary to assume that the intervals in which change were being measured were consecutive for all participants. This assumption, however, cannot be met with the current dataset; given the discrepancy in time between assessments across individual participants, it was not appropriate to pool consecutive variances when determining an overall standard error for each month. That is, the pooled consecutive variances would not be consistently reflective of the true interval in which changes were being measured for any participant. As such, the Jacobson and Truax (1991) formula was used to account only for the initial month's variance in order for the standard error to remain generalizable across all participants.

Given the RCI formula is influenced by the reliability and standard deviation of the measure, it is important to consider these in the RCI interpretation. The test-retest reliability of the measure is reasonable (i.e., greater than .75) for all 36 months except for three: 1month, 10months, and 30months. The 1-month interval, though, only had eight assessments recorded. There was great variation in the standard deviation of the measure

across the 36 month period; however, this is likely related to the fact that the number of assessments recorded at each month also varied greatly, ranging from 8 to 49 assessments.

One of the limitations of this study is that the density of the data set was low given the number of participants. Ideally, each participant would have completed at least one assessment per month across the 36month time span, resulting in a total of 3456 completed assessments. In reality, there were 953 assessments completed, representing a 27.6% response rate. This is lower than the 55.6% that has been reported as a reasonable response rate in academic studies (Baruch, 1999). When the assessments are distributed among the monthly intervals, the relatively small *n*'s do have a negative effect on the accuracy of the RCI (Maassen, Bossema, & Brand, 2009), such that the standard deviations of some of the monthly means are quite large (range: 1.81 to 39.63). Crawford and Howell (1998) did suggest, however, that the standard error of change does not fluctuate greatly when the sample size ranges from 25 to 500.

One reason for the lower response rate could be that the inclusion criteria allowed for participants to enter the study at any point between 0-24 months, resulting in not all participants having their first assessment at 1month of age. Additionally, some participants' families opted to complete the POEMS at various times throughout the assessment period, so that RCI intervals were quite variable (range: 1-16months). This saw some children represented only in small windows of the assessment period. Trajectory classifications of these children may have differed had they recorded assessments in a less condensed time frame.

CHAPTER 5

Analysis 3: Longitudinal Analysis of Overall Scores Using a Cut-Off Score of 70

Methods & Results

As reported in Feldman et al. (2012), a cut-off score of 70 was used to establish the sensitivity of the POEMS. This score was chosen "because it was about midway between the means in the full sample at each age examined" (p.19). Using this score as the deciding value of a clinically significant or insignificant score, the raw data from the eligible 96 participants were examined for longitudinal patterns. Again, if a participant had multiple assessments recorded during a given month, an average score was calculated for that month. If the score was equal to or greater than 70, it was considered a 'high' score; if it was below 70, it was considered a 'low' score. Assessments that were completed after the age of 36 months were included in the final age category as a consolidated '36+ month' age group.

Five groups were established based on the distribution of 'high' or 'low' scores across the duration of the study for each participant: Stable Low, Stable High, Decreasers, Increasers, and Variable. See Table 9 for a summary of the groupings along with diagnostic status. Figures A1-A5 in the Appendix depict visual representations of the participants' assessment trajectories that comprise each group.

Table 9

Results of the longitudinal analysis of raw scores using a cut-off score of 70.

	No ASD Diagnosis	**ASD Diagnosis Frequency (%)**	**Lost/Provisional ASD Diagnosis**	**Total**

	Frequency (%)		Frequency (%)	
Stable Low	44 (100)	0	0	44
Stable High	3 (38)	5 (62)	0	8
Decreasers	15 (100)	0	0	15
Increasers	1 (25)	2 (50)	1 (25)	4
Variable	22 (88)	0	3 (12)	25
Total	85	7	4	96

Stable Low. This group was defined as having all scores below 70 across all assessment months. It was the largest group with 44 participants, all of whom were from the No ASD diagnostic group. Despite six of the participants having a significant increasing RCI interval and five having a significant decreasing RCI interval, none of the scores went above 70.

Stable High. In contrast, this group was defined as having all scores equal to or greater than 70 across all assessments. Five of the eight participants that comprised this group were from the ASD Diagnosis diagnostic group. Six members of this group had decreasing RCI intervals and four had increasing RCI intervals, but all scores remained above 70. Only one member did not have any significant RCI intervals.

Decreasing. This group was defined as having only 'high' scores at the start of their assessment trajectory and having only 'low' scores at the end. There were 15 participants in this group, all from the No ASD diagnostic category. Within each participant's trajectory, only 4 (27%) had significant decreasing RCI intervals at the point at which the scores went from 'high' to 'low'. One other participant had a significant

decreasing RCI interval, but it was well after his change from ‘high’ to ‘low’. Four participants had significant increasing RCI intervals either before or after the change in scores.

