Distinct patterns of treatment response using the Early Start Denver Model with young children with autism spectrum disorder: a cluster analysis

Godel Michel*[1] • Robain François*[1], Kojovic Nada[1], Franchini Martina[1], Wood de Wilde Hilary[1], Schaer Marie[1]

*Equal contribution

[1] Department of Psychiatry, University of Geneva School of Medicine, Geneva, Switzerland

Abstract

Evidence-based early intervention approaches have been shown to significantly improve many areas of development in young children with autism. Despite having an overall positive impact for most children, there is high inter-individual heterogeneity of response to treatment, and it is currently not possible to predict which child will benefit from which specific intervention. In this study, we analyze developmental trajectories of young children with autism who received two-years of Early Start Denver Model intervention in Geneva, Switzerland (n=47). Using cluster analyses, we distinguish between 3 subgroups based on cognitive level at baseline and rates of change over time. The first subgroup of children had cognitive scores at baseline only slightly below what would be expected for their age and were found to have nearly no cognitive delay by the end of treatment (High at Baseline: HB). The children in the two other subgroups all presented with severe cognitive delay at baseline, with the second group's cognitive scores improving significantly in the first year of intervention, along with increased communication and daily living skills (High Responders: HR). The third subgroup showed little to no cognitive level change over the course of treatment (Minimal Responders: MR). Furthermore, the rates of change in cognition and adaptative behavior in the HR subgroup were already significantly higher compared to the MR subgroup within the first 6 months of intervention. A rapid and early response to intervention might thus represent a predictor of optimal outcome. Inversely, slower progress by the sixth month of intervention may predict that the child will have a slower response to the treatment overall and may need alternative supports to reach their learning objectives quickly.

Keywords: Autism Spectrum Disorders, Early intervention, Predictors of treatment response, Heterogeneity, Minimal responder

Introduction

Autism spectrum disorders (ASD) are characterized by the presence of restricted and repetitive behaviors along with impairments in communication and social interactions (1). Over the last decades, many types of comprehensive, evidence-based early interventions have been developed for children with ASD to improve social-communication, cognitive functioning and adaptive skills with the ultimate goal of increasing lifelong autonomy (2–7). Despite a common objective to alleviate the difficulties of children with ASD, all early intensive intervention approaches result in a very heterogeneous response between children (8,9). Whereas some children show significant improvements after the intervention, others only display a minimal response. This finding constituted the starting point of a discussion about the definition of “response to treatment” (RTT) , and how early in the intervention process should a child be qualified as being "minimal responders" (10). Despite all the research performed to identify responders’ characteristics, it is currently impossible to predict to what extent a child will respond to an early intervention approach based on their behavioral characteristics at intake (10). Thus, early intervention is generally recommended for all young children diagnosed with ASD (11) independently from their specific behavioral or symptom profile.

During the last decade, the Early Start Denver Model (ESDM, (2)) has emerged as a promising early intervention in the framework of Naturalistic Developmental Behavioral Interventions (NDBI) notably because of its effectiveness, its natural environment implementation and individualized learning objectives (7). A seminal randomized controlled trial (RCT, (2)) reported a mean increase of 17 IQ points in a sample of 24 toddlers receiving the ESDM intervention. Some studies have replicated these results (for a review see (12)) and put forward the cost-effectiveness of such an intervention (13,14). Overall, ESDM intervention significantly increases cognitive and language skills but is equivalent to other types of early interventions in terms of changing ASD symptoms (12). However, literature about response to treatment in early intervention let alone a specific approach such as ESDM remains sparse, and most studies published to date comprised a limited sample size. Indeed, to the best of our knowledge, no single site study using ESDM with a 1:1 child/therapist ratio for more than 12 hours per week and for a 2 years period exceeded the sample size of the original RCT (2) to date. Two other studies achieved larger samples, respectively of 42 (15) and 55 children with ASD receiving ESDM intervention (15), using a multicenter study design. As a result, there is scarce knowledge about the ESDM response predictors. Indeed, despite the undeniable positive outcome of ESDM when considering groups of children, the inter-individual variability in the response is high, as for all types of early interventions (16). To better understand this heterogeneity in the outcome, authors like Vivanti et al. (17) attempted to identify predictors of response specific to the ESDM. Their study showed that developmental gains after one year of treatment were best predicted by higher imitation skills, goal understanding and more advanced skills in

functional use of objects before the intervention start. However, outcome measures were collected after only one year of intervention. Also, baseline measures used in this study were based on original tasks making it difficult to compare them with other studies. Therefore, to date, there are no identified characteristics or profiles based on habitual/common clinical assessments, which could predict the response to the ESDM intervention. Nevertheless, given the considerable investment of time and resources that the intervention represents for therapists and families (13)(18), identifying these predictors is important. The identification of intervention-specific, rather than general, predictors could allow clinicians to refer children to a certain type of intervention based on their profile, thus increasing the likelihood of positive treatment response.

