HHS Public Access

Author manuscript

Autism Res. Author manuscript; available in PMC 2020 December 08.

Published in final edited form as:

Autism Res. 2020 November ; 13(11): 1867–1879. doi:10.1002/aur.2397.

The Autism Impact Measure (AIM): Examination of Sensitivity to Change

Micah O. Mazurek, PhD,
Department of Human Services, School of Education and Human Development, University of Virginia

Coleen Carlson, PhD,
Department of Psychology, Texas Institute for Measurement, Evaluation, and Statistics University of Houston

Mary Baker-Ericzén, PhD,
Child and Adolescent Services Research Center, Autism Discovery Institute, Rady Children's Hospital

Eric Butter, PhD,
Department of Pediatrics and Psychology, The Ohio State University, Nationwide Children's Hospital

Megan Norris, PhD,
Child Development Center, Nationwide Children's Hospital

Christopher Barr, PhD,
Texas Institute for Measurement, Evaluation, and Statistics, University of Houston

Stephen Kanne, PhD
Department of Health Psychology, Thompson Center for Autism & Neurodevelopmental Disorders, University of Missouri

Abstract

The Autism Impact Measure (AIM) was designed specifically for treatment-outcome assessment in children with ASD, focusing on treatment-relevant aspects of symptom presentation and efficient detection of short-term improvement. The AIM demonstrated strong reliability and validity in initial psychometric studies. The current study evaluated the AIM's sensitivity to change across well-established treatments. The sample included 471 children with ASD (ages 2-14) participating in one of six treatments. The AIM was administered at baseline and 6-week intervals and a battery of domain-specific concurrent measures was also administered. A longitudinal repeated measures design examined the degree to which: 1) AIM domain scores changed over time in response to treatment, and 2) change in AIM domains was associated with

Please address correspondence to: Micah Mazurek, PhD, University of Virginia, PO Box 400267, Charlottesville, VA 22903, Phone: (434) 924-5692, mazurekm@virginia.edu.

Conflict of Interest Statement

Drs. Mazurek and Kanne will receive royalties from Western Psychological Services for the Autism Impact Measure. Drs. Carlson, Baker-Ericzén, Butter, and Norris declare that they have no conflicts of interest.

change in measures of similar constructs. Results across growth curve models indicated that AIM domains are sensitive to change in symptoms across treatment. Across all models, symptoms decreased over time, with some deceleration in rate of improvement. For all AIM domains except Repetitive Behavior, symptoms improved as a function of treatment group. Correlations of change between AIM and other measures varied across domains (from .01-.43 across measures). This was the first large-scale study to systematically evaluate sensitivity to change in a measure of core ASD symptoms. The results provide support for the AIM's ability to detect short-term improvement across symptom domains and indicate that AIM domains are sensitive to change overall and as a function of different treatment conditions. The brief repeated assessment window also highlights the AIM's utility for detecting improvements across short-term treatments.

Lay Summary

Good measures are important for assessing outcomes in children with autism. However, there are few tools for tracking short-term changes in autism symptoms. This study tested a new measure, the Autism Impact Measure (AIM), in a large group of children with autism. The results showed that the AIM appears to be a valid and accurate tool for measuring autism symptoms. The AIM may be a helpful tool for researchers and clinicians interested in tracking short-term improvements in autism symptoms.

Keywords

Autism spectrum disorder; treatment outcome; autism symptoms; assessment; measurement

Introduction

Autism spectrum disorder (ASD) is a complex neurodevelopmental disorder characterized by social impairment and restricted and repetitive behaviors (American Psychiatric Association, 2013). Increases in prevalence, cost, and societal impact over recent years (Buescher, Cidav, Knapp, & Mandell, 2014; Christensen et al., 2016) have prompted a surge of interest in developing new psychosocial and pharmacological treatments for ASD (Canitano, 2014; Kasari, Shire, Factor, & McCracken, 2014; Wong et al., 2015). However, a lack of psychometrically sound treatment outcome tools has limited progress in the field, resulting in calls for development of appropriate autism outcome measures that are sensitive to change in primary symptom domains (Anagnostou et al., 2015; Baker-Ericzén, Brookman-Frazee, & Brodkin, 2018; Bolte & Diehl, 2013; McConachie et al., 2015; Scahill et al., 2015).

To date, ASD treatment outcome research has employed a wide variety of outcome tools, ranging from measures of associated symptoms to study-specific measures of discrete behaviors/skills. While informative for assessment of unique aspects of treatment effectiveness, this lack of consistent outcome measures has limited comparisons of treatment effects across studies and treatments, and likely limits the ability to find strong evidence of efficacy across interventions (Rogers & Vismara, 2008; Tachibana et al., 2017; Warren et al., 2011). Similarly, researchers often employ multiple measures to assess various aspects of symptom presentation, contributing to burden for participants and researchers, and reducing

the feasibility of repeated measurement across a typical course of treatment. Many existing tools require significant time to administer, and yield scores that are not comparable across domains and measures. Most importantly, they were generally not designed to be sensitive to short-term improvements.

Measures of core ASD symptoms most commonly used in treatment-outcome studies were originally designed for either screening or diagnosis. As such, they were designed primarily to determine categorical risk status or presence or absence of disorder. For example, the widely-used Autism Diagnostic Observation Schedule – 2nd Edition (ADOS-2) shows good sensitivity and specificity with regard to ASD diagnosis (Lord et al., 2012); however, it comprises five different modules with differing numbers and types of items and activities. Despite the development of calibrated severity scores (CSS; now termed the comparison score or CS) that account for differences in age, IQ and language (Esler et al., 2015; Gotham, Pickles, & Lord, 2009), neither ADOS-2 CSS nor raw scores have proven successful in measuring short-term improvements (Brian, Smith, Zwaigenbaum, Roberts, & Bryson, 2016; Estes et al., 2015; Thurm, Manwaring, Swineford, & Farmer, 2015). Similarly, the Autism Diagnostic Interview – Revised (ADI-R) has strong psychometric properties with regard to diagnostic accuracy (Lord, Rutter, & Le Couteur, 1994), but is not well-suited to repeated measurement of incremental change in current symptoms. In addition, both measures carry a significant administrative burden in terms of the time required for reliable administration and scoring. Another commonly-used measure of core ASD symptoms, the Social Responsiveness Scale (SRS) (Constantino & Gruber, 2012) has a relatively low administrative burden, but assesses symptoms as they have occurred over the previous 6 months and is conceptualized as a measure of stable autism traits, limiting its utility for assessing short-term change.

Various research teams are working to develop measures that are sensitive to change in core symptoms of ASD. The new Brief Observation of Social Communication (BOSCC) measures treatment response in minimally verbal children with ASD (Grzadzinski et al., 2016) through observation of brief interactions between children and caregivers. Initial psychometric properties are strong (Grzadzinski et al., 2016) and demonstrate responsiveness to change following treatment (Kitzerow, Teufel, Wilker, & Freitag, 2016). Although promising, the BOSCC is only appropriate for young children with minimal language, and only assesses social communication rather than the full range of core symptoms. The Social Communication Checklist (SCC) is another tool specifically targeting social communication skills in young children with ASD. An initial psychometric study showed good reliability and construct validity, and demonstrated evidence of sensitivity to change following treatment in a small sample (Wainer, Berger, & Ingersoll, 2017). However, the SCC is most appropriate for young children (up to age 7) and does not assess repetitive behaviors or restricted interests. The Adapted Skillstreaming Checklist (ASC) is another study-specific rating scale designed to assess outcomes of social skills interventions for children with ASD. The ASC demonstrated good reliability and validity in a recent psychometric study, but does not asses the full range of autism symptomatology and is restricted with regard to age (6 to 12) and functional level (children without significant cognitive or language impairment) (Lopata, Rodgers, Donnelly, Thomeer, McDonald, & Volker, 2017). The new caregiver report Autism Behavior Inventory (ABI) (Bangerter et al.,

2017) was designed as an outcome tool for clinical trials and covers two core symptom domains Social Communication (27-items) and Restricted Behaviors (16 items) and three associated symptom domains. Items are rated on two of four scales (depending on the item): frequency, quality, context, or intensity. Initial psychometric properties examined in a small initial pilot sample ($n = 43$) suggest adequate test-retest reliability and concurrent validity (Bangerter et al., 2017). However, evidence of inter-rater reliability, structural validity, and sensitivity to change have not yet been reported, and to date, scores do not appear to be comparable across domains. Additionally, it is not clear whether scores will be comparable across domains, given differences in rating scale and anchors across items.