Increasing. This group had the opposite pattern to the Decreasing group, with only ‘low’ scores in their early assessments and only ‘high’ scores in their later assessments. There were four participants in this group; two of the four were from the ASD Diagnosis group and one was from the Lost/Provisional ASD diagnostic group. All four participants (100%) had significant increasing RCI intervals at the point at which the scores changed from ‘low’ to ‘high’. Three of the four participants also had significant decreasing RCI intervals either before or after the point of change.

Variable. This final group was comprised of 25 participants, of whom three were from the Lost/Provisional ASD diagnostic group. This group was loosely defined as having ‘high’ and ‘low’ scores alternating throughout their assessments. Across all members, there were 36 points at which the scores went from ‘high’ to ‘low’, and of these, 15 (42%) were represented by a significant decreasing RCI interval. Additionally, there were 32 points at which the scores went from ‘low’ to ‘high’, and 9 of these points (28%) corresponded to significant increasing RCI intervals.

ASD Diagnosis and Lost/Provisional ASD Diagnostic Groups Only. Separating out the participants with an ASD diagnosis, five of the seven children have stable high scores across their assessment period. Although all five participants have significant decreasing RCI intervals at some point in their assessment period, none of their scores reach 70 or below. The remaining two children with an ASD diagnosis each had only one point within their assessment period where their score went from below 70 to above

70, and both of these changes correspond to significant increasing RCI intervals. Therefore, of the two points at which the scores went from 'low' to 'high', 100% were represented by a significant increasing RCI interval.

Looking at the Lost/Provisional ASD Diagnosis group, one of the participants had only one change from below 70 to above 70, and this change was represented by a significant increasing RCI interval. The other three members from this group had multiple points within their assessment period where their scores fluctuated above or below 70. In total, across all four members of this diagnostic group, there were four points in time where scores went from 'high' to 'low', and 2 of them (50%) corresponded to significant decreasing RCI intervals. On the other hand, there were five points at which the scores went from 'low' to 'high', and four of them (80%) were represented by significant increasing RCI intervals.

Relationship Between RCI and Cut-Off Score of 70. Across all participants, there were 56 significant decreasing RCI intervals. Of these, 19, or 34%, represented a change from a 'high' score to a 'low' score, using 70 as the cut-off score. There were 23 intervals (41%) where the scores remained at 70 or above (i.e., had 'high' scores on both ends of the significant interval) and the remaining 14 intervals (25%) had scores remain below 70 (i.e., had 'low' scores on both ends). See Table 10 for a summary of the results.

Table 10

Proportion of the significant RCI intervals that represent a change in scores across the cut-off score of 70.

Decreasing RCI Interval			Increasing RCI Interval		
'High' to	**'High' to 'Low'**	**'Low' to 'Low'**	**'High' to**	**'Low' to 'High'**	**'Low' to 'Low'**

	'High' Frequency (%)	**Frequency (%)**	**Frequency (%)**	**'High' Frequency (%)**	**Frequency (%)**	**Frequency (%)**
ASD Diagnosis	6 (26)	0	1 (7)	9 (45)	2 (14)	2 (13)
Lost/Provisional ASD Diagnosis	2 (9)	2 (11)	2 (14)	4 (20)	4 (29)	1 (7)
No ASD Diagnosis	15 (65)	17 (89)	11 (79)	7 (35)	8 (57)	12 (80)
Total	23	19	14	20	14	15

Note. 'High' scores are scores equal to or above 70 based on the overall POEMS score. 'Low' scores are scores below 70. The scores being considered are those that start and end a given RCI interval.

There were 49 significant increasing RCI intervals, of which 14 (29%) corresponded with a change from below 70 to above 70. Twenty intervals (41%) had scores that remained at 70 or above (i.e., 'high' scores on both ends) and 15 intervals (31%) with scores remaining below 70 (i.e., had 'low' scores on both ends).

Looking at the ASD diagnostic group exclusively, there were seven significant decreasing RCI intervals. Six of the intervals had scores remain above 70 and one had scores remain below 70. None of the intervals represented a change from a 'high' score to a 'low' score. There were also 13 significant increasing RCI intervals, of which two represented a change from a 'low' to a 'high' score. Nine of the intervals had scores remain above 70, and two of the intervals had scores remain below 70.