Given the absence of evidence about the superiority of one early intervention approach over the others, and a lack of solid predictors of response to specific treatment models, the currently available guidelines adopt a "one size fits all" therapeutic recommendation. However, in the global framework of precision medicine (19), there is an urge to develop more individualized guidelines for intervention in ASD. One way of promoting individualized interventions relies on identifying early predictors of response, i.e. patterns of behavioral and/or demographic characteristics displayed before the start of the intervention associated with a positive outcome after the intervention completion (10). Until now, methodological strategies to achieve this goal include whole-sample correlations between baseline and outcome measures (20)(21), comparison between subgroups defined on the basis of an arbitrary cut-off such as rapid *vs* slow learners (22) or best *vs* non-best outcome (23)(24). A promising alternative relies on/upon the identification of distinct phenotypic subgroups within ASD (25). This can be achieved by applying cluster analyses (CA), a strategy that has already been applied to samples of preschoolers with ASD to investigate ASD heterogeneity (for a review see 26). Recently Frazier et al. (27) used CA to identify subgroups of language development during an applied behavioral analysis (ABA) intervention. However, no study to date has applied CA on a sample of preschoolers who received early intensive intervention to define data-driven subgroups based on developmental or symptom levels at intervention start and rates of change over time.

As highlighted above, identifying intervention-specific predictors of response has major clinical implications. One of the most relevant ways to explore predictors of outcome lies in subgroup phenotyping. In the present study, we explore the developmental trajectories of 47 preschoolers with ASD who completed two years of intensive ESDM intervention in our program in Geneva, Switzerland. We first investigated if our sample's outcome, in terms of cognition, symptom severity and adaptive functioning, reproduced those described in the literature. Then, we explored the heterogeneity in the response to intervention within our sample by using cognition as our main outcome. CA highlighted three different subgroups based on cognitive trajectories during the period of intervention. We further explored baseline differences between the subgroups to identify potential predictors of cognitive change over time. Finally, we investigated for differences in the early rates of change between subgroups (first six months of treatment) to look at the early response to

intervention as potential predictors of treatment outcome (after completing the two-year intervention).

Method

Participants

Our original sample included 52 participants who completed 2-years of ESDM intervention in Geneva, Switzerland. Four participants were not included in the analyses because of missing data regarding their developmental assessment at baseline and one participant because of missing data at the end of the intervention. Our analyses where thus based on the data collected from 47 participants (See Table 1). All children were referred to the intervention program after receiving a clinical diagnosis of ASD according to Diagnostic and Statistical Manual of mental disorders, 5th edition (28) criteria and Autism Diagnosis Observation Schedule-Generic (ADOS-G) (29) or ADOS-2 (30) diagnosis cut-offs. Inclusion criteria to the therapeutic program comprised a confirmed diagnosis of ASD, having two complete years before the age of entrance to school (4 years old in Geneva) and parental consent. The enrollment process was based on the room availability at the time of referral and there was no restriction based on the parental socio-economic status or the child's developmental level. The parents of all participants gave their written consent to the research protocol that was approved by the institutional review board of the University of Geneva. All participants belong to the ongoing longitudinal Geneva Autism Cohort. From this sample, 22 children were already included in a previous study measuring outcome after one year of ESDM intervention (31). Baseline evaluations were performed at the start of the intervention and comprised behavioral measures that are detailed below. Parents also filled questionnaires about medical history (including pregnancy, delivery, and drug history), as well as demographic information that are detailed below. Children were then assessed after 6, 12 and 24 months of therapy, resulting in a total of 4 visits. Ultimately, post intervention schooling information were collected as well as the amount of support provided there. Types of schooling and educational support comprised regular public schooling without any specialized support; regular public schooling with specialized educational support (a psychologist present in the classroom); private schooling with its specific special needs programs; public special education schooling with intensive support.

Intervention

The 52 participants were enrolled in one of the 4 units of the Centre d'Intervention Précoce en Autisme in Geneva, Switzerland (CIPA, Center for Early Intervention in Autism), where they received 20 hours a week of daily, individual and group intervention sessions using the Early Start Denver Model (ESDM). The ESDM is a comprehensive, evidence-based, early intervention approach that promotes child learning through naturalistic, developmental, and behavioral techniques (2,7). Parents of the participants were provided with 12 hours of once-a-week parent coaching sessions in the use of the ESDM model at the start of their child's program. The children were evaluated every 3 months using the Early Start Denver Model Curriculum Checklist for Young Children with Autism (ESDM-CC) to establish targeted and measurable learning objectives. The intervention services were provided by graduate level therapists, who were trained within the CIPA program in the use of the ESDM

approach, meeting ESDM fidelity on the ESDM Fidelity Rating System (32). Today, the team consists of 20 credentialed ESDM therapists, and the program is overseen by an ESDM certified trainer.