The Autism Impact Measure (AIM) is another tool developed specifically for treatment-outcome assessment in children with ASD (Kanne et al., 2014; Mazurek et al., 2020). The AIM was designed to address limitations of existing tools by focusing on treatment-relevant aspects of symptom presentation and facilitating detection of short-term improvement. For example, the AIM assesses both frequency and functional impact of symptoms, uses 5-point rather than dichotomous response options, and includes a recent (2-week) rather than longer-term reporting period. The measure is appropriate for children and adolescents across functional levels and is brief and easy to administer and score. It also provides assessment of symptoms across distinct empirically-derived domains: Repetitive Behavior, Atypical Behavior, Communication, Social Reciprocity, and Peer Interaction. The AIM has demonstrated strong test-retest and inter-rater reliability and convergent and structural validity in previous large multi-site studies (Houghton et al., 2019; Kanne et al., 2014; Mazurek et al., 2020). However, the ability of the AIM to detect symptom improvements has not yet been examined. Sensitivity to change, or responsiveness, is an important aspect of validity for measures designed for treatment outcome study (Guyatt, Kirshner, & Jaeschke, 1992) and is critical to examine in order to evaluate the AIM's utility as a treatment-outcome tool.

Current Study

The purpose of this study was to evaluate the extent to which the AIM demonstrates sensitivity to change across distinct well-established treatments for core symptoms of ASD. To our knowledge, this is the first large-scale study to systematically evaluate sensitivity to change in a measure of core autism symptoms. Using a longitudinal repeated measures design, we examined the degree to which: 1) specific AIM symptom domain scores changed over time in response to separate treatments; and 2) change in AIM domain scores was associated with change as assessed by other measures of similar constructs.

Empirical examination of sensitivity to change presents a number of different challenges. Treatment-outcome measures should be able to detect small but meaningful changes in order to inform progress monitoring and clinical decision-making. However, an individual may experience change in symptoms as a direct result of an intervention or as a result of the passage of time and natural development. In order to address this challenge, sensitivity was examined within the context of specific interventions with prior evidence of efficacy within particular symptom domains. It is important to note that the purpose of this study was not to examine treatment efficacy, but to determine the extent to which the AIM is sensitive to

change when it occurs. As such, the treatments selected were chosen because they 1) have demonstrated prior efficacy, and 2) differ in their primary symptom targets. Within each treatment group, we anticipated relatively greater gains in symptom domains specifically targeted by that treatment (described in more detail in the Methods section). Although we are unable to determine whether improvement is due to the intervention itself as opposed to other factors or the passage of time, this design was chosen to provide a context that would allow for the greatest likelihood for change to occur within specific symptom domains.

A second significant challenge to examination of sensitivity to change is the absence of a fully satisfactory criterion measure of change or improvement in core ASD symptoms. Ideally, change as assessed by the AIM should be compared to an established external standard of improvement. Unfortunately, as described above, there is no established "gold standard" measure of "true change" in ASD symptoms within or across domains, which prompted the need for such a tool in the first place. However, even in the absence of a "gold standard" criterion measure of change, a battery of concurrent measures purporting to measure similar constructs within each domain was administered in order to examine some degree of convergent evidence for change across domains. Only moderate concordance was expected, due to the fact that none of these measures were designed to assess short-term change.

Method

Participants

The sample included 471 children with ASD recruited from one of three sites: 1) University of Missouri, 2) Rady Children's Hospital San Diego, and 3) Nationwide Children's Hospital. Eligibility requirements included having a previous ASD diagnosis, meeting DSM criteria for ASD, meeting or exceeding clinical cut-off scores on the ADOS-2 at baseline, and being 2-14 years old at baseline (M = 6.8 years, SD = 3.8). Children referred for one of six treatment programs were recruited across sites, and data were collected between 2013 and 2016. The study was approved by the Institutional Review Board at each site and written informed consent was obtained from all participants. The majority of participants were male (82.2%) and non-Hispanic Caucasian (64.8%). Full Scale IQ ranged from 30 to 141 (M = 84.5, SD = 23.5). See Table S1 for sample characteristics by treatment group.

Prospective data were collected for the purposes of the current study at baseline and 6-week intervals across six different treatment modalities. The 6-week measurement interval was chosen because clinically meaningful treatment effects were expected by that point in treatment. Because the AIM was intended for use across a range of treatments, age groups, and functional levels, treatments were selected to be representative of a range of delivery formats, duration, and modalities, and all had established evidence of efficacy/effectiveness.

Treatment Groups

Improving Parents as Communication Teachers (ImPACT; n=69) is an evidence-based manualized parent-training intervention for improving social-communication skills based on developmental and naturalistic behavioral strategies (Ingersoll & Wainer, 2013). ImPACT

consists of 24 weekly/biweekly sessions (approximately 12 weeks in this project) with the parent, child, and therapist. Sessions include discussion, modeling, and practice of new strategies using interactive (e.g., modeling and expanding language, creating natural opportunities for communication) and direct methods (e.g., prompting, reinforcement, teaching/ modeling, functional behavioral strategies). ImPACT is effective in improving social communication skills among children with ASD (Ingersoll & Wainer, 2013; Stadnick, Stahmer, & Brookman-Frazee, 2015).

Positive Behavior Supports and Pivotal Response Training (PBS/PRT; n=33) is a manualized parent training program that incorporates principles of applied behavior analysis (ABA), naturalistic strategies, and pivotal response therapy (PRT) to target pragmatic language, social communication, and behavior. The duration is approximately 20 weeks, and intervention techniques focus on maximizing child motivation and responding, reinforcing positive behavior, reducing challenging behavior, and expanding the complexity of the child's skills in communication, social interaction, play, and daily living. PRT is an effective and well-established treatment for autism (National Autism Center, 2015) and naturalistic behavioral approaches are effective for improving imitation, gesture use, play, and language skills among children with ASD (Charlop-Christy & Carpenter, 2000; Ingersoll, Lewis, & Kroman, 2007).

Social Competence Intervention (SCI; n=89) is a group-based, manualized intervention program designed for children ages 6-14 with IQ ≥75, lasting approximately 15 weeks in this study (Stichter et al., 2010). For elementary-school children, SCI consists of 22 (60-minute) sessions; while the middle-school version consists of 32 (45-minute) sessions. SCI focuses on key social cognitive domains using cognitive-behavioral strategies and scaffolded instruction combining didactic instruction, behavior modeling, rehearsal, and in vivo practice with group-mates to teach and/or modify social behavior. SCI is effective in improving social behavior, interactions, and cognitive processes among youth with ASD (Stichter et al., 2010; Stichter, O'Connor, Herzog, Lierheimer, & McGhee, 2012).

Early Intensive Behavioral Intervention (EIBI; n=122) is an individualized ABA-based intervention targeting multiple developmental domains. Procedures include discrete-trial teaching, differential reinforcement, prompting, and incidental teaching. Interventions are typically delivered one-on-one, with intense (multiple hours/week) and long-term (multi-year) duration. EIBI is a well-established evidence-based treatment (National Autism Center, 2015) that is effective in improving communication and cognitive skills among children with ASD (Reichow & Wolery, 2009; Reichow, Barton, Boyd, & Hume, 2012).

Playful Learning Academy for Young Children (PLAYC; n=40) is an inclusive preschool program for children with ASD (and typically developing peers). Early Inclusion classes enroll 15 children (30-42 months-old) per class (5 with ASD), and Preschool Inclusion classes enroll 20 children (42 months to kindergarten) per class (4 with ASD). Children with ASD are generally enrolled in a half day program for 6 to 18 months. The program utilizes evidence-based treatment strategies, including naturalistic strategies, principles of applied behavior analysis, and PRT (Akshoomoff, Stahmer, Corsello, & Mahrer, 2010; Stahmer & Ingersoll, 2004).