When the Lost/Provisional ASD diagnostic group was examined, there were six significant decreasing RCI intervals. There were two that represented a change from 'high' to 'low' scores, two where the scores remained above 70, and two where the

scores remained below 70. Finally, there were nine significant increasing RCI intervals, of which four corresponded with a change from a 'low' score to a 'high' score. Four of the intervals had scores remain above 70 and one of the intervals had scores remain below the cut-off score.

Discussion

In this analysis, a total of 105 significant RCI intervals were identified across 96 participants, of which 33, or 31%, corresponded with a change across the clinical cut-off score of 70 on the POEMS measure. The remaining significant intervals were defined by scores that either remained above or below the cut-off 70 score. For example, although there were significant intervals – either increasing or decreasing – within the Stable Low and Stable High groups, their scores did not cross the cut-off score of 70 in either direction.

Analyzing the data from the reverse perspective, across all participants, there were 36 points at which the scores went from 'sub-clinical' (below 70) to 'clinical' (above 70). Of these, 36% corresponded to significant increasing RCI intervals. In the remaining 64% of the points of change, the situation occurred such that a score reached clinical significance without a corresponding statistically significant change. There were also 51 points at which the scores went from 'clinical' to 'sub-clinical', and 37% of these were represented by significant decreasing RCI intervals. These discrepancies could indicate a change in a participant's symptoms between assessments at a rate too slow to be considered statistically significant. Alternatively, it could suggest that the change between sub-clinical and clinical classification, or vice versa, could be just a random fluctuation due to error in the POEMS measurement.

In terms of the ASD Diagnosis diagnostic group, it is interesting to note that five of the seven group members had stable high scores and the remaining two had increasing trajectories. Of the 85 members of the No ASD Diagnosis group, there were only three participants that had stable high scores and one participant that had an increasing trajectory. Thus, this analysis provides preliminary data to support the hypothesis that those later diagnosed with an ASD are more likely to have scores above 70 throughout their development or have a point in time in which their scores increase into the clinical range and then remain there, as per an 'increasing' trajectory. Diagnostic follow-up beyond 36 months would determine if the three No ASD Diagnosis participants with stable high scores and the one No ASD Diagnosis participant with an increasing trajectory were eventually diagnosed with ASD.

CHAPTER 6

General Discussion

This is the first study to use the RCI statistic on longitudinal data of infants at biological risk for ASD and both the group and individual longitudinal analyses of the POEMS indicate that RCI has clinical utility. Although the use of the RCI statistic has not yet been explored in the autism literature, it has been applied to studies in neuropsychological research to identify measurable change in an individual's scores attributable to treatment response (Chelune, Naugle, Luders, Sedlak, & Awad, 1993), injury recovery (Iverson, Brooks, Collins, & Lovell, 2006), and disease progression (Hensel, Angermeyer, & Riedel-Heller, 2007). Given ASD is a developmental disorder whose symptoms appear to emerge progressively over time, it seems reasonable to use the RCI statistical model in a similar way to identify changes on a scored assessment such as the POEMS.

Recently, attempts have been made to classify the trajectories of infants at risk of ASD using regression procedures (Fountain et al., 2012; Lord et al., 2012). These latent class models provide estimates of the rate of change associated with different group memberships, and endeavor to identify variables contributing to the variation in trajectories (Singer & Willett, 2003). They are able to answer questions about what factors contribute to a given trajectory and which model is most predictive of a given outcome. What they fail to provide is a means by which significant change can be detected in practice on a case-by-case basis in order to make decisions around the necessity for more intensive assessment or intervention. That is, the detection of a significant change from one point in time to the next may give clinicians a 'red flag' that

warrants enhanced monitoring of a child's development. This may then lead to recommendations for earlier intervention that can facilitate a maximally positive developmental trajectory (Freeman & Perry, 2010).

As mentioned, one of the limitations of this study is the scarcity of data across assessment months. With a more robust dataset, the RCI could be applied to establish a normative understanding of what changes may be expected and when, within the development of high-risk infants. We could then use this information and compare it to the changes seen in infants who receive preventative early intervention to determine whether expected changes are delayed, or even expunged due to treatment effects. When compared to multiple regression procedures, the RCI statistic tends to be more apt to Type I errors (Maassen et al., 2009); however, it can be argued that false positives may be clinically reasonable when examining the utility of a screening measure such as the POEMS. In reality, though, conclusions made from this dataset provide preliminary evidence that the POEMS is reasonably sensitive in terms of detecting significant change. The POEMS also demonstrates good specificity, given that none of the individuals with stable low scores were diagnosed with ASD as of 36 months of age.