Measures

The ADOS, which refers to one of the two recent versions the ADOS-G and its later version ADOS-2, is a standardized assessment which comprises series of semi-structured social presses aimed to elicit and measure of ASD symptoms (30,33). The schedule comprises 5 different modules (30), adapted to the child's age and level of language. The calibrated severity score (CSS) was used to compare the total severity score as well as the restricted and repetitive behaviors (RRB) and social affect (SA) symptoms severity scores (34,35). The ADOS-G/ADOS-2 were performed by a trained examiner and videotaped. The members of the team who rated the videotapes were not implicated in the delivery of the ESDM intervention.

The Mullen Scales of Early Learning (MSEL) is a standardized assessment for children aged from birth to 68 months (36). It measures the child's development in four developmental domains: expressive language (EL), receptive language (RL), visual reception (VR), fine motor (FM), as well as gross motor skills (GM).

The Psychoeducational Profile – third edition (PEP-3) is a standardized assessment tool that evaluates various cognitive, motor and adaptative domains in children from 6 months to 7 years of age (37). These domains include EL, RL, cognitive verbal and preverbal (CVP) and FM.

Developmental quotient scores (DQ) were computed for each subdomain of the MSEL by dividing the individual developmental age by the chronological age and multiplying by 100 as described in (38). The composite DQ was computed by computing the average of all four subdomains developmental ages then divided by the chronological age and multiplying by 100. Similarly, DQ were computed for the subdomains of the PEP-3 that assess cognitive domains that are equivalent to the ones of the MSEL, namely EL, RL, CVP and FM. Composite DQ was derived using the same formula as described for the MSEL. DQ have already been used for the PEP-3 subdomains in previous literature (31). For our analyses of cognitive skills, we used the Early Learning Composite ELC DQ obtained with the MSEL. As the MSEL was not administered for some participants (n = 7 at baseline, n= 7 after 6 months of therapy, n=3 after 12 months of therapy, n=2 after 24 months of therapy) we replaced the missing DQ by their equivalent DQ derived from the PEP-3. It is important to keep in mind that DQ is normalized for the age at the time of evaluation. Hence, a DQ that remains stable over time does not reflect a stagnation but rather a progress in the acquisitions. Also, a loss of DQ over time does not necessarily means a loss of acquisitions (or regression) but rather slower acquisitions leading to a widening of the gap compared to typical development.

The Vineland Adaptive Behavior Scales – 2nd edition (VABS-II) is a parental questionnaire that was administered by a trained clinician to assess the child's adaptive behavior (39). The domains assessed comprise communication, socialization, daily living skills (DLS) and motor

skills. Furthermore, an overall adaptive behavior composite score (ABC) of all these domains is computed.

The ADOS, VABS-II, PEP-3 and MSEL were administered at baseline, after 12 months and after 24 months of therapy. Assessment at six months only comprised the VABS-II and the MSEL.

We measured the parental socio-economic status of the participants using the total household yearly income and the highest level of education achieved by parents at baseline. The household income was divided into three subgroups that are detailed in Table 1. Parental educational level was first coded using the seven categories of the four factor index of social status developed by Hollingshead (40). We then parsed these categories into two groups: 1) elementary school or high school completed, and 2) college degree or PhD completed.

Rate of change

For all behavioral measures acquired longitudinally (ADOS, the VABS-II and the DQ), we computed an individual rate of change using the following symmetrized percentage change (SPC) formula:

$$SPC\ [\%/year] = 100 \times \frac{(B_y - B_x)/[(B_x + B_y)/2]}{(age_y - age_x)}$$

Where B_x and B_y represent the behavioral measure acquired when the participant was aged of age_x and age_y respectively. In other words, SPC is the behavioral difference between two timepoints relatively to the mean of the scores across these two timepoints, then divided by the time interval (in years). This results in a yearly rate of change that can be expressed as a percentage when multiplied by 100. The main advantages of using symmetrized measures of change over absolute differences (such as $B_y - B_x$) or non-symmetrized percentages (such as $(B_y - B_x)\ /\ B_x$) comprise increased statistical robustness, higher reliability in small samples and balanced consideration of both measures B_x and B_y (41).

Statistical Analyses

IBM®SPSS® Statistics v26.0.0.0 for macOs (Armonk, NY: IBM Corp.) was used for all analyses. Statistical significance threshold was set at alpha = 0.05. Graphs were obtained with Prism® v8.3.0 (GraphPad Software, La Jolla California USA, www.graphpad.com) and Matlab R2018b for MacOs (MathWorks).

To test for an effect of time a repeated measure ANOVA was performed on the whole sample for each longitudinal behavioral measure using the scores collected at baseline, 12 months, and 24 months after the start of the therapy. Greenhouse-Geisser correction was applied whenever the assumption of sphericity was violated according to Mauchly test.