A **Medical Treatment** (*n*=118) group was included as a comparison group without a targeted treatment for core symptoms. Participants were seen by physicians at an academic medical center specializing in medical treatment of children with autism. This group received treatment as usual, which included ongoing medical management and monitoring of common co-occurring medical and psychiatric conditions.

Design

All participants were assessed at baseline and approximately 6-week intervals over the course of treatment. Final assessments were completed at either 2-3 weeks post-treatment for short- and medium-term treatment modalities (12-20 weeks post-baseline; ImPACT, SCI, PBS/PRT) or approximately 24 weeks post-baseline for long-term ongoing treatment modalities (EIBI, PLAYC, and Medical). The full assessment battery was administered at baseline (*n*=471) and final visits (*n*=367). The AIM was administered at all assessment intervals. Given that interventions varied in length, the number of assessment time points also varied by group. Participants receiving the briefest intervention (ImPACT) were assessed at 3 time points, medium-term interventions (SCI and PBS/PRT) were assessed at 4 time points, and ongoing long-term interventions (EIBI, PLAYC and Medical) were assessed at 5 time points. As illustrated in Table 1, the mean number of weeks at each target assessment point varies somewhat both by assessment time and treatment group. As such, in subsequent analyses, we calculated true time (TIME) by setting baseline to 0 and each subsequent assessment time point as the number of weeks from baseline.

Primary Measure

The **Autism Impact Measure (AIM)** was administered at each assessment time point. The AIM is a parent-report questionnaire that includes 41 core-symptom items rated on two corresponding 5-point scales: frequency (ranging from “never” to “always”) and impact (ranging from “not at all” to “severely”) over the previous 2-weeks. AIM scores are calculated by combining frequency and impact ratings, with higher scores indicating greater symptom severity. Five empirically derived subdomain scores are generated using a subset of 29 items: Repetitive Behavior, Atypical Behavior, Communication, Social Reciprocity, and Peer Interaction (Mazurek et al., 2020). The AIM has demonstrated strong reliability and validity in previous studies (Houghton et al., 2019; Kanne et al., 2014; Mazurek et al., 2020).

Concurrent Measures

A battery of additional measures was included to examine the extent to which change as assessed by AIM domains was associated with change as assessed by measures of similar constructs.

Repetitive and Atypical Behavior Domains—Concurrent measures for AIM Repetitive Behavior and AIM Atypical Behavior domains included the Lethargy/Social Withdrawal and Stereotypic Behavior subscales from the *Aberrant Behavior Checklist (ABC)* (Aman & Singh, 1986) the *Repetitive Behavior Scale – Revised (RBS-R)* (Bodfish, Symons, Parker, & Lewis, 2000) overall and subscale scores (i.e. Compulsive, Stereotyped, Restricted, Ritualistic, Sameness, and Self-Injurious Behavior), the Restricted and Repetitive

Behavior domain score from the *Autism Diagnostic Observation Schedule – 2nd Edition (ADOS-2)* (Lord et al., 2012), the Restricted Interests/Repetitive Behavior subscale from the *Social Responsiveness Scale - 2nd Revision (SRS-2)* (Constantino & Gruber, 2012), and relevant subscales from the *Ohio Autism Clinical Global Impression Scale (OACIS)* (Butter & Mulick, 2006). Clinicians completed OACIS-Severity (OACIS-S) ratings at baseline and final time points with scores ranging from 1 (normal) to 7 (among the most severe), and then completed OACIS-Improvement (OACIS-I) ratings at the final visit as a measure of global change relative to baseline, ranging from 1 (very much improved) to 7 (very much worse). The OACIS-I Aberrant/Abnormal Behavior and Restricted/Narrow Interests subscales were examined in relation to AIM Repetitive Behavior and Restricted Interest domains.

Communication Domain—Concurrent measures for the AIM Communication domain included the *OACIS-I* Verbal and Nonverbal subscales, the *Vineland Adaptive Behavior Scales (Vineland-II)* Communication domain and subdomain scores (Sparrow, Cicchetti, & Balla, 2005), and the *MacArthur Communicative Development Inventory (CDI)* Total Gestures, Total Words Produced, and Total Words Understood subscales (Fenson et al., 1993). Because it is not appropriate for children with fluent language, the CDI was only administered to a subset of the sample with limited language ability (n = 104).

Social Reciprocity and Peer Interaction Domains—Concurrent measures for the AIM Social Reciprocity and Peer Interaction domains included the ADOS-2 Social Affect (SA) domain score, the SRS-2 Overall and subscale scores for Social Awareness, Social Communication, and Social Responsiveness, the OACIS-I Social Interaction subscale score, the Vineland-II Socialization domain and subdomain scores, and the Social Functioning Domain Score from the *Pediatric Quality of Life Inventory (PedsQL) Version 4.0* (Varni, Seid, & Kurtin, 2001).

Data Analysis

To ease interpretation, we transformed baseline AIM domain raw scores into T-scores and created a raw score look-up table aligned to these scores (Mazurek et al., 2020). This was used to convert scores for each subsequent measurement time point to this new metric (mean of 50 and *SD* of +/− 10) and allowed for examination of change estimates in terms of *SD* over time.

AIM Sensitivity to Change—To examine the nature of change over time across all individuals we focused solely on AIM domains. Growth was modeled separately for each AIM domain with SAS PROC MIXED (SAS 9.4) and the restricted maximum likelihood method (REML) for estimation. For each domain we first fit an unconditional random intercept model which served as the base for model fit comparisons of subsequent models. Incremental additions to subsequent model included to addition of: fixed linear slope, random linear slope, fixed quadratic slope, and random quadratic slope.

To select the model with best fit, the −2 log likelihood Information Criterion (AIC), Akaike Second Order Information Criterion (AICC) and Bayesian Information Criterion (BIC) were used. Generally, smaller values indicate better model fit. Given the varied spacing of

assessments across participants within treatment, we examined change over time using the time structured predictor (TIME) in lieu of measurement occasion. TIME at baseline is equal to zero for all participants and varies across participants for each subsequent assessment occasion.

Because treatment effects are not the primary focus of this paper, we focused instead on change in AIM as an indicator of AIM sensitivity to change across treatment groups. In each AIM domain model we used treatment group to examine the degree of differential change. Table 2 presents these models, by AIM domain, along with follow-up analyses of significant interactions with time. For each AIM domain model, Table 3 presents model estimated average T-score at baseline and each 6 week time point (with standard error and 95% CL estimates). Table 4 presents slope estimates at 6 and 12 weeks for each AIM domain by treatment group (significant t values indicate slopes significantly different from 0). Figures S1-S5 (available online) visually present performance trajectories over time for each treatment group by AIM domain from baseline to end of program implementation (24 weeks for EIBI, Medical, PLAYC; 20 weeks for PBS-PRT; 15 weeks for SCI; and 12 weeks for ImPACT).

AIM Change in Relation to Change in Other Measures—For the second set of analyses, we examined pre- to post-intervention change in AIM domains in relation to change over the same time period in concurrent measures. For AIM domains, we used T-scores (described above). For concurrent measures we used raw scores to calculate differences from baseline to final time point. Analyses examined the Pearson Product Moment correlations between AIM and concurrent validity measure change scores. The *p*-value was adjusted to account for the number of correlations being examined (i.e., AIM Repetitive Behavior and Atypical Behavior: $p < .004$ (.05/13 comparisons); AIM Communication: $p < .006$ (.05/8 comparisons); and AIM Peer Interaction and Social Reciprocity: $p < .005$ (.05/10 comparisons)).

