Research in Autism Spectrum Disorders 5 (2011) 592–603

Contents lists available at ScienceDirect

Research in Autism Spectrum Disorders

Journal homepage: http://ees.elsevier.com/RASD/default.asp

Predictors of outcome for children receiving intensive behavioral intervention in a large, community-based program☆

Adrienne Perry [a,*], Anne Cummings [b], Jennifer Dunn Geier [c], Nancy L. Freeman [d], Susan Hughes [e], Tom Managhan [f], Jo-Ann Reitzel [g], Janis Williams [h]

[a] *Department of Psychology, York University, 4700 Keele Street, Toronto, ON, Canada M3J 1P3*
[b] *Kinark Child and Family Services, Markham, ON, Canada*
[c] *Children's Hospital of Eastern Ontario, Ottawa, ON, Canada*
[d] *Surrey Place Centre, Toronto, ON, Canada*
[e] *Pathways for Children and Youth, Kingston, ON, Canada*
[f] *Child Care Resources, Sudbury, ON, Canada*
[g] *Hamilton-Niagara Early Autism Initiative, Hamilton, ON, Canada*
[h] *ErinoakKids, Mississauga, ON, Canada*

ARTICLE INFO

Article history:
Received 3 July 2010
Accepted 12 July 2010

Keywords:
Early Intensive Behavioral Intervention (EIBI)
Predictors
Outcomes
Effectiveness Study

ABSTRACT

This study reports on predictors of outcome in 332 children, aged 2–7 years, enrolled in the community-based Intensive Behavioral Intervention (IBI) program in Ontario, Canada. Data documenting children's progress were reported in an earlier publication (Perry et al., 2008). The present paper explores the degree to which four predictors (measured at intake to IBI) are related to children's outcomes: age at entry, IQ, adaptive scores, and autism severity. Outcome variables examined include: post-treatment scores for: autism severity, adaptive behavior, cognitive level, rate of development in IBI, and categorical progress/ outcomes (seven subgroups). All four types of predictors were related to children's outcomes, although initial cognitive level was the strongest predictor. In addition, two subgroups of the sample are examined further. Children who were most successful in the program and achieved average functioning had higher developmental levels at intake, were considerably younger than the rest of the children, and were in treatment longer than children in other outcome categories. Children who were least successful in the program and made essentially no progress did not differ appreciably from the remainder of the group. Implications of these results for decision-making are discussed.

1. Introduction

1.1. Intensive Behavioral Intervention

Early Intensive Behavioral Intervention (EIBI or, as typically designated in Ontario, IBI) is an intensive application of the principles of Applied Behavior Analysis designed for young children with autism. The intervention is comprehensive in scope and is typically provided in a 1-to-1 format (at least initially) for 20–40 h per week for about 2 years. It is intended to change

☆ Portions of this study were presented in poster or presentation formats at: the Association on Behavior Analysis Autism Conference in February 2009 in Jacksonville, FL; the Association for Behavior Analysis in May 2007 in San Diego, CA; the Society for Research in Child Development in April 2007 in Boston, MA; the Ontario Association on Developmental Disabilities in April 2007 in Barrie, ON; the Association on Behavior Analysis Autism Conference in February 2007 in Boston, MA; and at the Ontario Association for Behaviour Analysis in November 2006 in Markham, ON.

* Corresponding author. Tel.: +1 416 736 5115; fax: +1 416 736 5814.
E-mail address: perry@yorku.ca (A. Perry).

1750-9467/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.rasd.2010.07.003

children's developmental trajectory such that they can transition to learning in a more typical way in school (Lovaas, 1987; McEachin, Smith, & Lovaas, 1993). The behavioral nature of the intervention, not just its intensity, appears to be linked to superior outcomes (Eikeseth, Smith, Jahr, & Eldevik, 2002; Howard, Sparkman, Cohen, Green, & Stanislaw, 2005). Although IBI is now widely regarded as best practice for children with autism, outcomes are decidedly variable with good outcomes for about half the children at best. Although this is remarkable compared to the results of autism intervention prior to the widespread use of IBI, clearly IBI is not a panacea for all children. Research, to date, has not been able to account very well for this heterogeneity in outcomes.

Several investigators have focused on examining more closely which children make very dramatic gains and achieve "recovery", "best outcomes" or average functioning (e.g., Eikeseth, Smith, Jahr, & Eldevik, 2007; Lovaas, 1987). However, it is also important to investigate further which children do not benefit, which has rarely been done. As noted by Reichow and Wolery (2009), "...it is imperative children not responding to intervention are identified early so additional and/or different treatments can begin." (p. 39). Considerations regarding resource allocation in the face of limited budgets and active advocacy efforts loom large. It is clear that greater understanding is urgently needed regarding which children will benefit, to what degree, and why.

Questions regarding moderators or predictors of outcome have been increasingly addressed in the literature recently. Factors that have been theorized to be related to the heterogeneity in outcome include child, family, and treatment characteristics. The majority of research, including the present study, focuses on child characteristics, such as cognitive and adaptive levels, age at IBI onset, and severity of autism symptomatology. However, it should be noted that some consideration is also being given recently to treatment characteristics (Granpeesheh, Dixon, Tarbox, Kaplan, & Wilke, 2009; Koudys & Perry, 2010; Makrygianni & Reed, 2010; Reichow & Wolery, 2009) and family variables (Remington et al., 2007; Shine & Perry, 2010; Solish, 2010).

1.2. Child factors related to outcome

There are good reasons to assume that starting IBI younger might be beneficial and this could be argued from several different theoretical perspectives (e.g., behavioral theory, neural plasticity, and developmental theory). A number of studies have reported on the question of children's age when treatment commences and whether the widespread "earlier the better" belief is empirically borne out. Results are somewhat equivocal, in fact. Studies which have wide age ranges and have divided their sample into younger versus older subgroups have typically found that younger (however defined; under 4 or under 5) children are more likely to show better outcomes than older children (Fenske, Zalenski, Krantz, & McClannahan, 1985; Granpeesheh et al., 2009; Harris & Handleman, 2000) and the same is true of Anderson, Avery, DiPietro, Edwards, and Christian (1987) based on the individual data in the paper. However, other studies (typically using correlational type statistics) have looked for and not found a relationship with age. This is true in very young samples (2–3.5 years; Hayward, Eikeseth, Gale, & Morgan, 2009; Lovaas, 1987) and also one somewhat older sample (4–7 years; Eikeseth et al., 2002, 2007). Small samples with restricted age ranges may preclude correlations from emerging as significant in these studies. Other studies appear not to have examined the question of age at all, or at least do not report such analyses. Thus, as noted by Matson and Smith (2008), the precise nature and power of age as a predictor remains less clear than one might think.

Children's initial cognitive level has also been examined as a predictor of outcome in a number of studies. Initial IQ has often been reported to be moderately to highly correlate with outcomes (Eikeseth et al., 2002, 2007; Harris & Handleman, 2000; Hayward, Eikeseth, Gale, & Morgan, 2009; Sallows & Graupner, 2005). However, this is likely the case regardless of treatment as exemplified by results reported by Gabriels, Hill, Pierce, Rogers, and Wehner (2001) for more generic treatment and the Eikeseth et al. (2007) eclectic comparison group. Although most studies find initial IQ related to outcome, a few studies have examined this relationship and found it not to be significant (Birnbrauer & Leach, 1993; Cohen, Amerine-Dickens, & Smith, 2006; Smith, Groen, & Wynn, 2000).