Then, we performed a cluster analysis (CA) using the DQ as our main outcome. There are several reasons to choose DQ over other parameters as the main outcome measure of an early intervention. First, it is generally the main outcome reported in the different studies and also the one that displays the most variability (42) (43) (44). Second, cognition has been shown to be the domain that improves the most after an early intervention (45). Third, studies investigating possible ASD subtypes within ASD showed that the most salient group differences emerge when categorized using cognitive skills (46). We used a *k*-means clustering approach to identify subgroups in terms of DQ trajectories with a maximal number of iterations set to 10 (47). We chose two variables that capture individual DQ trajectories: the composite DQ at baseline and the composite DQ SPC over the two-year intervention program. To objectively determine the number of clusters *k* we used a two-step clustering approach as suggested by Kodinariya (48). We used the two-step clustering algorithm developed by Chiu et al. 2001 (49) as it is implemented in IBM®SPSS® Statistics. Briefly, this method firstly divides the sample into a set of sub-clusters through a sequential approach and secondly merges the sub-clusters through a hierarchical technique based on the log-likelihood distance between them. Ultimately, the Akaike's information criterion is used to objectively determine the optimal number of clusters.

The two-step cluster analysis (CA) yielded 3 optimal clusters based on the baseline composite DQ and the composite DQ SPC over 2 years (Fig. 1). The ANOVA revealed that one of these clusters exhibited significantly greater composite DQ at baseline compared to the others and was thus named "high at baseline" (HB, n = 18). Its average DQ at baseline was 78.8 ± 10.9 with a range comprised between 64 and 108 with a SPC of 10.2 ± 5.8 %/yr. This corresponds to an average 19.2 gain for a final DQ of 98.0 ± 14.7 with a range comprised between 64 and 124. The second cluster was characterized by high rates of progress within the two-year program and was thus named "high responders" (HR, n = 20). DQ at baseline was 50.0 ± 10.8 with a range comprised between 22 and 67, and its average SPC was 23.8 ± 8.4 %/year. This corresponds to an average 33.9 gain for a final DQ of 83.9 ± 21.8 with a range comprised between 33 and 130. The third cluster was characterized by decreased rates of progress compared to the two other clusters with an average SPC of -10.3 ± 11.3 %/yr. Its composite DQ at baseline was 46.0 ± 8.3 with a range comprised between 32 and 60. The average loss was 8.5 for a final DQ of 37.5 ± 8.9 with a range comprised between 27 and 58. We labelled this cluster "minimal responders" (MR, n = 9). HR and MR did not differ in composite DQ at baseline, with an average of 50.0 and 46.0 respectively. Together, they form a group of Low functioning children (LF). Clusters differences over composite DQ at baseline and composite DQ SPC are illustrated on Fig 1. Detailed analyses are reported in supplementary table 1 and 2.

Figure 1: A. Composite DQ trajectory of the total sample over the two years of intervention. Significant results of repeated measure ANOVA are displayed. **B.** Composite DQ trajectories of the three subgroups parsed by the cluster analysis. **C.** Individual values of the two measures used in the clustering analysis algorithm (composite DQ at baseline and DQ SPC over the two years of interventions). Color code represents the cluster membership of each participant after the application of the cluster analysis. **D.** Differences between the three subgroups on the two measures that were used to parse them. ***p<.001. DQ: Developmental Quotient. SPC: Symmetrized Percent Change.

Demographic, socio-economic measures and behavioral measures at baseline were compared between clusters using one-way analysis of variance (ANOVA) or chi-square test. We used a Bonferroni correction for multiple testing on the subdomains of a same clinical evaluation (e.g., the subdomains of the VABS-II), setting the statistical significance at 0.05/number of subdomains. When an ANOVA reached statistical significance, post-hoc comparisons between clusters were performed using multiple T-test with Bonferroni correction and statistical significance set at 0.05/number of clusters.

We then applied the same strategy to compare the SPC between clusters. We performed analyses on the following SPC: from baseline to 6 months, from baseline to 12 months, from baseline to 24 months of therapy.

Ultimately, we focused on the two LF clusters which showed no differences in their composite DQ at baseline to explore whether any other behavioral measure could help classifying them. To do so, we used binary logistic regression models. More specifically, we selected all behavioral measures that differed between HR and MR on post-hoc T-tests at baseline. Then, we performed a multivariate logistic regression using the selected measures. Whenever a composite score as well as one or more subdomain scores of the same test were selected, we preferred the composite measure to minimize potential collinearity between variables in the model. Then, we used the same strategy for the SPC measures during the 6 first months of intervention, and ultimately with those of the 12 first months.