Results

AIM Sensitivity to Change

Fit statistics across unconditional models indicated that across AIM domains, the best fit model contained fixed and random intercept and linear slope and fixed quadratic components. Table S2 presents the fit statistics for the base (random intercept only) and final models. These results suggest that all AIM domains are sensitive to change in core ASD symptoms over time. In Table 2, significant liner slope estimates (Time) indicate a decrease in core symptoms over time and significant quadratic slope (Time*Time) estimates indicate deceleration in the rate of change over time. For all AIM domains, except Repetitive Behavior, symptoms improved as a function of treatment group, as indicated by significant linear slope and treatment group interactions and quadratic slope by treatment group interactions. Examination of residuals plots for all models indicated no cause for concern regarding the appropriateness of these models. Table 3 presents model estimated means across time points. Table 4 presents estimated slopes at 6 and 12 weeks and their difference from 0. We selected 6 and 12 week time points to examine slope estimates as those time

points align to the length of many traditional treatments. Figures S1-S5 (available online) illustrate the model estimates of change for each AIM domain separately, as a function of treatment group. Results indicate significant slopes (change) at 6 and/or 12 weeks for most treatment groups, although differentially by AIM domain. In general, when significant change was seen in AIM behavior domain scores, the domains aligned to the focus of the specific treatment. Results in Table 4 also indicate that change is variable (in regard to onset and degree) across treatment groups and behavioral domains in respect to time, suggesting that the AIM is sensitive to shorter-term improvements in ASD symptoms across symptom domains.

AIM Change in Relation to Change in Other Measures

Correlations of change in AIM domains and change in theoretically relevant measures are shown in Table 5, and the full correlation matrix is shown in Table S3. Change in AIM Repetitive Behavior and AIM Atypical Behavior domains were significantly correlated with change in measures of similar constructs, including ABC, RBS-R, and SRS-2 subscales. Change in AIM Communication was significantly correlated with clinician-rated improvement in relevant OACIS domains. Change in AIM Peer Interaction and AIM Social Reciprocity domains were also significantly correlated with change in relevant SRS-2, PedsQL subscales and clinician-rated improvement in relevant OACIS domains. Change in the Vineland Socialization and Interpersonal domains was related to AIM Peer Interaction (although not significant after p-value adjustment; $p < .01$), but not AIM Social Reciprocity. Change in AIM scores was not correlated with change in either domain of the ADOS-2. Significant correlations were not observed between change in AIM Communication and change on either the Vineland Communication domain or the MacArthur CDI.

Discussion

Efficient and psychometrically sound treatment outcome tools are greatly needed to enhance consistency and rigor in measuring outcomes across behavioral, pharmacological, and other treatment modalities. This is the first large-scale study examining the extent to which a new measure of core ASD symptoms, the Autism Impact Measure (AIM), is sensitive to change across different types of treatments with established efficacy/effectiveness. The current results provide support for the AIM's ability to detect short-term improvement across core symptom domains. Analyses of growth over time suggest that each AIM domain is sensitive to change overall and as a function of different treatment conditions in that symptoms decreased over time, with some deceleration in rate of improvement. The relatively brief repeated assessment window (6-week intervals) highlights the AIM's utility for detecting improvements over the course of short-term treatments, which should prove useful for informing mid-treatment changes and course corrections. Another advantage of the AIM is that it provides information about response to treatment across domains in a single measure with interpretable scores that can be examined over time within the same individual and between treatment groups.

Results by symptom domain indicate that repetitive behaviors improved somewhat across all treatment groups. Because none of the selected treatments targeted repetitive behaviors,

specifically, this suggests that overall symptom improvements in this domain may have been secondary to general improvement in other behaviors and skills (the specifics of which varied by treatment modality). In contrast, all other specific symptom domains demonstrated differential improvement by treatment group. Not surprisingly, treatments that specifically focused on social communication and social interaction skills (ImPACT, PBS-PRT, and SCI) resulted in the largest magnitude of improvement in those domains. For example, the largest improvements in social reciprocity, communication, and peer interaction skills were observed in the ImPACT, PBS-PRT and SCI interventions, with little to no improvement in the interventions that did not specifically target those skills (i.e., medical follow-up and preschool programs).

Regarding the AIM correlations with other measures, results indicated some variability in how strongly change as detected by each measure was correlated with change on the AIM. The strongest correlations were observed between AIM domains and similar measures of target constructs, particularly the SRS-2, RBS-R, ABC, OACIS and PedsQL. Change detected by the AIM did not correspond strongly with change on the Vineland-2, MacArthur Communicative Development Inventory (CDI) or the ADOS-2. Regarding communication skills, it is important to note that the Vineland-2 and CDI are measures of expressive and receptive language use and skill, not ASD-specific communication impairment. Thus, it is not surprising that change in these measures did not correspond exactly to change in AIM Communication. It is interesting to note that neither the ADOS-2 Social Affect domain nor the Vineland-2 Socialization Domain correlated strongly with improvements in social skills as measured by the AIM Peer Interaction and Social Reciprocity domains. Neither of these measures was designed for repeated assessment across brief intervals, emphasizing instead diagnostic properties or functional strength-based assessment, and may not be appropriately designed for detection of short-term improvements in these skills. Because both measures are frequently used in current treatment-outcome research, these results suggest that their utility may be limited in some treatment-outcome contexts. An additional challenge with regard to the ADOS-2 is that CSS scores adjust for potential effects of language and IQ and are intended to enhance comparability across modules; however, they are likely to be less sensitive to change than raw scores. For both measures (ADOS-2 and Vineland), while using raw scores can be a limitation due to the non-interval nature of the response scale, we elected to use raw scores rather than CSS or standard scores due to the greater potential for detection of incremental change. Although ADOS-2 modules vary slightly in the number of items per subscale, this did not affect our ability to examine within-subject change in the current study, as the same module was administered at both time points for the vast majority of participants.

Across concurrent measures, observed correlations were generally small in size (even when statistically significant), suggesting that the AIM may be detecting a different type or degree of change than other widely used measures in the field. Without an existing “gold-standard” treatment outcome instrument, there is no established criterion by which to compare new instruments. Thus, it is difficult to determine which instrument, from among those examined, is detecting “true” change. An additional challenge for assessment of concurrent validity of change is that there is not a one-to-one correspondence between available measures in terms of the specific constructs assessed (Mazurek et al., 2020).

Future research is needed to determine the relative utility of the AIM in treatment-outcome research and clinical practice. The current study only included children who met or exceeded clinical cut-off scores on the ADOS-2. Thus, the extent to which the AIM is able to detect change in children with more subtle or subclinical symptomatology requires investigation in future studies. It should also be noted that the treatment groups included in the current study varied substantially in terms of sample size. Future studies of the measurement properties of the AIM across multiple treatment groups would benefit from larger and more equally-sized groups. The inclusion of measures of treatment delivery would also be helpful in order to account for variation in implementation fidelity across groups.

Another potential limitation is that the AIM may be subject to reporter bias. This is especially relevant in non-blinded observational studies, highlighting the need for rigorous clinical research designs that incorporate blinding, multi-method assessment batteries, and other methods to minimize potential bias in outcome assessment. However, parent-report provides valuable information about the occurrence and functional impact of symptoms in real-world settings, which may be different from what can be observed in a clinical or research setting. Although the AIM is intended to provide one type of assessment of the frequency and impact of ASD symptoms over time, a multi-modal, multi-informant approach to measuring improvement may provide the most comprehensive understanding of individual treatment response. Future research is needed to inform the development of corresponding clinician-report and observational tools that can efficiently and accurately assess short-term improvement in core ASD symptoms.

Conclusion

The current study provides the first large-scale examination of sensitivity to change in a measure of core ASD symptoms. Strengths of the study include a large, well-characterized sample, and examination of change over time across well-established treatments with different treatment targets, levels of intensity, and duration. Results indicate that the AIM is sensitive to short-term improvements in ASD symptoms across symptom domains. These findings provide support for the AIM's utility as an efficient and psychometrically sound treatment-outcome tool for use across different types of ASD treatments.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

The authors are extremely grateful to all the families who participated in this study. Research reported in this publication was supported by the National Institute of Mental Health of the National Institutes of Health under award number R01MH097726. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

Akshoomoff N, Stahmer AC, Corsello C, & Mahrer NE (2010). What happens next? follow-up from the children's toddler school program. Journal of Positive Behavior Interventions, 12(4), 245–253. [PubMed: 21113315]

Aman MG, & Singh NN (1986). Aberrant behavior checklist: Manual Slosson.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders, fifth edition (DSM-5). Washington, DC: American Psychiatric Association.