Although adaptive behavior measures are frequently used as outcome measures, less attention has been paid to initial adaptive levels as predictors of outcome but there is some evidence that children with better adaptive skills tend to have better outcomes (Remington et al., 2007; Sallows & Graupner, 2005). Since cognitive and adaptive levels are correlated, one might expect similar results, but because of the interesting relationship between cognitive and adaptive scores at different cognitive levels (Perry, Flanagan, Dunn Geier, & Freeman, 2009), it might be worthwhile examining adaptive levels as a predictor of outcomes in a heterogeneous sample.

Surprisingly, severity of autism symptoms or diagnosis (Autism versus PDD-NOS) has rarely been included in predictor analyses or even as an outcome measure (Matson, 2007; Matson & Smith, 2008). Sallows and Graupner (2005) showed that lower pre-treatment ADI-R scores (together with higher IQ and more rapid early skill mastery) were quite accurate in predicting which children would be "rapid responders". Remington et al. (2007), on the other hand, found that their good responders initially had more severe autism symptoms. Smith et al. (2000) found that their Autism subgroup showed only a 4-point IQ gain versus a 16-point IQ gain in the PDD-NOS subgroup.

In summary, for all four child factors (initial age, cognitive level, adaptive skills, and diagnostic severity), results are not completely consistent and further research is needed. However, most of these studies are quite small (e.g., 10–25 children) and, thus, power limitations preclude many analyses which might shed light on why some children do better than others. The small samples likely result in Type II error (i.e., interesting findings may have been missed). Also, possible sampling issues could result in spurious findings as well, as some of the samples are quite restricted in range (e.g., the children in Lovaas [1987] were all less than 3.5; the children in Eikeseth et al. (2002) all had IQs over 50). Therefore, systematic and meta-analytic studies of this body of literature have started to appear, which seek to provide greater clarity by systematically combining the existing literature.

1.3. Meta-analyses

Howlin, Magiati, and Charman (2009) conducted a systematic review, identifying 11 studies which met their criteria. They concluded that IBI resulted in improved outcomes at a group level but that there was considerable individual variability. In terms of predictors, they concluded, based on their descriptive summary of the studies, that initial IQ and receptive language were important predictors of IQ at follow-up but that initial age and diagnosis were unrelated to outcomes. Eldevik et al. (2009) used a more specific criterion for inclusion of studies which resulted in examining nine controlled studies. They calculated effect sizes for IQ and adaptive behavior and conducted a meta-analysis of individual children's data from all studies combined. Results indicated a large effect size for IQ (1.10) and medium effect size for adaptive behavior (.66). However, Eldevik et al. (2009) did not conduct predictor analyses.

Reichow and Wolery (2009) conducted a systematic review of 13 studies, comprising a total of 251 children receiving IBI, which included ratings of methodological rigor, participant characteristics, intervention quantity and quality. They computed both mean change effect sizes (comparing pre- and post-treatment scores) and mean difference effect sizes (comparing the treatment group with a control or comparison group) for IQ, adaptive behavior, expressive language, and receptive language, and conducted a meta-analysis, using corrections for small sample sizes and other procedures to ensure conservative conclusions. Of greatest relevance to the present paper, Reichow and Wolery (2009) also conducted moderator analyses, examining the studies' weighted effect sizes for IQ (the main dependent variable) as a function of several child and treatment variables. Model of supervision was the only significant moderator across studies with the UCLA-model supervision associated with superior effect sizes. Other treatment variables reflecting amount of treatment (intensity and duration) were unrelated to child outcome, at least within the ranges studied. Neither of the child characteristics analyzed, that is, pre-treatment age and pre-treatment IQ, was found to significantly affect children's outcomes (post-treatment IQ). Diagnostic severity was not examined because most studies did not include it.

Most recently, Makrygianni and Reed (2010) conducted a meta-analysis of 14 studies, again examining both mean change effect sizes and mean difference effect sizes, but using slightly different procedures, for IQ, language, and adaptive behavior. They also examined the relationship of effect sizes with program variables (intensity, duration, and parent training) and child characteristics (age at intake, initial levels of cognitive, language, and adaptive skills), controlling for methodological quality of the studies. Their results were not necessarily consistent with those of Reichow and Wolery (2009). For example, they found intensity of treatment to be significantly related to several of the effect sizes. Their analyses for age at intake revealed moderately high negative correlations for age at intake with several of their effect size variables. Interestingly, a scatter plot of effect sizes for all variables by age at intake suggested that studies with children who began very early (roughly under 3) tended to have more uniformly large effect sizes whereas studies with children beginning later had more variable effect sizes. In terms of developmental level as a predictor, Makrygianni and Reed (2010) found that intellectual ability at intake was not correlated with effect sizes but that intake IQ was very highly correlated with outcome IQ. Initial adaptive behavior was related to effect sizes for language outcomes and adaptive behavior outcomes. Autism severity was not examined in this study (again because of limited data).

1.4. Current study

The purpose of the present paper was to report on our analyses of predictors in a very large ($n = 332$) and diverse sample of children receiving IBI. The sample is larger than the combined sample used by the authors of the meta-analyses and it is very diverse in terms of children's developmental and diagnostic characteristics as well as family background. As such, there is sufficient range and statistical power to explore prediction of outcomes more fully than has been possible before. Thus, we hoped that these analyses would help to shed some light on the debate in the literature regarding the relative importance of different child predictor variables. In particular, we set out to examine the relationship of children's outcomes to age at entry, initial cognitive level, adaptive functioning, and diagnostic severity. In addition, we examined correlates and predictors more closely in two subgroups of children at the two extremes of outcome classification: those with outcomes in the average range and those with poor outcomes.

Data for this paper are drawn from the effectiveness study of the Ontario province-wide IBI initiative, a large, community-based publicly funded IBI program. The earlier paper (Perry et al., 2008) addressed the question of whether children showed statistically significant and clinically significant improvement on developmental and diagnostic measures and describes the range of progress/outcomes seen in children. Briefly, results indicated that, overall, children improved significantly across all measures from entry (T1) to exit (T2) but there was substantial heterogeneity. More specifically, they demonstrated significantly milder autistic symptomatology at T2. Adaptive behavior age equivalents increased substantially in all areas, but standard scores changed only modestly (higher for Socialization and Communication but lower for Daily Living Skills). Cognitive level (when available) increased significantly. Children's rate of development during IBI was roughly double what it had been at T1. Based on a combination of all available measures, children were classified as falling into one of seven categories of progress/outcome: Average, Substantially Improved, Clinically Significantly Improved, Less Autistic, Minimally Improved, No change, and Worse. A subgroup of children who were more similar to children from model programs (younger with milder developmental delays) had similar outcomes to those reported in efficacy studies.

2. Methods

2.1. Participants

Psychological assessment file data were available for a total of 332 children (83% boys) with an entry assessment (T1) and another assessment (T2, usually at exit). More details about the sample may be found in the earlier report (Perry et al., 2008). The children's initial status on all diagnostic and developmental variables is shown in Table 1, as well as their age (which was, on average, about 4.5) and the duration of intervention (which was, on average, 18 months).