Although regression procedures are able to correct for alternative explanations of variance – such as regression towards the mean, demographic factors, and test-retest intervals – "the relative simplicity of simple regression and RCI methods outweighs the small benefit of multiple regression in most instances" (Levine et al., 2007, p. 670). In fact, some have argued that the simpler RCI calculation may be a more practical standardized approach to analyzing clinical change over time compared to overly complex regression procedures (Lambert & Ogles, 2009).

Two types of analysis were conducted using the RCI statistic: group and individual. The group analysis in Chapter 3 provided evidence that significant changes can be detected by the POEMS on a group level and that these changes correspond to the developmental timelines reported in the autism literature. Group comparisons provide insight into the 'normal' course of ASD and what might be expected in terms of an average trajectory. As aforementioned, the disadvantage to group analyses is that individual changes may be masked when data are averaged. This is especially true with smaller data sets where the variation of a few may be over- or under-represented in the group evaluation (Schoenberg et al., 2012).

The within-subject analysis in Chapters 4 and 5 provided further evidence that the POEMS is sensitive to significant changes on an individual level. It is important, particularly given the heterogeneity of ASD characteristics, for a measure to be able to quantify changes as meaningful and beyond the normal variation expected over the course of time (Heaton et al., 2001). The results reported in Chapter 4 support the use of POEMS as a screening measure; further studies are warranted to test the generalizability of these results to more robust datasets. It may also be appropriate to use the POEMS to explore what changes may be anticipated in children who are not at risk for an ASD or who may present with non-ASD specific developmental concerns.

Finally, further social validity research is recommended to gain an understanding of the low rate of response among families in this study. This measure is intended to be a reasonably straight-forward parent checklist; however, the 27.6% monthly response rate suggests that parents may find monthly reporting to be impractical or cumbersome. More indepth component analyses of the POEMS could also provide insight into which items

are most predictive of ASD diagnosis, leading to the potential for an abbreviated assessment measure that parents could feasibly complete on a more regular basis.

CHAPTER 7

Conclusion

In Analysis 1, the RCI statistic was applied on a group level. The results suggest that the POEMS is capable of detecting significant changes within pre-defined groups, and that the timeline of these changes is consistent with the current ASD literature. Significant relationships or trends in prognoses may be used to influence intervention decisions for infants and children presenting with similar characteristics at given ages.

In Analysis 2, meaningful change was detected on an individual level at various points within a child's assessment period. Unfortunately, given the large amount of missing data, it was difficult to make conclusions on a month-to-month basis for every child; however, the results provide acceptable evidence that the RCI statistic can be applied to an at-risk ASD population assessed through parent report.

If we were to make predictions based on the RCI patterns, the data from Analysis 3 would suggest that those with increasing trajectories and repeated high scores (i.e., above 70) are more likely to be diagnosed later with an ASD. In practice, if a child was identified as having a significant increasing RCI interval or consecutive high scores, they could be flagged as a higher-risk, and further observation may be recommended.

This study provides preliminary evidence for a simple statistical means by which individual change can be monitored in typical practice. Neither the POEMS nor the RCI calculation require intensive assessment, and pending studies exploring the acceptability of the POEMS within a low risk population, the potential for a reasonably simple, yet sensitive and specific screening measure is exciting. A hypothetical example of an application of these results follows: parents could input the scores into a computer

program which calculates RCI values and stores the results from repeated measures. These results could be taken to clinicians (or sent remotely) to monitor for significant changes on a month-to-month basis. This program could also potentially alert clinicians to repeated high overall scores and, coupled with significant RCI intervals, they may be able to more efficiently identify those children who warrant more specific monitoring. Further research with low risk infants will provide more accurate and standardized reliability measures for each of the months during infant development. These would be required for a population-based application of the RCI procedure with the POEMS assessment. Without a similar application of the RCI statistic to low risk populations, conclusions are limited to infants and toddlers genetically at risk for an ASD diagnosis.

References

Atkins, D.C., Bedics, J.D., McGlinchey, J.B., and Beauchaine, T.P. (2005). Assessing clinical significance: Does it matter which method we use? *Journal of Consulting and Clinical Psychology, 73*(5), 982-989. doi:10.1037/0022-006X.73.5.982

Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E., and Rutter, M. (1995). Autism as a strongly genetic disorder: Evidence from a British twin study. *Psychological Medicine, 25*(1), 63-77.

Baruch, Y. (1999). Response rate in academic studies – A comparative analysis. *Human Relations, 52*(4), 421-438.