Anagnostou E, Jones N, Huerta M, Halladay AK, Wang P, Scahill L, … Choi D (2015). Measuring social communication behaviors as a treatment endpoint in individuals with autism spectrum disorder. Autism, 19(5), 622–636. [PubMed: 25096930]

Baker-Ericzén MJ, Brookman-Frazee L, & Brodkin ES (2018). Accelerating research on treatment and services for transition age youth and adults on the autism spectrum. Autism, doi: 10.1177/1362361317738646

Bangerter A, Ness S, Aman MG, Esbensen AJ, Goodwin MS, Dawson G, … Opler M (2017). Autism behavior inventory: A novel tool for assessing core and associated symptoms of autism spectrum disorder. Journal of Child and Adolescent Psychopharmacology, 27(9), 814–822. [PubMed: 28498053]

Bodfish JW, Symons FJ, Parker DE, & Lewis MH (2000). Varieties of repetitive behavior in autism: Comparisons to mental retardation. Journal of Autism and Developmental Disorders, 30(3), 237–243. [PubMed: 11055459]

Bolte EE, & Diehl JJ (2013). Measurement tools and target symptoms/skills used to assess treatment response for individuals with autism spectrum disorder. Journal of Autism and Developmental Disorders, 43(11), 2491–2501. [PubMed: 23479074]

Brian JA, Smith IM, Zwaigenbaum L, Roberts W, & Bryson SE (2016). The social ABCs caregiver-mediated intervention for toddlers with autism spectrum disorder: Feasibility, acceptability, and evidence of promise from a multisite study. Autism Research, 9(8), 899–912. [PubMed: 26688077]

Buescher AV, Cidav Z, Knapp M, & Mandell DS (2014). Costs of autism spectrum disorders in the united kingdom and the united states. JAMA Pediatrics, 168(8), 721–728. [PubMed: 24911948]

Butter E, & Mulick J (2006). The ohio autism clinical impressions scale (OACIS). Columbus, OH: Children's Research Institute,

Canitano R (2014). New experimental treatments for core social domain in autism spectrum disorders. Frontiers in Pediatrics, 2, 61. [PubMed: 24999471]

Charlop-Christy MH, & Carpenter MH (2000). Modified incidental teaching sessions: A procedure for parents to increase spontaneous speech in their children with autism. Journal of Positive Behavior Interventions, 2(2), 98–112.

Christensen DL, Bilder DA, Zahorodny W, Pettygrove S, Durkin MS, Fitzgerald RT, … Yeargin-Allsopp M (2016). Prevalence and characteristics of autism spectrum disorder among 4-year-old children in the autism and developmental disabilities monitoring network. Journal of Developmental and Behavioral Pediatrics : JDBP, 37(1), 1–8. doi:10.1097/DBP.0000000000000235 [doi] [PubMed: 26651088]

Constantino JN, & Gruber CP (2012). Social responsiveness scale (SRS) Western Psychological Services Torrance, CA.

Esler AN, Bal VH, Guthrie W, Wetherby A, Weismer SE, & Lord C (2015). The autism diagnostic observation schedule, toddler module: Standardized severity scores. Journal of Autism and Developmental Disorders, 45(9), 2704–2720. [PubMed: 25832801]

Estes A, Munson J, Rogers SJ, Greenson J, Winter J, & Dawson G (2015). Long-term outcomes of early intervention in 6-year-old children with autism spectrum disorder. Journal of the American Academy of Child and Adolescent Psychiatry, 54(7), 580–587. doi:10.1016/j.jaac.2015.04.005 [doi] [PubMed: 26088663]

Fenson L, Dale PS, Reznick JS, Thal D, Bates E, Hartung JP, … Reilly JS (1993). MacArthur communicative development inventories: User's guide and technical manual.

Gotham K, Pickles A, & Lord C (2009). Standardizing ADOS scores for a measure of severity in autism spectrum disorders. Journal of Autism and Developmental Disorders, 39(5), 693–705. [PubMed: 19082876]

Grzadzinski R, Carr T, Colombi C, McGuire K, Dufek S, Pickles A, & Lord C (2016). Measuring changes in social communication behaviors: Preliminary development of the brief observation of social communication change (BOSCC). Journal of Autism and Developmental Disorders, 46(7), 2464–2479. [PubMed: 27062034]

Guyatt GH, Kirshner B, & Jaeschke R (1992). Measuring health status: What are the necessary measurement properties? Journal of Clinical Epidemiology, 45(12), 1341–1345. [PubMed: 1460470]

Houghton R, Monz B, Law K, Loss G, Le Scouiller S, de Vries F, & Willgoss T (2019). Psychometric validation of the autism impact measure (AIM). Journal of Autism and Developmental Disorders, 49(6), 2559–2570. [PubMed: 30968318]

Ingersoll B, Lewis E, & Kroman E (2007). Teaching the imitation and spontaneous use of descriptive gestures in young children with autism using a naturalistic behavioral intervention. Journal of Autism and Developmental Disorders, 37(8), 1446–1456. [PubMed: 17033930]

Ingersoll B, & Wainer A (2013). Initial efficacy of project ImPACT: A parent-mediated social communication intervention for young children with ASD. Journal of Autism and Developmental Disorders, 43(12), 2943–2952. [PubMed: 23689760]

Kanne SM, Mazurek MO, Sikora D, Bellando J, Branum-Martin L, Handen B, … Warren Z (2014). The autism impact measure (AIM): Initial development of a new tool for treatment outcome measurement. Journal of Autism and Developmental Disorders, 44(1), 168–179. [PubMed: 23748386]

Kasari C, Shire S, Factor R, & McCracken C (2014). Psychosocial treatments for individuals with autism spectrum disorder across the lifespan: New developments and underlying mechanisms. Current Psychiatry Reports, 16(11), 512-014–0512-6. doi:10.1007/s11920-014-0512-6 [doi] [PubMed: 25248342]

Kitzerow J, Teufel K, Wilker C, & Freitag CM (2016). Using the brief observation of social communication change (BOSCC) to measure autism-specific development. Autism Research, 9(9), 940–950. [PubMed: 26643669]

Lopata C, Rodgers JD, Donnelly JP, Thomeer ML, McDonald CA, & Volker MA (2017). Psychometric properties of the Adapted Skillstreaming Checklist for high-functioning children with ASD. Journal of Autism and Developmental Disorders, 47(9), 2723–2732. [PubMed: 28593595]

Lord C, Rutter M, DiLavore P, Risi S, Gotham K, & Bishop S (2012). Autism diagnostic observation schedule–2nd edition (ADOS-2). Los Angeles, CA: Western Psychological Corporation,

Lord C, Rutter M, & Le Couteur A (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. Journal of Autism and Developmental Disorders, 24(5), 659–685. [PubMed: 7814313]

Mazurek MO, Carlson C, Baker-Ericzén M, Butter E, Norris M, & Kanne S (2020). Construct validity of the autism impact measure (AIM). Journal of Autism and Developmental Disorders, 50, 2307–2319. [PubMed: 29344761]

McConachie H, Parr JR, Glod M, Hanratty J, Livingstone N, Oono IP, … Charman T (2015). Systematic review of tools to measure outcomes for young children with autism spectrum disorder.