2.2. Measures

Assessment measures for the children included severity of autism, cognitive level, and adaptive level, using standardized, well-accepted measures for this population. *Autism Severity*. Autism severity was measured using the *Childhood Autism Rating Scale* (*CARS*; Schopler, Reichler, & Renner, 1988), which results in a Total Score (higher scores indicating greater severity) and a trichotomous classification: not autism, mild/moderate autism, and severe autism. DSM-IV diagnoses (Autism, PDD-NOS, or broader ASD/PDD) were also available. *Adaptive Levels*. Adaptive Behavior levels were assessed using the *Vineland Adaptive Behavior Scales* (*VABS*; Sparrow, Balla, & Cicchetti, 1984). We report Standard Scores and Age Equivalents for the three principal domains, Communication, Daily Living Skills, and Socialization, as well as the Motor domain for children under 6 years. There is also an overall Adaptive Behavior Composite (ABC) score which, in this report, is based on a mean of the three principal domains (not including Motor). *Cognitive Level*. A cognitive measure was available for 163 children (49%) at intake and 265 (80%) at exit, based on various tests. Lack of a cognitive measure on file was associated with resource limitations in the programs during start-up and not to child variables. Independent *t*-tests comparing children with and without an initial IQ score indicated no difference in age at entry, but those without an IQ scored significantly lower on the ABC and significantly higher on the CARS, although neither difference was large in magnitude. *Developmental Rate*. Estimates of Developmental Rates were based on the VABS ABC Age Equivalent scores. *Initial Developmental Rate* was calculated by dividing this score by the child's age (similar to ratio IQ). *Developmental Rate During IBI* was calculated by taking the difference between the exit and intake age equivalents divided by the time interval between them (i.e., duration of IBI). See Perry et al. (2008) for further details regarding the measures.

Table 1
Developmental and Diagnostic Status of Participants at Intake ($n = 332$).

	M (*SD*)	Range
Autism Severity ($n = 304$)		
CARS Total score	36.58 (5.49)	22–53
VABS Standard Scores ($n = 295$)		
Communication	53.98 (11.05)	24–105
Daily Living	54.17 (11.92)	19–111
Socialization	56.95 (7.84)	43–92
Motor ($n = 278$)	61.85 (14.93)	19–113
ABC	54.94 (9.13)	33–97
VABS Age Equivalents (months) ($n = 297$)		
Communication	15.58 (8.82)	1–62
Daily Living	20.83 (6.60)	8–50
Socialization	13.72 (6.40)	2–45
Motor ($n = 285$)	28.87 (8.70)	11–71
ABC	16.76 (6.42)	6–42
Cognitive level ($n = 151$)		
IQ Estimate	45.50 (19.24)	11–96
MA Estimate (months)	22.92 (10.96)	3–60
Rate of Development ($n = 297$)		
Initial rate (ratio ABC age/CA)	.32 (.12)	.10–.86
Program variables ($n = 332$)		
Age at Intake (months)	53.56 (12.60)	20–86
Duration (months)	18.43 (8.38)	4–47

Table 2
Correlations of outcome variables with predictors.

Outcome variables at T2	Predictors at T1			
	Age	IQ	ABC	CARS
CARS Total	.18	−.42	−.51	.52
VABS Standard Scores				
Communication	−.38	.64	.71	−.33
Daily Living	−.44	.56	.74	−.32
Socialization	−.40	.65	.71	−.30
Motor	−.26	.51	.53	−.27
ABC	−.43	.67	.77	−.34
VABS Age Equivalents				
Communication	−.17	.65	.60	−.34
Daily Living	−.09[a]	.54	.58	−.29
Socialization	−.20	.63	.53	−.29
Motor	−.12[a]	.48	.43	−.22
ABC	−.17	.67	.62	−.33
Cognitive Level				
IQ estimate	−.39	.73	.72	−.43
MA estimate	−.21	.76	.62	−.44
Rate of Development	.11[a]	.50	.26	−.18

Note: n varies. For VABS scores and CARS at T1 and T2, 273 to 279, except for motor (167 for AEs and 138 for SSs). For cognitive at T1 $n = 163$, for T2 $n = 265$. All correlations significant at $p < .01$ except those with [a] superscript are *ns*.

3. Results

3.1. Predictors of outcome in the whole sample

3.1.1. Age at entry

Do children beginning IBI earlier show greater gains? This question was examined in several different ways. First, as shown in Table 2, the adaptive and cognitive variables at outcome tended to be significantly negatively correlated with age at entry (i.e., children who started IBI younger tended to score higher at discharge), and younger age at entry was correlated with milder autism severity at exit. These correlations are small to medium in magnitude.

Second, data were compared using independent t-tests for two subgroups who were Younger and Older for the absolute levels of the exit assessment scores. Results are shown in Table 3. Children who were under 4 years of age at program entry (only about one-third of the sample) had significantly lower CARS scores at exit and significantly higher scores on most other measures compared to older children. The magnitude of the differences was quite large for standard scores (e.g., 20 points for

Table 3
Scores on outcome variables for younger versus older age groups.

	Younger age <48 ($n = 97$)	Older age ≥48 ($n = 199$)	t	p
CARS				
Total Score	30.54 (5.63)	32.30 (5.49)	2.56	.01
VABS Standard Scores	($n = 93$)	($n = 209$)		
Communication	70.71 (24.23)	54.67 (18.99)	−6.21	< .001
Daily Living	60.61 (17.58)	46.62 (16.35)	−6.71	< .001
Socialization	65.84 (15.16)	56.52 (9.85)	−6.37	< .001
Motor ($n = 84$ and 77)	72.68 (21.42)	63.65 (18.60)	−2.85	.005
ABC	65.72 (17.75)	52.60 (13.88)	−6.94	< .001
VABS Age Equivalents	($n = 95$)	($n = 216$)		
Communication	38.25 (21.15)	30.16 (18.81)	−3.36	.001
Daily Living	34.92 (14.02)	31.89 (11.56)	−2.00	.05
Socialization	29.26 (17.15)	22.83 (11.97)	−3.80	< .001
Motor ($n = 85$ and 108)	44.89 (13.92)	42.85 (14.69)	−0.98	*ns*
ABC	34.14 (16.27)	28.32 (13.07)	−3.36	.001
Cognitive	($n = 78$)	($n = 177$)		
FS IQ Estimate	71.42 (30.53)	51.00 (24.70)	−5.66	< .001
MA Estimate	45.23 (18.48)	37.80 (17.65)	−3.04	.003
Rate of Development	($n = 83$)	($n = 195$)		
	.68 (.58)	.80 (.83)	1.17	*ns*

Table 4
Categories of outcome for younger versus older children and mean intake age of each outcome group.

	n	Younger Age <48 (n = 95)	Older Age ≥48 (n = 201)	Mean age at entry[a] M (SD)
(1) Average Functioning	32	24 (25.3%)	8 (4.0%)	41.91 (12.54)
(2) Substantial Improvement	43	13 (13.7%)	30 (14.9%)	53.16 (12.17)
(3) Clinically Sig. Improvement	90	22 (23.2%)	68 (33.8%)	55.44 (11.02)
(4) Less Autistic	31	11 (11.6%)	20 (10.0%)	54.45 (13.56)
(5) Minimal Improvement	25	4 (4.2%)	21 (10.4%)	56.08 (12.85)
(6) No change	55	15 (15.8%)	40 (19.9%)	54.27 (11.90)
(7) Worse	20	6 (6.3%)	14 (7.0%)	55.20 (11.91)
Total	296			53.36 (12.58)

[a] Post hoc tests indicate Group 1 significantly different from all other groups, none of which differ from each other.