Bricker, D. and Squires, J. (1999). The ages and stages questionnaires. Baltimore, MD: Paul H. Brookes.

Bryson, S.E., Zwaigenbaum, L., Brian, J., Roberts, W., Szatmari, P., Rombough, V., and McDermott, C. (2007). A prospective case series of high-risk infants who developed autism. *Journal of Autism and Developmental Disorders, 37*, 12-24. doi: 10.1007/s10803-006-0328-2.

Chawarska, K., Paul, R., Klin, A., Hannigen, S., Dichtel, L.E., and Volkmar, F. (2007). Parental recognition of developmental problems in toddlers with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37*, 62-72. doi: 10.1007/s10803-006-0330-8.

Chelune, G.J., Naugle, R.I., Luders, H., Sedlak, J., and Awad, I.A. (1993). Individual change after epilepsy surgery: Practice effects and base-rate information. *Neuropsychology, 7*(1), 41-52.

Christensen, L.C. and Mendoza, J. (1986). A method of assessing change in single-subject designs: An alteration of the RC index. *Behavior Therapy, 17*, 305-308.

Crawford, J.R. and Howell, D.C. (1998). Regression equations in clinical neuropsychology: An evaluation of statistical methods for comparing predicted and obtained scores. *Journal of Clinical and Experimental Neuropsychology, 20*(5), 755-762.

Feldman, M.A., Ward, R.A., Savona, D., Regehr, K., Parker, K., Hudson, M., Penning, H., and Holden, J.J. (2012). Development and initial validation of a parent report measure of the behavioral development of infants at risk for autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42*, 13-22. doi: 10.1007/s10803-011-1208-y.

Fountain, C., Winter, A.S., and Bearman, P.S. (2012). Six developmental trajectories characterize children with autism. *Pediatrics, 129*(5), 1-9. doi: 10.1542/peds.2011-1601

Freeman, N. and Perry, A. (2010). Outcomes of intensive behavioural intervention in the Toronto Preschool Autism Service. *Journal on Developmental Disabilities, 16*(2), 17-32.

Gamliel, I., Yirmiya, N., and Sigman, M. (2007). The development of young siblings of children with autism from 4 to 54 months. *Journal of Autism and Developmental Disorders, 37*, 171-183. doi: 10.1007/s10803-006-0341-5

Gamliel, I., Yirmiya, N., Jaffe, D.H., Manor, O., and Sigman, M. (2009). Developmental trajectories in siblings of children with autism: Cognition and language from 4

months to 7 years. Journal of Autism and Developmental Disorders, 39, 1131-1144. doi: 10.1007/s10803-009-0727-2

Heaton, R.K., Temkin, N., Dikmen, S., Avitable, N., Taylor, M.J., Marcotte, T.D., and Grant, I. (2001). Detecting change: A comparison of three neuropsychological methods using normal and clinical samples. *Archives of Clinical Neuropsychology, 16,* 75-91.

Hensel, A., Angermeyer, M.C., and Riedel-Heller, S.G. (2007). Measuring cognitive change in older adults. Do reliable change indicies of the SIDAM predict dementia? *Journal of Neurology, 254,* 1359-1365. doi:10.1007/s00415-007-0549-0.

Hsu, L.M. (1989). Reliable changes in psychotherapy: Taking into account regression toward the mean. *Behavioral Assessment, 11*, 459-467.

Iverson, G.L., Brooks, B.L., Collins, M.W., and Lovell, M.R. (2006). Tracking neuropsychological recovery following concussion in sport. *Brain Injury, 20*(3), 245-252.

Jacobson, N.S., Follette, W.C., and Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy, 15,* 336-352.

Jacobson, N.S. and Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12-19.

Lambert, M.L. and Ogles, B.M. (2009). Using clinical significance in psychotherapy outcome research: The need for a common procedure and validity data. *Psychotherapy Research, 19*(4-5), 493-501.

Landa, R. and Garrett-Mayer, E. (2006) Development in infants with autism spectrum disorders: A prospective study. *Journal of Child Psychology and Psychiatry, 47*(6), 629-638.

Levine, A.J., Hinkin, C.H., Miller, E.N., Becker, J.T., Selnes, O.A., and Cohen, B.A. (2007). The generalizability of neurocognitive test/retest data derived from a nonclinical sample for detecting change among two HIV+ cohorts. *Journal of Clinical Experimental Neuropsychology, 29*(6), 669-678. doi: 10.1080/13803390600920471

Lord, C., Luyster, R., Guthrie, W., & Pickles, A. (2012, April 16). Patterns of developmental trajectories in toddlers with autism spectrum disorder. *Journal of Consulting and Clinical Psychology*. Advance online publication. doi: 10.1037/a0027214

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P., and Rutter, M. (2000). The Autism Diagnostic Observation Schedule - Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*, 205–223. doi:10.1023/A:1005592401947

Lord, C., Rutter, M., and 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.