Results

Whole sample trajectories

Descriptive measures collected at each visit are reported in Table 1 for the total sample. The children were aged from 15.3 to 42.0 months at the beginning of the intervention (average: 28.9 ± 5.3 months). The average composite DQ of the entire group at baseline was 60.3 ± 18.0 (range: 22 - 108). As a group, all 47 children receiving ESDM showed a significant decrease in their total level of symptom severity (ADOS CSS) (See Table 1). This improvement was driven by a decrease in the social affect (SA) domain. On the contrary, the RRB symptom severity increased over time. We found that these changes occurred mainly during the first year of intervention and that CSS were stable during second year of intervention. In parallel, participants partially alleviated their cognitive delay. This improvement was significant in all cognitive subdomains (i.e. VR, RL and EL) with the exception of fine motor skills (FM). As for the measures of symptom severity, all changes in cognition were significant during the first year of therapy but not the second one. Finally, there was no significant change regarding the composite adaptive behavior scores over time. Nevertheless, participants made significant gains in the subdomain of communication which occurred both during the first and the second year of intervention. Besides, participants increased their delay in the motor skills subdomain compared to typical development. This effect on motor skills was only significant during the first year of intervention and its effect size was the smallest (see Table 1) of all the mentioned significant behavioral changes. All statistically significant results are detailed in Table 1. Concerning the type of schooling after the intervention, 29 participants (61.7%) joined a regular public-school classroom with an in-class specialized educational support. One participant (2.1%) joined regular public-school without any support. Four

children (8.5%) joined a private school that provided an in-class special support. Finally, 13 participants (27.7%) entered a special education public-school.

Table 1 Sample characteristics over the two years of ESDM intervention*; *p<.05 ** p<.01 ***p<.001.* G*: Greenhouse-Geisser correction applied , ESDM = Early Start Denver Model*

MEASURE	At Baseline	*+ 6 months*	+ 12 months	+24 months	Pval (R.M. ANOVA)	Partial Eta Squared	0 - 24mo	0 - 12mo	12mo - 24mo
Clinical Description									
ADOS CSS total [Mean (SD)]	**8.0 (1.9)**		**6.5 (2.0)**	**7.0 (1.9)**	**<0.001*** **	.241	**.011***	**<0.001*** **	.055
ADOS CSS SA	**7.7 (2.0)**		**5.32 (1.8)**	**5.8 (1.9)**	**<0.001*** (G)**	.365	**<0.001*** **	**<0.001*** **	.262
ADOS CSS RRB	**7.9 (2.1)**		**9.0 (1.4)**	**9.0 (1.3)**	**<0.001*** (G)**	.199	**.003** **	**.001** **	1.0
ADI-R subdomains [Mean (SD)] (n = 43)									
ADI-R Social interactions	13.7 (4.9)								
ADI-R RRB	4.0 (2.3)								
VABS-II Adaptative Behavior Composite [Mean (SD)]	79.8 (9.1)	*81.0 (12.2)*	80.1 (12.6)	82.6 (15.5)	.102 (G)				
VABS-II socialization	81.1 (9.2)	*81.0 (10.5)*	78.6 (10.5)	79.5 (12.9)	.201 (G)				
VABS-II communication	**75.2 (12.1)**	*80.2 (15.4)*	**83.7 (18.5)**	**87.7 (18.5)**	**<0.001*** (G)**	.359	**<0.001*** **	**<0.001*** **	**.009** **
VABS-II daily living skills	84.0 (10.8)	*84.9 (11.4)*	82.7 (12.4)	83.1 (16.5)	.569				
VABS-II motor skills	**90.5 (9.6)**	*89.4 (11.1)*	**85.1 (10.3)**	**88.3 (14.7)**	**.031* (G)**	.092	1.0	**.002** **	.176
Composite DQ [Mean (SD)]	**60.3 (18.0)**	*72.6 (21.0)*	**77.8 (25.5)**	**80.4 (27.9)**	**<0.001*** (G)**	.480	**<0.001*** **	**<0.001*** **	.290
Fine Motricity DQ	75.0 (16.9)	*76.4 (16.8)*	76.7 (19.1)	81.9 (25.3)	.058 (G)				
Visual Reception DQ	**75.3 (24.1)**	*86.7 (23.8)*	**89.3 (31.9)**	**91.2 (31.8)**	**<0.001*** **	.235	**<0.001*** **	**<0.001*** **	1.0
Expressive Language DQ	**43.6 (19.6)**	*57.9 (26.3)*	**66.9 (29.6)**	**68.8 (27.5)**	**<0.001*** (G)**	.449	**<0.001*** **	**<0.001*** **	1.0
Receptive Language DQ	**47.0 (25.8)**	*69.3 (28.7)*	**78.1 (31.4)**	**79.6 (33.4)**	**<0.001*** (G)**	0.591	**<0.001*** **	**<0.001*** **	1.0
Demographics									
Chronological age [months Mean (SD)]	28.9 (5.3)								
Gender [females Number (percentage)]	6 (12.8 %)								
Parental Education [Number (percentage)] (n=46)									
Elementary School or High School	18 (39.1%)								
University or PhD	26 (60.9 %)								
Household Income [Number (percentage)] (n=45)									
<60k	12 (26.7%)								
60-140k	17 (37.8%)								
>140k	16 (35.6 %)								

Parsing the heterogeneity in treatment response

Difference between subgroups at baseline

We found no differences between clusters in parental education or household income (See Supplementary Table 1). There were also no differences regarding age at baseline. When looking at DQ at baseline, we found that HB showed higher scores in all DQ subdomains compared to both other clusters. Considering adaptive behavior, HB exhibited higher scores in ABC as well as in the communication subdomain compared to both other clusters. HB also showed a higher score in socialization compared to MR. All statistically significant results of analyses on the DQ and the VABS-II across the three subgroups at baseline are illustrated on Fig 2. There was no difference in the total ADOS CSS. Looking at ADOS subdomains, we found that HB exhibited lower RRB compared to MR. Besides, there was no difference in any clinical measure between MR and HR.