IQ, 16 points for Communication, and so on). These would certainly be considered large effect sizes and clinically significant as well as statistically significant. For Age Equivalent scores, however, differences were more modest (e.g., an "8-month" difference in Communication) and were nonsignificant (using a conservative p value) for Daily Living Skills and Motor Skills, as well as for Rate of Development during IBI (which is based on Age Equivalents).

Third, we examined the proportion of children from the Younger and Older groups who fell into the seven categories of progress/outcome at exit. As shown in Table 4, Younger children were more likely to be one of the better outcome groupings compared to Older children (e.g., 39% versus 19% in Categories 1 and 2 combined).

Fourth, a one-way ANOVA was calculated comparing the mean age at intake for children in each of the seven outcome categories. This was significant ($F(6,289) = 5.68$, $p < .001$) but post hoc tests indicated that the only group differences which were significant were between the Average Functioning group and every other group (all $p < .001$). No other groups differed significantly from each other. See Table 4, last column.

Finally, regressions were computed for eight primary dependent variables at T2 (CARS Total score, VABS Standard Scores for Communication, Daily Living, Socialization, Motor and ABC, IQ, and Rate of Development During IBI) to see whether age accounted for unique variance in outcomes over and above initial levels on the outcome variables. At Step 1, we entered the T1 score for the same variable (as a way of controlling for it) and report the R^2 for the initial step. Then age was entered at Step 2 to determine whether it accounted for any additional variance. Results are shown in Table 5 (first two columns). Age was shown to account for a significant amount of unique variance for most of the outcome variables, although the proportion of variance accounted for was very small in some cases, ranging from 1 to 6% of incremental variance over and above the T1 scores on these variables. Age made no unique contribution to the T2 outcomes of VABS Daily Living Skills or ABC.

3.1.2. Initial cognitive level

Is initial IQ a good predictor of outcome? Returning to Table 2, the second column shows the correlations of initial full-scale IQ estimate with all outcome variables. In the subset of children who had an IQ score available at intake ($n = 151$), there were significant and strong correlations between initial IQ and all outcome variables. To further explore the role of IQ as a predictor, regressions were computed as described above. As shown in Table 5 (third and fourth columns), initial IQ accounted for a significant amount of incremental variance for all eight outcome variables, beyond that associated with the initial value of that variable, accounting for an additional 5–12.5% of the variance in outcome and 53% of the variance in T2 IQ.

Table 5
Variance accounted for by T1 predictor variables in regressions on dependent variables at T2.

Dependent variable T2	Age		IQ		ABC		CARS	
	R^2 Step 1	R^2 change Step 2	R^2 Step 1	R^2 change Step 2	R^2 Step 1	R^2 change Step 2	R^2 Step 1	R^2 change Step 2
CARS Total	.269	.015*	.251	.074***	.261	.118***	.269	–
VABS Com	.530	.009*	.453	.063***	.530	.019**	.531	.004
VABS DLS	.543	.003	.563	.058***	.543	.039***	.544	.012
VABS Soc	.378	.028***	.429	.116***	.378	.116***	.412	.011
VABS Mot	.466	.033**	.421	.066***	.466	.006	.455	.001
VABS ABC	.596	.000	.576	.053***	.596	–	.599	.003
IQ Estimate	.532	.063***	.532	–	.535	.059***	.543	.038**
Rate During	.067	.045***	.127	.125***	.074	.001	.064	.006

Note: Com = Communication; DLS = Daily Living Skills; Soc = Socialization; Mot = Motor Skills domains on VABS. Rate During = Rate of Development during IBI.
All ANOVAS are significant at $p < .001$.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table 6
Categories of outcome for children with different diagnostic severity.

Progress/outcome group	Clinical diagnosis *n* (%)		CARS category *n* (%)			CARS Total *M* (*SD*)
	Autism	PDD-NOS	Not Autism	Mild/Mod.	Severe Autism	
(1) Average Functioning	19 (7.7%)	11 (25.6%)	9 (37.5%)	17 (12.1%)	5 (4.5%)	32.4 (4.7)
(2) Substantial Improvement	37 (15.0%)	6 (14.0%)	5 (20.8%)	24 (17.1%)	12 (10.9%)	35.1 (4.8)
(3) Clinically Sig. Improvement	75 (30.5%)	13 (30.2%)	2 (8.3%)	48 (34.3%)	34 (30.9%)	36.5 (4.6)
(4) Less Autistic	30 (12.2%)	1 (2.3%)	0	6 (4.3%)	24 (21.8%)	40.9 (4.7)
(5) Minimal Improvement	21 (8.5%)	3 (7.0%)	3 (12.5%)	11 (7.9%)	8 (7.3%)	36.5 (5.5)
(6) No change	45 (18.3%)	8 (18.6%)	4 (16.7%)	28 (20.0%)	16 (14.5%)	36.7 (6.0)
(7) Worse	19 (7.7%)	1 (2.3%)	1 (4.2%)	6 (4.3%)	11 (10.0%)	38.4 (5.4)

3.1.3. Initial adaptive level

Are outcomes strongly related to initial level of adaptive functioning? Initial Vineland ABC scores were significantly and quite highly correlated with all outcome variables, as seen in Table 2, third column. Regressions, shown in Table 5 (fifth and sixth columns), indicate that initial ABC scores accounted for significant incremental variance, beyond that associated with the initial value of that variable, for most outcome variables (accounting for 2–12% of the variance), but were not significant for Motor scores at T2 or Rate of Development during IBI.

3.1.4. Diagnostic severity

Do children with more severe autism symptoms or with a particular categorical diagnosis show different outcomes? At entry into the program, three diagnostic indicators were available: CARS Total score; CARS category (not autism, mild/moderate autism, severe autism), and DSM-IV clinical diagnosis. The latter was initially categorized as Autistic Disorder, PDD-NOS, or a less specific diagnosis of ASD/PDD. Further examination suggested the AD and ASD/PDD groups could be combined as the distinction appeared to be accounted for by regional variations in diagnostic nomenclature rather than different child characteristics (they had similar CARS Total scores and both were significantly higher than the group with PDD-NOS).

Each of the diagnostic indicators was examined in relation to children's progress/outcome category, as shown in Table 6. A Mann–Whitney *U* comparing the proportion of children in the seven outcome groups as a function of having AD versus PDD-NOS was significant ($z = -2.45$, $p = .014$). Children with a diagnosis of Autistic Disorder were relatively less likely to achieve Average Functioning, more likely to be in the Less Autistic Group and somewhat more likely to fall into the Worse group at exit, relative to children with PDD-NOS.

Looking at the CARS Categories, a much larger proportion of children who had been rated as below the autism cutoff on the CARS at T1 achieved Average Functioning or Substantial Improvement (about 60% of the Non-autism classified children) achieved these outcomes versus about 30% of the mild/moderate and 15% of the severe children. The children who ended up in the Less Autistic Group (category 4) were likely to have been in the severe autism range on the CARS at intake.