National Autism Center,. (2015). Findings and conclusions: National standards project, phase 2. . (). Randolph, MA.:

Reichow B, Barton EE, Boyd BA, & Hume K (2012). Early intensive behavioral intervention (EIBI) for young children with autism spectrum disorders (ASD). The Cochrane Library,

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(1), 23. [PubMed: 18535894]

Rogers SJ, & Vismara LA (2008). Evidence-based comprehensive treatments for early autism. Journal of Clinical Child & Adolescent Psychology, 37(1), 8–38. [PubMed: 18444052]

Scahill L, Aman MG, Lecavalier L, Halladay AK, Bishop SL, Bodfish JW, … Cook EH (2015). Measuring repetitive behaviors as a treatment endpoint in youth with autism spectrum disorder. Autism, 19(1), 38–52. [PubMed: 24259748]

Sparrow SS, Cicchetti DV, & Balla D (2005). Vineland adaptive behavior scales:(Vineland II), survey interview form/caregiver rating form. Livonia, MN: Pearson Assessments,

Stadnick NA, Stahmer A, & Brookman-Frazee L (2015). Preliminary effectiveness of project impact: A parent-mediated intervention for children with autism spectrum disorder delivered in a

community program. Journal of Autism and Developmental Disorders, 45(7), 2092–2104. [PubMed: 25633920]

Stahmer AC, & Ingersoll B (2004). Inclusive programming for toddlers with autism spectrum disorders: Outcomes from the children's toddler school. Journal of Positive Behavior Interventions, 6(2), 67–82.

Stichter JP, Herzog MJ, Visovsky K, Schmidt C, Randolph J, Schultz T, & Gage N (2010). Social competence intervention for youth with asperger syndrome and high-functioning autism: An initial investigation. Journal of Autism and Developmental Disorders, 40(9), 1067–1079. [PubMed: 20162344]

Stichter JP, O'Connor KV, Herzog MJ, Lierheimer K, & McGhee SD (2012). Social competence intervention for elementary students with aspergers syndrome and high functioning autism. Journal of Autism and Developmental Disorders, 42(3), 354–366. [PubMed: 21503797]

Tachibana Y, Miyazaki C, Ota E, Mori R, Hwang Y, Kobayashi E, … Kamio Y (2017). A systematic review and meta-analysis of comprehensive interventions for pre-school children with autism spectrum disorder (ASD). PloS One, 12(12), e0186502. [PubMed: 29211740]

Thurm A, Manwaring SS, Swineford L, & Farmer C (2015). Longitudinal study of symptom severity and language in minimally verbal children with autism. Journal of Child Psychology and Psychiatry, 56(1), 97–104. [PubMed: 24961159]

Varni JW, Seid M, & Kurtin P (2001). PedsQLTM 4.0: Reliability and validity of the pediatric quality of life InventoryTM version 4.0 generic core scales in healthy and patient populations. Medical Care, 39, 800–812. [PubMed: 11468499]

Wainer AL, Berger NI, & Ingersoll BR (2017). Brief report: The preliminary psychometric properties of the Social Communication Checklist. Journal of Autism and Developmental Disorders, 47(4), 1231–1238. [PubMed: 28105548]

Warren Z, McPheeters ML, Sathe N, Foss-Feig JH, Glasser A, & Veenstra-Vanderweele J (2011). A systematic review of early intensive intervention for autism spectrum disorders. Pediatrics, 127(5), e1303–11. doi:10.1542/peds.2011-0426 [doi] [PubMed: 21464190]

Wong C, Odom SL, Hume KA, Cox AW, Fettig A, Kucharczyk S, … Schultz TR (2015). Evidence-based practices for children, youth, and young adults with autism spectrum disorder: A comprehensive review. Journal of Autism and Developmental Disorders, 45(7), 1951–1966. [PubMed: 25578338]

Table 1.

Participants by Treatment Group and Time Point

	Assessment Time Point Sample Size (*n*) and Weeks After Baseline (*M, SD*)								
	T1/Baseline	**T2**		**T3**		**T4**		**T5**	
Treatment Group	***N***	***n***	***M (SD)***	***n***	***M (SD)***	***n***	***M (SD)***	***n***	***M (SD)***
ImPACT	69	60	6.7 (2.3)	56	14.6 (3.5)				
PBS/PRT	33	24	6.2 (1.8)	25	14.6 (2.6)	27	22.1 (3.3)		
SCI	89	77	8.0 (1.3)	43	14.4 (1.4)	68	17.4 (3.6)		
EIBI	122	103	6.6 (1.3)	105	12.6 (1.3)	96	18.6 (1.4)	96	25.9 (3.2)
PLAYC	40	36	6.7 (1.7)	37	13.1 (2.0)	37	19.0 (2.3)	36	26.3 (2.7)
Medical	118	76	7.3 (1.6)	75	13.1 (1.4)	82	19.5 (1.7)	84	26.6 (4.1)
TOTAL	471	376	6.9 (1.7)	341	13.7 (2.0)	310	19.3 (2.5)	216	26.3 (3.3)

Autism Res. Author manuscript; available in PMC 2020 December 08.

Table 2.

Type III Fixed Effects for AIM Domains Over Time

AIM Domain	Variables	Num DF	Den DF	F Value	Pr > F
Repetitive Behavior	Time	1	418	36.08	<.0001
	Time*Time	1	781	10.84	0.0010
	Treatment Group	5	781	4.81	0.0002
	Time*Treatment Group	5	781	1.97	0.0810
	Time*Time*Treatment Group	5	781	1.27	0.2735
Atypical Behavior	Time	1	420	47.19	<.0001
	Time*Time	1	792	17.35	<.0001
	Treatment Group	5	792	6.71	<.0001
	Time*Treatment Group	5	792	4.12	0.0011
	Time*Time*Treatment Group	5	792	3.29	0.0060
	Follow-Up: Time*Time*Treatment Group		***Group***	***t Value***	***Pr > |t|***
			EIBI	*18.4*	*<.0001*
			ImPACT	*54.13*	*<.0001*
			Medical	*11.11*	*0.0010*
			PBS-PRT	*0.75*	*0.3901*
			PLAYC	*0.26*	*0.6145*
			SCI	*21.86*	*<.0001*
Communication	Time	1	420	31.42	<.0001
	Time*Time	1	780	11.1	0.0009
	Treatment Group	5	780	23.97	<.0001
	Time*Treatment Group	5	780	3.87	0.0018
	Time*Time*Treatment Group	5	780	3.08	0.0092
	Follow-Up: Time*Time*Treatment Group		***Group***	***t Value***	***Pr > |t|***
			EIBI	*14.15*	*0.0002*
			ImPACT	*33.89*	*<.0001*

AIM Domain	Variables	Num DF	Den DF	F Value	Pr > F
			Medical	*4.25*	*0.0405*
			PBS-PRT	*0.03*	*0.8677*
			PLAYC	*0.01*	*0.9406*
			SCI	*15.48*	*0.0001*
Peer Interaction	Time	1	419	15.68	<.0001
	Time*Time	1	787	0.82	0.3645
	Treatment Group	5	787	6.88	<.0001
	Time*Treatment Group	5	787	2.62	0.0233
	Time*Time*Treatment Group	5	787	2.22	0.0504
	Follow-Up: Time*Time*Treatment Group		***Group***	***t Value***	***Pr > |t|***
			EIBI	*8.99*	*0.0030*
			ImPACT	*16.36*	*<.0001*
			Medical	*0.11*	*0.7455*
			PBS-PRT	*0.70*	*0.4079*
			PLAYC	*0.01*	*0.9049*
			SCI	*7.66*	*0.0066*
Social Reciprocity	Time	1	421	11.66	0.0007
	Time*Time	1	779	4.28	0.0389
	Treatment Group	5	779	5.04	0.0001
	Time*Treatment Group	5	779	2.52	0.0284
	Time*Time*Treatment Group	5	779	2.32	0.0421
	Follow-Up: Time*Time*Treatment Group		***Group***	***t Value***	***Pr > |t|***
			EIBI	*4.75*	*0.0301*
			ImPACT	*13.45*	*0.0003*
			Medical	*0.45*	*0.5036*
			PBS-PRT	*1.61*	*0.2107*
			PLAYC	*2.05*	*0.1566*
			SCI	*8.22*	*0.0050*

Autism Res. Author manuscript; available in PMC 2020 December 08.

Table 3.