Figure 2: A. Statistically significant differences in DQ subdomains between subgroups at baseline. **B.** Statistically significant differences in VABS ABC and VABS subdomains between subgroups at baseline. *p<.05. **p<.01. ***p<.001. ABC: Adaptative Behavior Composite. DQ: Developmental Quotient. EL: Expressive Language. FM: Fine Motor. RL: Receptive Language. VABS: Vineland Adaptative Behavior Scale. VR: Visual Reception.

Differences between subgroups in rates of change

We found that over the two years of therapy, HR exhibited higher rates of change compared to both other subgroups in cognition (composite DQ as well as VR, FM and RL subdomains) as well as in adaptive behavior compared to MR (VABS-II ABC, in socialization, communication and DLS subdomains) (See Supplementary table 2). Except for socialization, we found that all these differences were already present within the 12 first months of therapy. Also, we found that MR exhibited slower rates of change during the total time of intervention compared to both other subgroups in cognition (composite DQ, VR, FM and EL) as well as in adaptive behavior (VABS-II ABC, socialization, communication and DLS). These differences were already present during the first 12 months of therapy, except for socialization, VR and EL.

Finally, we highlighted that HR already exhibited faster rates of change in composite DQ and adaptative communication compared to MR (Fig 3) after 6 months of intervention. During the same early period, HB exhibited higher SPC in receptive language (RL) DQ compared to HR.

We did not find any difference in the rates of change of symptom severity (total ADOS, SA and RRB) between the three subgroups during the time of intervention.

Figure 3: A. Statistically significant differences between subgroups in the rates of change of behavioral measures within the first 6 months of intervention. **B.** Statistically significant differences between subgroups in the rates of change of behavioral measures within the first 12 months of intervention. *p<.05. **p<.01. ***p<.001. ABC: Adaptative Behavior Composite. DLS: Daily Living Skills. DQ: Developmental Quotient. EL: Expressive Language.

FM: Fine Motor. RL: Receptive Language. SPC: Symmetrized Percent Change. VABS: Vineland Adaptative Behavior Scale. VR: Visual Reception.

Logistic regressions

Minimal responders (MR) and high responders (HR) showed no difference on the composite DQ at baseline. They were thus selected for our classification analyses to address the potential of clinical measures at baseline as well as their early rates of progress to classify them. Since we found no difference in any of the clinical measures at baseline, no logistic regression was performed on this timepoint. Within the first 6 months of therapy, MR showed slower SPC in the VABS-II communication subdomain and in the composite DQ. Logistic regression based on these two variables allowed a partition of MR and HR with a 95.5 % overall correct classification rate (Fig 4 A). All the 17 HR included in the model and 4 out of the 5 MR included were classified correctly. The logistic regression model was statistically significant ($\chi^2 = 10.501, p = .005$) and explained 38% of the variance (Nagelke R^2). The prediction equation was the following: *0 = -.046 * DQ SPC - .065 VABS-II communication SPC + .561*. Within the first 12 months of therapy, HR exhibited higher SPC in both the VABS-II ABC and the composite DQ. Logistic regression performed with both measures reached a 96.4 % rate of overall correct classification between HR and MR (fig 4B). All the 19 HR included in the model and 8 out of the 9 MR included were classified correctly. The logistic regression model was statistically significant ($\chi^2 = 27.274, p < .001$) and explained 87% of the variance (Nagelke R^2). The prediction equation was the following: *0 = -.198 * DQ SPC - .222 VABS-II ABC SPC + 1.838*. In other words, it would have already been possible for a clinician to classify HR and MR with more than 95% of accuracy after six months of therapy based on the participant's rates of change in VABS-II communication and composite DQ, and after 12 months of intervention based on their rates of change in VABS ABC and composite DQ.

Figure 4: A. Logistic regression between MR and HR using their rates of change within the 6 first months of intervention in Composite DQ and VABS communication. Color code corresponds to subgroup membership (red: MR, green: HR). Decision boundary is represented in dotted line. **B.** Logistic regression between MR and HR using their rates of change within the 12 first months of intervention in Composite DQ and VABS ABC. ABC: Adaptative Behavior Composite. DQ: Developmental Quotient. SPC: Symmetrized Percent Change. VABS: Vineland Adaptative Behavior Scale.