A one-way ANOVA for CARS Total score in the seven outcome groups was significant ($F(6,267) = 8.10$, $p < .001$) and post hoc tests suggested that the Average Functioning Group had entered the program with lower CARS scores than all other groups. The Less Autistic group had entered with the highest CARS scores (significantly higher than all other groups except the Category 7, the group who got Worse).

Correlations of initial CARS Total scores with the T2 outcome variables were shown previously in Table 2 and indicate modest negative correlations with all the outcome variables. However, regression analyses indicated that, in general, initial autism severity did not contribute to the prediction of the outcome variables, as shown in Table 5. The one exception was IQ: initial CARS score accounted for an additional 4% of the variance in T2 IQ Estimate, controlling for T1 IQ Estimate.

3.1.5. Simultaneous exploration of predictors

Which predictors are most important and account for most variance? Since the majority of studies in the literature examining predictors have sought to predict outcome levels of IQ, we carried out a final set of prediction analyses to predict T2 IQ Estimate in the subsample who had all necessary scores ($n = 105$). Given that all four categories of predictors were found to be important when examined in isolation, but that all the variables are also inter-correlated, this hierarchical regression involved a simultaneous examination of the role of age, initial IQ, ABC score, and CARS in predicting T2 IQ scores. As per the logic outlined earlier, IQ at T1 was entered first (accounting for 54% of the variance); then age at entry, which accounted for an additional 5.3% of the variance ($p < .001$), then Vineland ABC score, which accounted for an additional 2.3% of the variance ($p < .05$), then CARS Total score which accounted for an additional 2.0% of the variance ($p < .05$). The ANOVA for this final model was significant ($F(4,108) = 47.48$, $p < .001$) and the final R^2 was .637. Thus, 64% of the variance in outcome IQ can be predicted based on the combination of these four variables.

3.2. Further examination of the average functioning group

Given the considerably important and perhaps surprising result of finding *any* children resembling "best outcomes" in a community effectiveness study (see Perry et al., 2008; Remington et al., 2007 for further discussion of this distinction), closer

Table 7
Developmental and diagnostic results for Average Outcome group ($n = 32$) compared to remainder of sample at intake and change from intake to exit.

	Remainder of sample at intake ($n = 300$)[a]	Average outcome group ($n = 32$)[b]		Independent t average versus others at intake	Paired t change intake–exit
		Intake *M (SD)*	Exit *M (SD)*		
Autism Severity CARS Total	37.05 (5.38)	32.44 (4.68)	24.24 (3.85)	4.58, $p < .001$	7.70, $p < .001$
VABS Standard Scores					
Communication	52.56 (9.87)	68.46 (12.74)	97.32 (20.52)	−7.55 $p < .001$	−7.25, $p < .001$
Daily Living	52.81 (11.21)	66.82 (10.78)	76.75 (16.83)	−6.46 $p < .001$	−3.19, $p = .004$
Socialization	56.01 (7.27)	65.50 (8.39)	79.89 (14.18)	−6.33 $p < .001$	−4.97, $p < .001$
Motor	60.07 (14.36)	76.21 (12.72)	89.63 (12.06)	−5.87 $p < .001$	−4.54, $p < .001$
ABC	53.74 (8.30)	66.93 (8.38)	84.65 (14.87)	−7.89 $p < .001$	−6.15, $p < .001$
VABS Age Equivalents					
Communication	15.17 (8.36)	20.28 (12.09)	60.45 (17.28)	−2.89, $p = .004$	−10.86 $p < .001$
Daily Living	20.81 (6.44)	21.66 (8.60)	48.28 (14.25)	−0.77, *ns*	−10.76 $p < .001$
Socialization	13.44 (6.12)	16.96 (8.30)	46.00 (14.82)	−2.86, $p = .005$	−9.66 $p < .001$
Motor	28.91 (8.83)	26.72 (7.03)	57.00 (9.86)	0.10, *ns*	−13.74 $p < .001$
ABC	16.50 (6.16)	19.93 (8.60)	51.86 (13.07)	−2.71, $p = .007$	−11.62 $p < .001$
Cognitive Level					
IQ Estimate	43.45 (18.70)	64.08 (14.75)	102.38 (13.13)	−3.68 $p < .001$	−6.89, $p < .001$
MA Estimate	22.56 (10.98)	25.77 (9.58)	62.12 (11.34)	−0.78, *ns*	−11.51 $p < .001$
Rate of Development	.31 (.11)	.48 (.14)	1.33 (.74)	−7.35 $p < .001$	−5.89, $p < .001$
Program variables					
Age at Intake	54.75 (11.87)	41.91 (12.54)		5.75 $p < .001$	
Duration	17.70 (7.91)	25.92 (8.94)		−5.47 $p < .001$	

[a] n varies: CARS $n = 243$; VABS $n = 252$ except Motor $n = 238$; Cognitive $n = 130$; Rate $n = 256$; Program variables $n = 300$.
[b] Paired data: CARS $n = 31$; VABS $n = 28$ except Motor $n = 25$; Cognitive $n = 13$; Rate $n = 28$.

examination of the assessment data on these 32 children (28 boys and 4 girls) seems warranted. Table 7 (second and third columns) shows their scores at intake and exit. CARS Total scores were significantly lower at exit, a change of close to 2 *SD*s, a very substantial change, clinically. Four children, however, were still in the autism range on the CARS, though close to the cut score (30–31). It is important to note that these four children (at least) likely continued to display residual features of ASD in spite of their high cognitive and/or adaptive functioning (they had a mean ABC of 94 and FSIQ of 104). Adaptive behavior standard scores also improved significantly. The magnitude of gain was especially impressive (close to 30 points) for the Communication domain, and was most modest for Daily Living (10 points). Overall, however, their Adaptive Behavior outcomes were considerably more variable and not quite as good overall as their cognitive outcomes. The rate of development for this subgroup of children increased almost 3-fold from a mean of .48 to a mean of 1.33, which is greater than a typical rate of development. Unfortunately cognitive scores at both intake and exit were not available for all children but IQ estimates were available for 13 pairs and indicated a very substantial mean gain of close to 40 IQ points for this small group of children. At exit, complete IQ data were available for 29 of the 32 children. Their Full Scale IQ was in the average range with a mean of 102.38 ($SD = 13.13$), with Performance IQ ($M = 113.00$; $SD = 14.18$) substantially higher than Verbal ($M = 92.52$; $SD = 12.41$). All cognitive scores were in the average range for individual children except one child who scored below 85 on all three IQ scores (but his ABC was over 85) and four additional children whose VIQs were below 85.

Possible factors related to Average Outcome status were explored, including developmental and diagnostic variables and program variables (age, duration) for these children compared to the remainder of the children, whose data are shown in the first column of Table 7. The children who achieved Average Functioning at outcome differed at intake significantly and substantially from the remainder of the sample. On the VABS, their standard scores were significantly higher (at least 1 *SD* difference). However, age equivalents were much more similar with no significant difference in Daily Living and Motor Skills and significant but small differences in other domains. Thus, the absolute level of functioning was not so different (but their age differed). The same pattern is seen on the cognitive variables, i.e., the Average outcome children had had higher IQ but similar mental age at intake.