Maassen, G.H. (2004). The standard error in the Jacobson and Truax reliable change index: The classical approach to the assessment of reliable change. *Journal of the*

International Neuropsychological Society, 10, 888-893. doi:10.10170S1355617704106097

Maassen, G.H., Bossema, E., and Brand, N. (2009). Reliable change and practice effects: Outcomes of various indices compared. *Journal of Clinical and Experimental Neuropsychology, 31*(3), 339-352.

Mullen, E.M. (1995). *Mullen scales of early learning.* Circle Pines, MN: American Guidance Services, Inc.

Ozonoff, S., Iosif, A.M., Baguio, F., Cook, I.C., Moore Hill, M., Hutman, T., Rogers, S.J., Roxga, A., Sangha, S., Sigman, M., Steinfield, M.B., and Young, G.S. (2010). A prospective study of the emergence of early behavioral signs of autism. *Journal of the American Academy of Child and Adolescent Psychiatry, 49*(3), 258-270.

Ozonoff, S., Young, G.S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., Bryson, S., Carver, L.J., Constantino, J.N., Dobkins, K., Hutman, T., Iverson, J.M., Landa, R., Rogers, S.J., Sigman, M., and Stone, W.L. (2011). Recurrence rise for autism spectrum disorders: A baby siblings research consortium study. *Pediatrics, 128*(3), e488-e495. doi:10.1542/peds.2010-2825.

Schoenberg, M.R., Rinehardt, E., Duff, K., Mattingly, M., Bharucha, K.J., Scott, J.G. (2012). Assessing reliable change using the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) for patients with Parkinson's disease undergoing deep brain stimulation (DBS) surgery. *The Clinical Neuropsychologist, 26*(2), 255-270.

Singer, J.D. and Willett, J.B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* Toronto: Oxford University Press.

Speer, D.C. (1992). Clinically significant change: Jacobson and Truax(1991) revisited. Journal of Consulting and Clinical Psychology, 60, 402-408.

Wetherby, A.M., Woods, J., Allen, L., Cleary, J., Dickinson, H., and Lord, C. (2004). Early indicators of autism spectrum disorders in the second year of life. *Journal of Autism and Developmental Disorders, 34*(5), 473-493.

Young, R.L., Brewer, N., and Pattison, C. (2003). Parental identification of early behavioural abnormalities in children with autistic disorder. *Autism, 7*(2), 125-143.

Zwaigenbaum, L., Bryson, S., Lord, C., Rogers, S., Carter, A., Carver, L., Chawarska, K., Constantino, J., Dawson, G., Dobkins, K., Fein, D., Iverson, J., Klin, A., Landa, R., Messinger, D., Ozonoff, S., Sigman, M., Stone, W., Tager-Flusberg, H., and Yirmiya, N. (2009). Clinical assessment and management of toddlers with suspected autism spectrum disorder: Insights from studies of high-risk infants. *Pediatrics, 123*, 1383-1391. doi:10.1542/peds.2008-1606.

Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J., and Szatmari, P. (2005). Behavioural manifestations of autism in the first year of life. *International Journal of Developmental Neruoscience, 23*, 143-152.