Discussion

By analyzing data from one of the largest samples of children who underwent two years of ESDM intervention, the current study, explores the predictors of treatment response. We observed that, on average, the preschoolers in our sample made cognitive progress and showed gains in adaptive skills after two years of intervention (Fig 1A). These results appear consistent with those reported in other studies on ESDM-based intervention (43,50). More specifically, our sample showed an average change in DQ (+20 points) that is very close to the one described in the randomized controlled trials (RCT) study by Dawson et al. (2), which reported around 17 points of cognitive gain, an average significantly greater than that of the

control group. These improvements were allowed most children to continue onto full inclusion into regular education classrooms, which in Geneva requires a child to have near peer-level functioning at the end of the intervention (72.3%). These results reinforce the idea that it is feasible to export this intervention in the European therapeutic landscape with similar level of efficiency as in the USA. Our results also support the cost-effectiveness of ESDM intervention, given that improvements in cognition and adaptive behavior, as well as the reduction of support needed in subsequent schooling offset the costs associated with early intensive intervention (13,14,51). However, it is important to keep in mind that the present study is not an RCT, which limits conclusions about better efficacy of ESDM based intervention compared to other interventions available in Europe. Nonetheless, given the similarity between our results and the ones reported in the ESDM intervention group of Dawson et al.(43), one can infer that ESDM had a causal effect on the improvements we measured.

In the present study, our first aim was to determine whether preschoolers with ASD who participated in a two-year ESDM intervention program could be separated into distinct subgroups based on their cognitive trajectories over time. We used a *k*-means cluster analysis (CA) approach with cognitive abilities at baseline and cognitive rates of change over time as variables. CA yielded three subgroups: ~38% of children with a mild cognitive delay at baseline that displayed a globally good outcome (High at baseline: HB), and two subgroups of low-functioning children (LF) at baseline with very different outcome trajectories. The first one, which represented ~43% of the sample, underwent great cognitive and adaptive skill improvements (High responders: HR) and the second one, which represented ~19% of our sample, saw a widening of the gap increased over time compared to typical development in cognition and adaptative behavior (Minimal responders: MR). The clear distinction between toddlers with mild cognitive delay (HB) on one side and others with the more severe cognitive delay at baseline (LF) on the other is very close to results reported by previous studies that applied CA on preschoolers with ASD (52)(53)(54). These studies identified at least two subgroups categorized as high- and low-functioning based on early cross-sectional cognitive measures. One of the main differences here is that we included a longitudinal variable in our CA (i.e., the rate of cognitive change) and were thus able to define a more fine-grained subgrouping within LF based on the individual cognitive trajectories over time that a cross-sectional CA would have failed to capture. As a second analysis, we aimed to uncover potential predictors of treatment response and evaluate how we can predict a participant's cluster membership. Amongst LF subgroups (MR and HR), we found no behavioral or demographic differences at baseline (see Supplementary Table 1). Nevertheless, there was a great difference between them in their rates of change in cognition and communication skills within the first 6 months of intervention. Using a logistic regression model, we showed that these early behavioral rates of change predicted at 95% attrition to either the MR or HR group.

Our analyses of the HB subgroup suggest that a small cognitive delay (78.8 $\pm$ 10.9 of composite DQ) at the start of an ESDM intervention is associated with an alleviation of the delay in cognitive skills (+10% DQ per year) and adaptive behavior (+4.1% ABC per year) over time. Also, we found that only one child from the HB group did not present a level of functioning for acceptance into a regular education classroom in Geneva. With a DQ of 64 at both the beginning and the end of the intervention, this child was the only one in the HB

group with a DQ value lower than 80 at the end of the intervention. These observations align with a recent review that concluded that higher cognitive abilities at baseline constitute a good predictor of positive outcome after an early intensive intervention (55). Also, the HB group exhibited higher levels of adaptive skills compared to other subgroups (84.9 ± 9.2 of ABC) at baseline, especially in the domain of communication (85.7 ± 11.0). This is in line with previous studies that reported that higher abilities in adaptive behavior (22)(56) as well as in language (57) constitute predictors of good outcome after an early intensive intervention. Overall, our results concerning the HB subgroup suggest a positive effect of ESDM intervention on preschoolers with a mild developmental delay at baseline, which is in line with the conclusion of a previous study that explored this subpopulation (58). This claim will need to be further confirmed with future RCT that assess the specific causality of ESDM intervention in these results. A practical implication of our findings concerning HB is that clinicians who refer a toddler with a mild developmental delay at baseline to an ESDM program can be confident that there will be a good outcome in cognition and adaptative behavior by the end of the intervention.