Diagnostically, the Average Functioning children also seem to have had somewhat milder autism at intake. They had significantly lower CARS scores than the remainder of the children (about 1 *SD* in magnitude), as shown in Table 7. In fact, 10 children (29%) were in the non-autism range (versus 6% of the remaining children) and only 5 (16%) were in the severe autism range (versus 43% of the remaining children). In terms of DSM-IV clinical diagnosis, 11 children (35%) had a diagnosis of PDD-NOS (versus 12% in the remaining children).

The program parameters of age and duration were also examined in these 32 children who achieved average functioning. They began IBI at significantly and substantially younger ages (42 months versus 55 months at intake) and received service for somewhat longer durations (26 months versus 18 months). See Table 7 for further details.

Table 8
Developmental and diagnostic results for Poor Outcome group ($n = 75$) compared to Remainder of Sample at Intake and Change from Intake to Exit.

	Remainder of sample at intake ($n = 257$)[a]	Poor outcome group ($n = 75$)[b]		Independent t poor versus others at intake	Paired t change intake–exit
		Intake *M (SD)*	Exit *M (SD)*		
Autism Severity CARS Total	36.24 (5.30)	35.90 (4.69)	35.48 (4.06)	1.24, *ns*	0.91, *ns*
VABS Standard Scores					
Communication	55.30 (11.33)	50.52 (9.92)	44.68 (11.15)	−2.99, $p = .003$	7.76, $p < .001$
Daily Living	55.31 (11.94)	50.93 (11.00)	39.20 (11.41)	−2.50, $p = .01$	13.39, $p < .001$
Socialization	57.59 (8.06)	54.82 (6.54)	51.52 (6.29)	−2.28, *ns*	6.50, $p < .001$
Motor	63.39 (14.99)	62.52 (12.07)	55.28 (16.07)	−2.90, $p = .004$	3.64, $p = .001$
ABC	56.00 (9.24)	52.09 (8.04)	45.13 (8.36)	−2.90, $p = .004$	12.74, $p < .001$
VABS Age Equivalents					
Communication	16.35 (9.12)	13.47 (7.61)	17.85 (11.02)	−2.17, *ns*	−5.00, $p < .001$
Daily Living	21.24 (6.81)	19.91 (6.19)	23.80 (6.97)	−1.35, *ns*	−6.94, $p < .001$
Socialization	14.11 (6.68)	12.62 (4.99)	14.12 (5.96)	−1.37, *ns*	−2.28, $p = .03$
Motor	29.45 (8.98)	25.13 (6.21)	32.58 (10.82)	−1.73, *ns*	−4.81, $p < .001$
ABC	17.26 (6.68)	15.41 (5.49)	18.63 (6.97)	−1.85, *ns*	−5.48, $p < .001$
Cognitive Level					
IQ Estimate	49.00 (17.95)	34.36 (19.55)	33.21 (15.48)	−4.51, $p < .001$	0.55, *ns*
MA Estimate	24.38 (10.46)	19.23 (10.99)	23.41 (9.05)	−3.35, $p = .001$	−3.92, $p = .001$
Rate of Development	.33 (.13)	.29 (.09)	.18 (.29)	−2.51, $p = .01$	2.85, $p = .006$
Program variables					
Age at Intake	52.97 (12.82)	54.52 (11.83)		0.92, *ns*	
Duration	18.97 (8.50)	17.48 (8.09)		−1.32, *ns*	

[a] n varies: CARS $n = 208$; VABS $n = 210$ except Motor $n = 198$; Cognitive $n = 111$; Rate $n = 210$; Program variables $n = 257$.
[b] Paired data: CARS $n = 56$; VABS $n = 71$ except Motor $n = 29$; Cognitive $n = 28$; Rate $n = 71$.

3.3. Further examination of the poor outcome group

At the other extreme, further analyses were conducted on the subgroup of children who did not appear to benefit from participation in the IBI program (i.e., those who fell into the No Change or Worse categories of Progress/Outcome). This set of analyses included 75 children (11 girls; 64 boys). Table 8 shows the intake and exit scores (in the second and third columns of numbers) for children in the Poor Outcome categories (paired data only). Note that the signs for the *t* statistics are opposite for the Age Equivalents and the Standard Scores. On average, these children made statistically significant *gains* in adaptive skills when considering the Age Equivalents and MA. However, these were very small gains considering their time in the program, which averaged 17 months. Consequently, their Standard Scores (which are corrected for age) *declined* significantly on all subscales of the VABS. The magnitude of this change varied but was approaching one *SD* in size for the ABC Composite score. This pattern of results is often seen clinically when children are aging faster than they are developing. CARS Total score and IQ Estimate (available for only a subset) did not change significantly in either direction.

Possible factors related to Poor Outcome status were explored, including developmental and diagnostic variables and program variables (age, duration) for these children compared to the remainder of the children, whose data are shown in the first column of Table 8. In most respects, the Poor Outcome children did not differ appreciably from the larger sample of which they were a part. A series of independent *t*-tests comparing the initial intake assessment scores for the group who had poor outcomes (Category 6 and 7) to the remaining children (Categories 1 through 5) indicated that the Poor Outcome children were significantly lower at intake on: IQ estimate (approximately 34 versus 49), MA (approximately 19 months versus 24 months), Initial rate of Development (.29 versus .33), and all VABS Standard Scores except Socialization. However, most of these differences are not very large in magnitude from a clinical significance perspective.

Diagnosis and diagnostic severity on the CARS did not differ and the proportion of children in the three CARS categories was similar in the poor outcome children versus the remainder of the sample, as was the proportion of children with a clinical diagnosis of PDD-NOS.

The program parameters of age and duration were also examined in reference to these Poor Outcome children. There were no significant differences in age at program entry between those who showed some progress and those in the Poor Outcome groups. In both cases the means were within a month of each other. One might wonder whether longer durations were of benefit to these children but the data suggest not. There was no correlation between duration and exit assessment scores or rate of progress during IBI. Furthermore, there were 16 children in the Poor Outcome Group who had received service for 2 years or longer and they did not differ significantly on any outcome variables from those with shorter durations.

4. Discussion

This paper reports on initial age at entry, cognitive level, adaptive skills, and diagnostic severity as predictors of outcome in a large group of children receiving community-based IBI (the same group whose outcomes were reported in Perry et al., 2008). All four predictors were found to be important, to some extent, in relationship to outcomes of cognitive and adaptive level, severity of autism, rate of development during intervention, and categorical outcome.

Earlier age at program entry was associated with better outcomes when comparing the children under 4 versus over 4 years of age. This is consistent with previous studies which include samples with wide age ranges (Anderson et al., 1987; Bibby, Eikeseth, Martin, Mudford, & Reeves, 2002; Fenske et al., 1985; Harris & Handleman, 2000). There were moderate significant negative correlations of age at entry with outcome variables, suggesting that, in general, younger children did better. Age accounted for a significant but modest amount of unique variance for most outcome variables beyond the initial value of the variable (e.g., controlling for initial IQ, age at entry accounted for an additional 6% of the variance in IQ outcomes). As noted earlier, several studies in the literature do not report a relationship with age (Eikeseth et al., 2002, 2007; Hayward et al., 2009; Lovaas, 1987; Smith et al., 2000). Narrower age ranges, limited power in the smaller samples in other studies, different statistical techniques (median splits versus correlations or regressions), and sample differences likely account for these discrepancies.