Model Estimated Means and Change in AIM Domain Scores Over Time

AIM Domain	Treatment Group	Model Based Estimates (LS Mean, Standard Error, 95% CL)																			
		Baseline				6 weeks				12 weeks				18 weeks				24 weeks			
		Mean	SE	Upper	Lower	Mean	SE	Upper	Lower	Mean	SE	Upper	Lower	Mean	SE	Upper	Lower	Mean	SE	Upper	Lower
Repetitive Behavior	EIBI	**51.56**	0.63	50.33	52.79	**50.42**	0.50	49.45	51.40	**49.68**	0.72	48.27	51.09	**49.35**	0.93	47.52	51.17	**49.42**	1.24	46.99	51.85
	ImPACT	**48.85**	0.89	47.10	50.59	**46.46**	0.93	44.63	48.29	**45.05**	1.16	42.78	47.32	**44.61**	1.96	40.76	48.45	**45.13**	4.34	36.62	53.64
	Medical	**49.96**	0.65	48.68	51.25	**49.25**	0.55	48.16	50.34	**48.71**	0.79	47.15	50.27	**48.35**	1.00	46.39	50.31	**48.16**	1.28	45.65	50.67
	PBS-PRT	**50.02**	1.25	47.57	52.47	**48.13**	1.16	45.85	50.41	**46.64**	1.64	43.43	49.85	**45.56**	1.97	41.70	49.42	**44.88**	2.82	39.33	50.42
	PLAYC	**43.89**	1.09	41.74	46.03	**42.64**	0.85	40.97	44.32	**41.71**	1.22	39.32	44.11	**41.09**	1.57	38.01	44.17	**40.79**	2.07	36.73	44.84
	SCI	**47.80**	0.76	46.30	49.30	**44.66**	0.71	43.26	46.05	**43.16**	0.92	41.36	44.97	**43.31**	1.33	40.71	45.91	**45.11**	2.60	40.00	50.22
Atypical Behavior	EIBI	**50.29**	0.64	49.03	51.54	**48.92**	0.47	48.00	49.84	**48.18**	0.63	46.95	49.40	**48.06**	0.75	46.58	49.54	**48.56**	0.99	46.62	50.50
	ImPACT	**48.62**	0.90	46.86	50.38	**45.73**	0.87	44.02	47.45	**43.88**	1.03	41.87	45.89	**43.06**	1.73	39.67	46.46	**43.28**	3.99	35.44	51.12
	Medical	**51.86**	0.66	50.56	53.16	**50.17**	0.53	49.14	51.20	**49.17**	0.70	47.80	50.53	**48.84**	0.82	47.23	50.45	**49.19**	1.02	47.20	51.19
	PBS-PRT	**51.47**	1.29	48.93	54.01	**49.33**	1.10	47.17	51.48	**47.56**	1.45	44.72	50.41	**46.18**	1.61	43.03	49.33	**45.18**	2.35	40.57	49.79
	PLAYC	**42.44**	1.11	40.27	44.61	**42.60**	0.81	41.01	44.18	**42.62**	1.06	40.54	44.71	**42.53**	1.27	40.04	45.02	**42.31**	1.65	39.08	45.54
	SCI	**50.95**	0.79	49.40	52.49	**47.40**	0.68	46.06	48.75	**46.11**	0.81	44.52	47.71	**47.08**	1.12	44.88	49.28	**50.29**	2.39	45.61	54.98
Commnication	EIBI	**55.69**	0.61	54.49	56.89	**54.33**	0.47	53.41	55.25	**53.56**	0.66	52.26	54.85	**53.38**	0.83	51.74	55.01	**53.79**	1.10	51.62	55.95
	ImPACT	**52.55**	0.87	50.85	54.26	**51.41**	0.88	49.68	53.15	**50.44**	1.07	48.34	52.54	**49.62**	1.81	46.07	53.18	**48.97**	4.09	40.95	56.99
	Medical	**46.40**	0.64	45.15	47.65	**45.40**	0.53	44.37	46.44	**44.88**	0.73	43.44	46.32	**44.84**	0.90	43.07	46.60	**45.28**	1.14	43.04	47.51
	PBS-PRT	**51.78**	1.30	49.23	54.33	**49.94**	1.11	47.76	52.13	**48.31**	1.52	45.32	51.30	**46.88**	1.77	43.42	50.34	**45.65**	2.54	40.68	50.63
	PLAYC	**47.88**	1.06	45.81	49.96	**47.47**	0.81	45.89	49.05	**47.06**	1.12	44.87	49.26	**46.67**	1.40	43.92	49.42	**46.29**	1.84	42.68	49.90
	SCI	**43.72**	0.76	42.24	45.21	**39.99**	0.68	38.66	41.32	**38.50**	0.85	36.83	40.16	**39.26**	1.20	36.90	41.62	**42.27**	2.44	37.48	47.05
Peer Interaction	EIBI	**53.66**	0.68	52.34	54.99	**51.86**	0.51	50.86	52.86	**50.58**	0.70	49.22	51.95	**49.85**	0.86	48.16	51.54	**49.65**	1.15	47.40	51.90
	ImPACT	**50.37**	0.96	48.50	52.25	**47.55**	0.94	45.70	49.39	**44.49**	1.12	42.29	46.70	**41.21**	1.92	37.45	44.97	**37.70**	4.36	29.14	46.26
	Medical	**48.70**	0.71	47.31	50.09	**47.97**	0.57	46.85	49.09	**47.36**	0.78	45.84	48.88	**46.87**	0.93	45.03	48.70	**46.49**	1.17	44.19	48.80
	PBS-PRT	**50.06**	1.37	47.37	52.75	**49.32**	1.19	46.98	51.66	**48.48**	1.63	45.29	51.67	**47.54**	1.87	43.88	51.21	**46.51**	2.70	41.22	51.80
	PLAYC	**44.20**	1.17	41.90	46.50	**43.92**	0.87	42.21	45.63	**43.75**	1.18	41.44	46.07	**43.71**	1.45	40.87	46.56	**43.79**	1.89	40.08	47.50
	SCI	**48.56**	0.83	46.93	50.20	**45.67**	0.74	44.23	47.12	**44.29**	0.90	42.53	46.06	**44.42**	1.26	41.95	46.89	**46.06**	2.61	40.93	51.19
Social Reciprocity	EIBI	**52.62**	0.66	51.32	53.93	**51.44**	0.50	50.46	52.43	**50.75**	0.69	49.40	52.11	**50.56**	0.86	48.88	52.24	**50.85**	1.13	48.63	53.07
	ImPACT	**50.46**	0.93	48.64	52.28	**49.37**	0.93	47.53	51.20	**48.96**	1.13	46.74	51.18	**49.23**	1.93	45.45	53.02	**50.19**	4.35	41.65	58.74

AIM Domain	Treatment Group	Model Based Estimates (LS Mean, Standard Error, 95% CL)																			
		Baseline				6 weeks				12 weeks				18 weeks				24 weeks			
		Mean	SE	Upper	Lower	Mean	SE	Upper	Lower	Mean	SE	Upper	Lower	Mean	SE	Upper	Lower	Mean	SE	Upper	Lower
	Medical	**48.62**	0.69	47.26	49.98	**47.94**	0.56	46.84	49.04	**47.53**	0.76	46.03	49.03	**47.38**	0.92	45.57	49.20	**47.50**	1.16	45.21	49.78
	PBS-PRT	**50.10**	1.34	47.46	52.73	**50.56**	1.18	48.24	52.89	**49.90**	1.59	46.78	53.03	**48.12**	1.81	44.56	51.68	**45.20**	2.63	40.05	50.36
	PLAYC	**46.07**	1.16	43.80	48.34	**45.22**	0.86	43.53	46.90	**44.93**	1.17	42.64	47.22	**45.21**	1.44	42.39	48.04	**46.07**	1.88	42.37	49.76
	SCI	**47.26**	0.81	45.67	48.86	**44.06**	0.72	42.65	45.48	**42.72**	0.89	40.97	44.46	**43.22**	1.25	40.77	45.67	**45.57**	2.58	40.52	50.62

Note: Shaded = Post treatment estimates. SE will typically be higher and CI broader for estimates as they move further away from actual data point. The values marked in bold indicate p < .0001.