Appendix

Age (months)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
Stable Low Pattern																																				
339					-	-		-														-					-	-	-	-	-		-	-	-	
621											-	-		-			-	-			-						-	-	-	-	-	-	-			-
1507						-												-	-	-	-		-	-	-	-	-			-						
1712																												-	-	-	-	-	-	-	-	
1741		-				-											-		-	-		-	-	-	-	-			-						
1856																											-						-		-	-
1941					-	-			-	-		-	-	-		-	-			-	-		-	-	-										
2043			-						-			-			-		-	-	-		-	-		-											
2307			-			-			-			-				-	-	-	-			-	-	-			-			-			-			-
2340				-	-	-			-			-					-		-	-		-		-		-				-				-		
2376																				-	-	-	-	-	-	-		-	-	-	-		-	-	-	
2417													-	-	-		-	-																		
2422		-		-			-			-	-	-		-		-		-	-			-		-			-		-				-			-
3273							-	-			-	-	-		-			-																		
3305					-	-	-	-		-	-	-	-	-	-		-																			
3327		-	-	-	-	-	-	-	-	-	-	-	-	-										-								-			-
3353		-	-	-	-		-	-		-	-	-	-																						
3360		-			-			-		-														-												
3667															-		-		-	-	-	-	-				-									-
3896		-	-	-		-		-	-	-										-					-	-	-			-			-			-
4232													-	-	-	-	-			-							-									
4490							-	-	-	-					-		-							-												
4566	-		-					-															-													
4572			-	-	-			-		-		-				-			-					-			-			-						
4726					-	-	-		-		-																			-			-		-	
4767		-	-		-	-	-											-												-		-				
5083	-	-		-	-	-				-										-		-		-			-			-			-			-
5649	-	-		-										-						-		-		-			-				-					-
6151							-	-		-	-	-		-		-																				
6238				-	-	-		-							-			-		-		-		-			-			-						
6644											-		-			-	-		-			-				-				-						
6666								-		-														-												-
6859																-			-	-		-		-												-
6929																-		-		-		-			-		-									-
6957														-		-		-		-		-	-													-
6976							-			-		-			-						-				-											-
7079				-		-										-		-		-		-		-		-										-
7245												-		-		-																				-
7324			-							-		-		-		-		-																		-
7718			-		-			-		-		-		-		-																				-
7936												-		-		-				-		-		-		-					-			-			-
8365				-		-		-		-		-		-			-	-		-		-						-		-			-			-
26148				-	-			-																-			-									-
35849									-	-		-		-				-				-		-			-			-				-		

Figure A1. Visual representation of participants with a stable low pattern of scores, using a cut-off score of 70.

Legend:

Score ≥ 70 on the POEMS

Score < 70 on the POEMS


Significant positive RCI interval (i.e., > 1.96)


Significant negative RCI interval (i.e., < -1.96)


Individual not diagnosed with an ASD as of 36 months of age


Individual diagnosed with an ASD as of 36 months of age


Individual with a lost or provisional ASD diagnosis as of 36 months of age

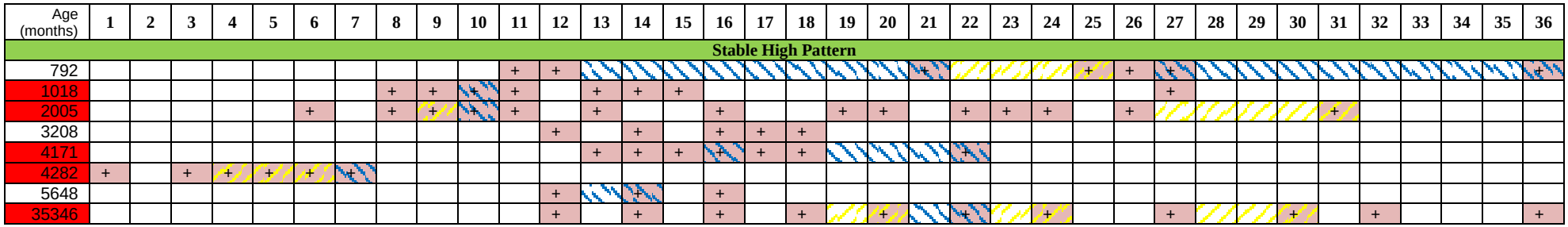

Figure A2. Visual representation of participants with a stable high pattern of scores, using a cut-off score of 70.

Legend:

+ Score ≥ 70 on the POEMS

\- Score < 70 on the POEMS

Significant positive RCI interval (i.e., > 1.96)

Significant negative RCI interval (i.e., < -1.96)

Individual not diagnosed with an ASD as of 36 months of age

Individual diagnosed with an ASD as of 36 months of age

Individual with a lost or provisional ASD diagnosis as of 36 months of age

Age (months)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
Decreasing Pattern																																				
58										+	+	+						-		-		-		-							-	-			-	
353													+											+		+	+	+	+		-		-	-		
1123										+		+	+		+	+	-	-						-	-		-				-		-			-
1797										+	+	+		+		-	-	-	-		-	-	-	-	-	-		-		-	-		-			-
1857						+	-		-	-	-	-	-	-	-	-	-		-	-		-	-	-	-	-	-	-	-	-	-					
2218																								+							+		+		-	-
2346													+	-						-			-	-	-		-	-		-			-			-
2552												+			-	-	-			-	-			-			-									
2690																		+			+			-	-	-		-	-	-	-	-	-	-	-	
2714										+								+			-															
3084										+		+	+	+	+		+			-	-	-													
3163										+	+	+	+	+	+	+		+	+	+		+		+						+			-			
6679											+					-				-		-					-		-		-		-			-
6944						+						+		+		+			-	-		-														-
7937										+		+							-	-		-		-			-			-			-			-

Figure A3. Visual representation of participants with a decreasing pattern of scores, using a cut-off score of 70.