Further examination of the age issue suggests a non-linear relationship, at least in terms of the seven progress/outcome categories. Children who ended up in the average functioning group had been substantially younger than children in all six other categories (who did not differ from each other). This finding aligns well with the conclusions of Makrygianni and Reed (2010) that effect sizes were uniformly large in studies of very young children. This supports the ‘‘sensitive period’’ notion and implies that, if the goal of IBI is to alter developmental trajectories and boost children into the average range, this may only be feasible if children begin IBI when they are very young. Thus, it is crucial that efforts are made to encourage early diagnosis and to reduce wait lists and ensure that children receive IBI early whenever possible. This is not to say, of course, that older children will not benefit or do not merit our best treatment efforts, but they are less likely to evidence highly successful outcomes such as average functioning. Also, it should be acknowledged that early entry into treatment may be associated with other factors which relate to outcome, such as parents' knowledge and resourcefulness.

Initial IQ (available in a subset) was very strongly correlated with outcome variables and accounted for significant unique variance (5 to 12% of the variance) in every regression. These results are consistent with those of a number of other studies (Eikeseth et al., 2002, 2007; Harris & Handleman, 2000; Hayward et al., 2009; Sallows & Graupner, 2005) but are inconsistent with findings of Cohen et al. (2006) and Smith et al. (2000). The discrepancy may again be related to statistical power and/or sample differences. Adaptive behavior levels at intake have been less of a focus as predictors of outcome but our results are very similar to those of Sallows and Graupner (2005), who reported moderate correlations of intake cognitive and adaptive scores with subsequent IQ. In the present study, adaptive behavior levels were also correlated with outcome variables and regressions indicated that the ABC score accounted for significant incremental variance in most outcome variables. These findings appear consistent with those of the meta-analysis by Makrygianni and Reed (2010), who reported initial adaptive levels correlated with effect sizes for language outcomes and adaptive behavior outcomes. However, it should be reiterated that initial cognitive and adaptive levels likely predict later cognitive and adaptive levels in children not receiving IBI as well (Eikeseth et al., 2007; Flanagan, 2009; Gabriels, Hill, Pierce, Rogers, & Wehner, 2001).

The few other studies that have examined autism severity as a predictor have used a variety of measures, making comparison difficult (Howlin et al., 2009). In the present study, children's initial level of autism severity, as measured by the CARS, was significantly and moderately correlated with children's outcome scores, somewhat similarly to Sallows and Graupner's report that ADI-R social and communication scores were among the predictors of post-treatment IQ. However, in our regression analyses, CARS scores typically did not account for much additional variance (except in the case of IQ, where initial CARS score accounted for unique variance beyond initial IQ levels). Children with particularly high autism severity seemed to show very large reductions in autism symptomatology which may reflect a regression to the mean phenomenon, at least in part. Children with a diagnosis of PDD-NOS were more likely to achieve average functioning, which is consistent with the finding of Smith et al. (2000) that bigger IQ gains were seen in children with PDD-NOS versus autism. Although the IQ gain in our sample did not differ between children with AD and PDD-NOS, those with PDD-NOS had initial IQs about 10 points higher.

Our final set of prediction analyses focused on the question of the relative strength of different predictors considered concurrently. Focusing on the prediction of outcome IQ only (which has been most studied as an outcome variable), these analyses involved simultaneous examination of initial age, cognitive level, adaptive skills, and diagnostic severity to determine the relative proportion of variance accounted for by each of these inter-correlated variables. The final regression to predict IQ outcome accounted for a substantial 64% of the variance. Clearly initial IQ accounts for the vast majority of the outcome variance and the other variables a much smaller proportion. Of the other variables, age appears to account for a greater proportion of the variance and adaptive behavior and autism severity a small but significant proportion. Thus, all four categories of predictors we explored seemed to be important. However, it should be noted that there was still some unexplained variance, which means other child, family, or intervention factors not measured in the present study may also be influential.

A number of secondary analyses explored correlates of outcome within the two extremes of the range of outcomes (average functioning and poor outcome). There appeared to be a constellation of factors associated with achieving average functioning. At intake to IBI, these 32 children had higher cognitive and adaptive skills for their age and milder autism symptomatology. In addition, they were, on average, a year younger than the other children (beginning IBI at 42 months on average versus 55) and they had somewhat longer treatment duration (26 months on average versus 18). However, it is important to state that, unfortunately, not all children with these characteristics necessarily achieved average functioning.

At the other end of the distribution, initial assessment scores were not necessarily helpful in predicting which children will do poorly. Although children who had poor outcomes were significantly lower on several intake variables, these differences were small and there were certainly many children with similarly low developmental levels whose outcomes were more positive. Other factors not measured in this study may account for the poor outcomes, such as other child characteristics not measured, family factors, and/or treatment quantity or quality.

It has been suggested that research on predictors will be helpful to decide which specific children are good candidates for IBI and which are not. However, these results suggest that when a child begins IBI, predictions regarding outcomes are unlikely to be very accurate. Possibly good candidates are more easily identified than poor candidates, although this hypothesis requires further empirical evidence. Some recent research has suggested, however, that very early response to treatment (within the first weeks or months of intake) may be an effective predictor (Goin-Kochel, Myers, Hendricks, Carr, & Wiley, 2007; Sallows & Graupner, 2005; Sullivan, 2010; Weiss, 1999). Resource allocation decisions are ethically and practically challenging when there are many children needing service, some children receiving it but not necessarily benefiting very much and others remaining on waitlists. Thus, it seems to us most responsible to offer children with autism an initial trial of IBI, as recommended by the Ontario Expert Clinical Panel (Szatmari et al., 2007) and to monitor their progress carefully using clear and specific benchmarks, as recommended by the Ontario Benchmark Development Expert Panel (Freeman et al., 2008).

There are significant limitations to the larger study which are discussed fully by Perry et al. (2008) but these are less relevant in the present paper since it is a within-group analysis. Limitations which do apply to the present study include the variable missing data, especially for cognitive scores. Because of the large sample, statistical power is certainly adequate but Type I error may be an issue because of the many analyses and repeated analyses with the same data. However, the magnitude of the various statistics used is typically reflective of very large effect sizes, so it seems unlikely that chance findings have been overinterpreted. One of the major limitations is that there are no measures of treatment quality, intensity, supervision approach, and so on, which are one class of factors which are likely to impact on effectiveness in general (Koudys & Perry, 2010; Makrygianni & Reed, 2010; Reichow & Wolery, 2009) and which could well interact with child characteristics such as age and IQ to predict outcome differentially.

Finally, it is important to acknowledge that, without a control group, we cannot be sure what progress children might have made without IBI or with a different treatment. As already noted, studies have shown correlations of initial IQ with children's outcomes, regardless of treatment. Furthermore, predictors may well have a differential effect (i.e., an interaction) as a function of what intervention children are receiving. For example, in another recent study in the Ontario program, Flanagan (2009) showed that higher adaptive levels at intake predicted higher adaptive levels at exit in both the IBI group and the waitlist comparison group, but that younger age was a significant predictor in the IBI group only.

In spite of these cautions, we believe this paper contributes to the literature on the predictors and correlates of outcome in IBI by virtue of its size and the fact that it reflects a community-based effectiveness sample of a culturally and socioeconomically diverse group of children with heterogeneous developmental and diagnostic characteristics.