Table 4.

Model Estimated Slopes at 6 and 12 weeks by AIM Domain and Treatment Group

AIM Domain	Treatment Group	Model Estimated Slopes											
		6 weeks						12 weeks					
		Slope	SE	t	Pr > |t|	Upper	Lower	Slope	SE	t	Pr > |t|	Upper	Lower
Repetitive Behavior	EIBI	−0.16	0.09	−1.79	0.0742	−0.33	0.02	−0.08	0.05	−1.74	0.0830	−0.18	0.01
	ImPACT	−0.32	0.14	−2.20	0.0277	−0.60	−0.04	−0.11	0.18	−0.65	0.5163	−0.46	0.23
	Medical	−0.10	0.09	−1.12	0.2610	−0.28	0.08	−0.08	0.05	−1.53	0.1265	−0.17	0.02
	PBS-PRT	−0.25	0.19	−1.27	0.2042	−0.63	0.13	−0.20	0.11	−1.85	0.0644	−0.41	0.01
	PLAYC	−0.21	0.15	−1.41	0.1578	−0.51	0.08	−0.09	0.08	−1.14	0.2530	−0.25	0.07
	SCI	−0.37	0.11	−3.20	0.0014	−0.59	−0.14	−0.20	0.11	−1.85	0.0641	−0.40	0.01
Atypical Behavior	EIBI	−0.18	0.08	−2.32	0.0205	−0.34	−0.03	−0.06	0.04	−1.37	0.1706	−0.15	0.03
	ImPACT	−0.40	0.13	−3.05	0.0023	−0.66	−0.14	−0.15	0.16	−0.95	0.3418	−0.47	0.16
	Medical	−0.21	0.08	−2.58	0.0099	−0.37	−0.05	−0.11	0.05	−2.52	0.0120	−0.20	−0.03
	PBS-PRT	−0.27	0.18	−1.52	0.1276	−0.62	0.08	−0.28	0.10	−2.83	0.0048	−0.47	−0.09
	PLAYC	−0.03	0.14	−0.24	0.8094	−0.30	0.23	0.00	0.07	0.04	0.9677	−0.14	0.15
	SCI	−0.38	0.11	−3.56	0.0004	−0.58	−0.17	−0.07	0.10	−0.77	0.4437	−0.26	0.12
Commnication	EIBI	−0.18	0.08	−2.15	0.0317	−0.34	−0.02	−0.07	0.04	−1.57	0.1163	−0.16	0.02
	ImPACT	−0.18	0.14	−1.29	0.1958	−0.44	0.09	−0.17	0.16	−1.01	0.3149	−0.49	0.16
	Medical	−0.14	0.08	−1.70	0.0903	−0.31	0.02	−0.07	0.05	−1.49	0.1370	−0.16	0.02
	PBS-PRT	−0.30	0.19	−1.58	0.1143	−0.67	0.07	−0.24	0.10	−2.28	0.0226	−0.44	−0.03
	PLAYC	−0.09	0.14	−0.65	0.5136	−0.37	0.18	−0.07	0.08	−0.87	0.3866	−0.21	0.08
	SCI	−0.42	0.11	−3.90	0.0001	−0.63	−0.21	−0.09	0.10	−0.89	0.3714	−0.28	0.11
Peer Interaction	EIBI	−0.29	0.09	−3.29	0.0010	−0.46	−0.12	−0.16	0.05	−3.31	0.0009	−0.25	−0.06
	ImPACT	−0.51	0.14	−3.53	0.0004	−0.79	−0.22	−0.43	0.17	−2.47	0.0137	−0.77	−0.09
	Medical	−0.15	0.09	−1.67	0.0945	−0.33	0.03	−0.10	0.05	−1.91	0.0560	−0.19	0.00
	PBS-PRT	−0.10	0.20	−0.49	0.6265	−0.48	0.29	−0.16	0.11	−1.50	0.1343	−0.37	0.05
	PLAYC	−0.07	0.15	−0.44	0.6611	−0.36	0.23	−0.02	0.08	−0.26	0.7923	−0.18	0.14
	SCI	−0.35	0.12	−3.05	0.0023	−0.58	−0.13	−0.07	0.11	−0.71	0.4797	−0.28	0.13
Social Reciprocity	EIBI	−0.16	0.09	−1.84	0.0656	−0.33	0.01	−0.06	0.05	−1.25	0.2122	−0.15	0.03
	ImPACT	−0.10	0.14	−0.71	0.4791	−0.38	0.18	−0.07	0.17	−0.40	0.6874	−0.41	0.27

AIM Domain	Treatment Group	Model Estimated Slopes										
		6 weeks						12 weeks				
		Slope	SE	t	Pr > |t|	Upper	Lower	Slope	SE	t	Pr > |t|	Upper	Lower
	Medical	−0.10	0.09	−1.17	0.2410	−0.28	0.07	−0.06	0.05	−1.27	0.2032	−0.16	0.03
	PBS-PRT	0.00	0.19	0.02	0.9836	−0.37	0.38	−0.21	0.11	−1.97	0.0496	−0.42	0.00
	PLAYC	−0.11	0.15	−0.76	0.4472	−0.40	0.18	0.01	0.08	0.11	0.9156	−0.15	0.16
	SCI	−0.38	0.11	−3.33	0.0009	−0.60	−0.15	−0.11	0.10	−1.05	0.2918	−0.31	0.09

Table 5.

Correlations: Change in AIM Domains and Change in Concurrent Measures

Measure	Subscale	N	RRB	AB	C	PI	SR
Aberrant Behavior Checklist (ABC)	Lethargy/Withdrawal	340	0.19*	0.25**			
	Stereotypic	340	0.14+	0.12			
Repetitive Behavior Scale – Revised (RBS-R)	Compulsive	340	0.22**	0.16*			
	Stereotyped	341	0.22**	0.16*			
	Restricted	332	0.19*	0.14			
	Ritualistic	341	0.18*	0.20*			
	Sameness	341	0.12	0.20*			
	Overall Score	341	0.26**	0.26**			
Autism Diagnostic Observation Schedule – 2nd Edition (ADOS-2)	Restricted/Repetitive	349	0.01	−0.01			
	Social Affect	349				0.003	0.06
Social Responsiveness Scale-2 (SRS-2)	Restricted/Repetitive	338	0.28**	0.43**			
	Social Awareness	338				0.09	0.19*
	Social Communication	338				0.29**	0.31**
	Social Responsiveness	338				0.31**	0.33**
Ohio Autism Clinical Global Impression Scale-Improvement (OACIS-I)	Aberrant/Abnormal	338	0.15*	0.16*			
	Restricted Interests	338	0.09	0.12			
	Verbal Communication	338			0.13+		
	Nonverbal Comm.	338			0.10		
Vineland Adaptive Behavior Scales-2	Communication	291			−0.12		
	Receptive	286			−0.08		
	Expressive	276			−0.18*		
	Socialization	284				−0.10	−0.13
	Interpersonal	273				−0.12	−0.13
	Play/Leisure	279				−0.09	−0.06
	Coping Skills	279				−0.06	−0.13

Measure	Subscale	N	RRB	AB	C	PI	SR
MacArthur Communicative Development Inventory (CDI)	Total Gestures	80			−0.20		
	Total Words Produced	80			−0.22		
	Total Words Understood	80			−0.17		
Pediatric Quality of Life Inventory V4 (PedsQL)	Social Functioning	346				−0.33 **	−0.19 **

RRB = AIM Repetitive Behavior domain, AB = AIM Atypical Behavior domain, C = AIM Communication domain, PR = AIM Peer Interaction domain, SR = AIM Social Reciprocity domain

For RRB/AB:

** p = <.0001

* p = < .004

+ p = < .01

for C:

** p = <.0001

* p = < .006

+ p = < .01

for PI/SR:

** p = <.0001

* p = < .004

+ p = < .01