Legend:

+ Score ≥ 70 on the POEMS

\- Score < 70 on the POEMS

Significant positive RCI interval (i.e., > 1.96)

Significant negative RCI interval (i.e., < -1.96)

Individual not diagnosed with an ASD as of 36 months of age

Individual diagnosed with an ASD as of 36 months of age

Individual with a lost or provisional ASD diagnosis as of 36 months of age

Age (months)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
Increasing Pattern																																				
1084										-	-	+		+	+		+	+	+	+	+		+		+	+	+	+								+
2731					-		-				+			+	+		+	+					+													
4343	-	-			-	-		-		-			+																							
6548					-					-								+		+				+	+										-	

Figure A4. Visual representation of participants with an increasing pattern of scores, using a cut-off score of 70.

Legend:

+	Score ≥ 70 on the POEMS
-	Score < 70 on the POEMS
	Significant positive RCI interval (i.e., > 1.96)
	Significant negative RCI interval (i.e., < -1.96)
	Individual not diagnosed with an ASD as of 36 months of age
	Individual diagnosed with an ASD as of 36 months of age
	Individual with a lost or provisional ASD diagnosis as of 36 months of age

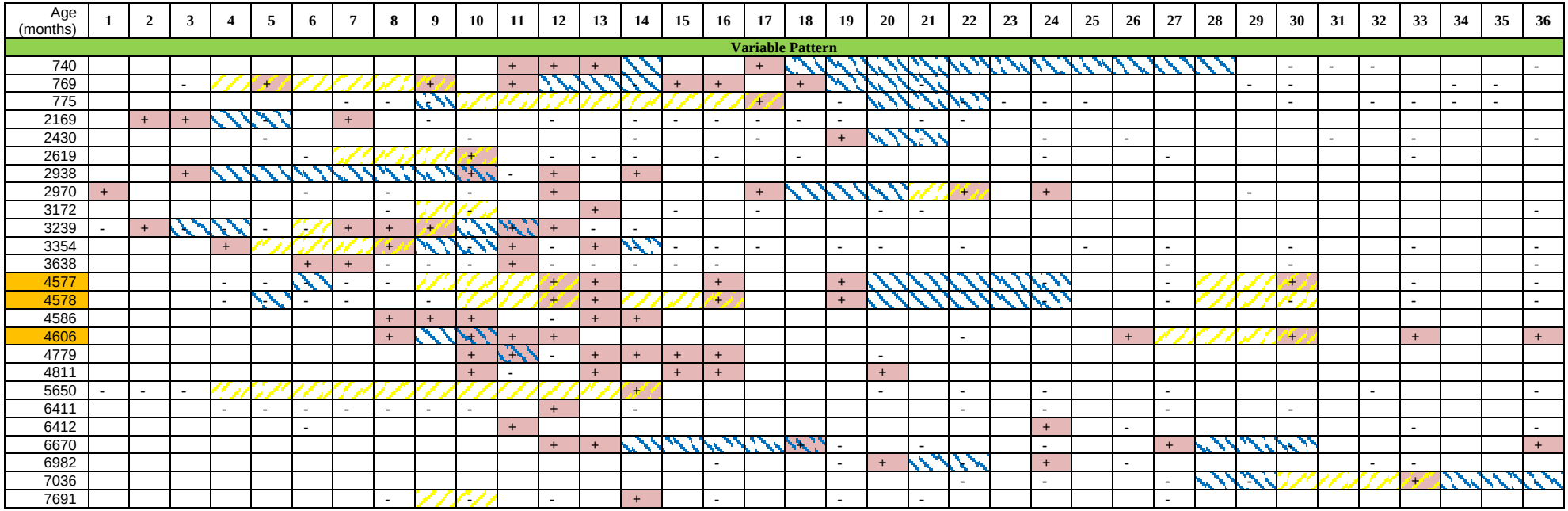

Figure A5. Visual representation of participants with a variable pattern of scores, using a cut-off score of 70.

Legend:

+ Score ≥ 70 on the POEMS

\- Score < 70 on the POEMS

Significant positive RCI interval (i.e., > 1.96)

Significant negative RCI interval (i.e., < -1.96)

Individual not diagnosed with an ASD as of 36 months of age

Individual diagnosed with an ASD as of 36 months of age

Individual with a lost or provisional ASD diagnosis as of 36 months of age