Acknowledgements

We are grateful to the Ontario Ministry of Children and Youth Services for funding this study. However, the views expressed here are those of the authors and do not necessarily represent the position of the Ministry. The Ministry in no way influenced the interpretation or writing of this paper. We appreciate the assistance of Alissa Levy, Helen Penn Flanagan, Alice Prichard, Abbie Solish, April Sullivan, and Kerry Wells with data collection, data entry and verification and Don Downer for data verification, scheduling, and other administrative support.

References

Anderson, S. R., Avery, D. L., DiPietro, E. K., Edwards, G. L., & Christian, W. P. (1987). Intensive home-based early intervention with autistic children. *Education and Treatment of Children, 10*, 352–366.

Bibby, P., Eikeseth, S., Martin, N. T., Mudford, O. C., & Reeves, D. (2002). Erratum to "Progress and outcomes for children with autism receiving parent-managed intensive interventions". *Research in Developmental Disabilities, 23*, 79–104.

Birnbrauer, J. S., & Leach, D. J. (1993). The Murdoch early intervention program after 2 years. *Behavior Change, 10*, 63–74.

Cohen, H., Amerine-Dickens, M., & Smith, T. (2006). Early intensive behavioral treatment: Replication of the UCLA model in a community setting. *Developmental and Behavioral Pediatrics, 27*, S145–S155.

Eikeseth, S., Smith, T., Jahr, E., & Eldevik, S. (2002). Intensive behavioral treatment at school for 4–7-year-old children with autism: A 1-year comparison controlled study. *Behavior Modification, 26*, 49–68.

Eikeseth, S., Smith, T., Jahr, E., & Eldevik, S. (2007). Outcome for children with autism who began intensive behavioral treatment between ages 4 and 7: A comparison controlled study. *Behavior Modification, 31*, 264–278.

Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2009). Meta-analysis of early behavioral intervention for children with autism. *Journal of Clinical Child and Adolescent Psychology, 38*, 439–450.

Fenske, E. C., Zalenski, S., Krantz, P. J., & McClannahan, L. E. (1985). Age at intervention and treatment outcome for autistic children in a comprehensive intervention program. *Analysis and Intervention in Developmental Disabilities, 5*, 49–58.

Freeman, N. L., Brown, R., Dunn Geier, J., Lindblad, T., Perry, A., & Reitzel, J., et al. (2008). The development of benchmarks for the delivery of intensive behavioural intervention for children with autism spectrum disorders in Ontario. Report submitted to the Ministry of Children and Youth Services.

Flanagan, H. E. (2009). The impact of community-based Intensive Behavioural Intervention. Unpublished doctoral dissertation, York University, Toronto, ON, Canada.

Gabriels, R. L., Hill, D. E., Pierce, R. A., Rogers, S. J., & Wehner, B. (2001). Predictors of treatment outcome in young children with autism. *Autism, 5*, 407–425.

Goin-Kochel, R. P., Myers, B. J., Hendricks, D. R., Carr, S. E., & Wiley, S. B. (2007). Early responsiveness to intensive behavioral intervention predicts outcomes among preschool children with autism. *International Journal of Disability, Development and Education, 54*, 151–175.

Granpeesheh, D., Dixon, D. R., Tarbox, J., Kaplan, A. M., & Wilke, A. E. (2009). The effects of age and treatment intensity on behavioral intervention outcomes for children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 3*, 1014–1022.

Harris, S. L., & Handleman, J. S. (2000). Age and IQ at intake as predictors of placement for young children with autism: A 4 to 6 year follow up. *Journal of Autism and Developmental Disorders, 30*, 137–142.

Hayward, D., Eikeseth, S., Gale, C., & Morgan, S. (2009). Assessing progress during treatment for young children with autism receiving intensive behavioral interventions. *Autism, 13*, 613–633.

Howard, J. S., Sparkman, C. R., Cohen, H. G., Green, G., & Stanislaw, H. (2005). A comparison of intensive behavior analytic and eclectic treatments for young children with autism. *Research in Developmental Disabilities, 26*, 359–383.

Howlin, P., Magiati, I., & Charman, T. (2009). Systematic review of early intensive behavioral interventions for children with autism. *American Journal on Intellectual and Developmental Disabilities, 114*, 23–41.

Koudys, J. L. & Perry, A. (2010). Identifying active ingredients in Intensive Behavioural Intervention programs for *children with autism*. Unpublished Research Paper, York University, Toronto, ON, Canada.

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.

Makrygianni, M. K., & Reed, P. (2010). A meta-analytic review of the effectiveness of behavioural early intervention programs for children with Autistic Spectrum Disorders. *Research in Autism Spectrum Disorders, 4*, 577–593.

Matson, J. L. (2007). Determining treatment outcome in early intervention programs for autism spectrum disorders. *Research in Developmental Disabilities, 28*, 207–218.

Matson, J. L., & Smith, K. R. M. (2008). Current status of intensive behavioral interventions for young children with autism and PDD-NOS. *Research in Autism Spectrum Disorders, 2*, 60–74.

McEachin, J. J., Smith, T., & Lovaas, O. I. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal on Mental Retardation, 97*, 359–372.

Perry, A., Cummings, A., Dunn Geier, J., Freeman, N., 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.

Perry, A., Flanagan, H. E., Dunn Geier, J., & Freeman, N. L. (2009). The Vineland Adaptive Behavior Scales in young children with autism spectrum disorders at different cognitive levels. *Journal of Autism and Developmental Disorders, 39*, 1066–1078.

Reichow, B., & Wolery, M. (2009). Comprehensive synthesis of early intensive behavioral interventions for young children with autism based on the UCLA Young Autism Project model. *Journal of Autism and Developmental Disorders, 39*, 23–41.

Remington, B., Hastings, R. P., Kovshoff, H., degli Espinosa, F., Jahr, E., Brown, T., et al. (2007). Early Intensive Behavioral Intervention: Outcomes for children with autism and their parents after two years. *American Journal on Mental Retardation, 112*, 418–438.

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*, 417–438.

Shine, R., & Perry, A. (2010). The relationship between parental stress and intervention outcome of children with autism. *Journal on Developmental Disabilities, 16*(2), 10–12.

Schopler, E., Reichler, R. J., & Renner, B. R. (1988). *The Childhood Autism Rating Scale (CARS)*. Los Angeles: Western Psychological Services.

Smith, T., Groen, A. D., & Wynn, J. W. (2000). Randomized trial of intensive early intervention for children with pervasive developmental disorder. *American Journal on Mental Retardation, 105*, 269–285.

Solish, A. (2010). Parents' involvement in behavioural interventions for their children with autism. Unpublished doctoral dissertation, York University, Toronto, ON, Canada.

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

Sullivan, A. (2010). Developmental trajectories of young children with autism enrolled in an IBI program: What the *ABLLS can tell us about their progress.* Unpublished doctoral dissertation, York University, Toronto, ON, Canada.

Szatmari, P., Freeman, N. L., Laredo, S., Nicolson, R., Perry, A., & Pryor, C., et al. (2007). Development of clinical practice guidelines for the delivery of intensive behavioural intervention for children with autism spectrum disorders. Report submitted to the Ontario Ministry of Children and Youth Services, ON, Canada.

Weiss, M. (1999). Differential rates of skill acquisition and outcomes of early intensive behavioral intervention for autism. *Behavioral Interventions, 14*, 3–22.