Apart from the HB subgroup, the rest of the sample was constituted of children with severe cognitive delay at baseline but made up two subgroups with drastically different cognitive trajectories of change over time. On the one hand, despite their severe cognitive delay at baseline (average DQ of 50.0 ± 10.8), the 15 children that composed the HR subgroup greatly improved their cognitive and adaptive skills over time and 75% of them were able to join regular schooling with some in-class support. On the other hand, the 9 children in the MR subgroup had a similar level of cognitive impairments at baseline (average DQ of 46.0 ± 8.3), but the gap in their cognitive and adaptive functioning scores increased over time compared to typical development, and this despite receiving early and intensive intervention. Only 2 out of 9 (22.2%) MR children met criteria to join a regular school following intervention. Here, we failed to identify any behavioral differences (i.e. in cognition, adaptative behavior or autism severity) that could accurately discriminate between HR and MR at baseline. One clinical implication of our analyses of HR and MR at baseline is that there is currently no clear clinical marker within LF toddlers that could help clinicians predict how the child will respond to ESDM intervention. Nonetheless, it is important to note that HR constituted the majority of the LF at baseline (~69% of the LF subgroup) thus supporting the *a priori* that most of LF toddlers respond positively to the ESDM. Nonetheless, a better understanding of the factors (behavioral, biological, and environmental) that are associated with MR is necessary to develop more targeted clinical recommendations. Moreover, the observation of two distinct trajectories of change in children with larger cognitive impairments at baseline could shed a new light on the inconsistencies that exist between various studies that measured the cognitive response to early intervention within LF preschoolers with ASD. For instance, one previous study concluded that children with this profile only improve in fine motor skills and receptive language but not in adaptive behavior (58). Other studies reported an association between low cognitive functioning at baseline and high cognitive gains (59) (60). One can hypothesize that the inter-individual heterogeneity of outcome reported by Ben-Itzchak et al., Devescovi et al. and Robain et al., as well as the differences in their results were due to the existence of two dynamic latent subgroups (MR and HR) that may have driven results in opposite directions. Our results thus advocate for a more systematic subgroup phenotyping, including longitudinal variables, in future studies focusing on the clinical outcome of early intensive

intervention to better describe the phenotypic heterogeneity within LF preschoolers with ASD.

Finally, our results suggest that despite the lack of significant differences at baseline within LF children, the outcome after two years of intervention can be predicted by the end of the first year of intervention with high accuracy levels. Indeed, the HR group's rates of change appear to be significantly higher than those observed in the MR group within the first year regarding their cognitive and adaptive skills (especially in communication and daily living skills). Thus, our results are consistent with those of Sallows et al. (61), who reported cognitive gain during the first year of intervention as one of the best predictors of outcome at the end of the intervention. Together, these conclusions provide answers to the question of the timing of "response to treatment" and when children can be considered "*non responders*" as raised by Vivanti et al. (10). Our analyses suggest that the first year of intervention is a crucial period for future therapy outcomes and that children who do not show progress by this time could be considered minimal responders. Moreover, based on the rates of change within the first six months of intervention in cognition and communication, we could infer the outcome after two years of intervention. The emphasis on this ultra-early response to intervention as a predictor of long-term outcome has several clinical implications. One of them is the interest of implementing a close standardized follow-up of children's cognition and adaptative behavior in the first 6 months in addition to the systematic ESDM *Curriculum Checklist* that is included in the model. This early standardized follow-up could help the clinician anticipating the ultimate outcome of intervention. Indeed, in view of our results, it is probable that an early widening of the developmental gap compared to typical development is followed by a very similar developmental pattern in cognitive and adaptative abilities during the rest of the intervention. Therefore, the implementation of a systematic standardized early follow-up could lead to an earlier adaptation of the therapeutic objectives and strategies with these children. Early identification of minimal responders should therefore be followed by a reassessment of the child's needs, specific strengths, and weaknesses to provide the best individualized therapeutic tools according to the child's developmental profile. Nonetheless, the question of which specific intervention suits which clinical profile the best still needs to be addressed as a recent multisite RCT failed to find any advantage between ESDM and applied behavior analysis (ABA) based interventions on preschoolers with an initial severe cognitive delay (15). Overall, our results show that despite the lack of reliable outcome predictors within children presenting severe cognitive delays at baseline, the consideration of their early dynamic behavioral parameters can efficiently help predicting the ultimate response to the intervention. Nonetheless, further RCTs that explore the trajectories of subgroups similar to ours are needed to determine the precise effect of ESDM on the differences observed between MR and HR. More specifically, it should be assessed whether ESDM helps HR improving their outcome or prevents MR from widening their developmental gap compared to typical development, or both at the same time. Another hypothesis to be addressed is whether ESDM has an influence in the relative number of participants that are affected to each subgroup – i.e., whether some HR participants would have been MR if they did not benefit from an ESDM based therapy. A better understanding of the specific effect of ESDM in each subgroup could result in future therapeutic guidelines that are more tailored for each child's individual developmental trajectory.

Limitations

Despite being one of the largest samples of preschoolers who benefited from a two-years intensive ESDM intervention, the number of participants in the present study limits the number as well as the size of subgroups that can be detected by a cluster analysis. Nevertheless, we took care to respect the commonly accepted prerequisites of cluster analyses, including the minimum sample size in each group or the number of factors in the analyses given the overall sample size (62,63). It is possible that studies performed on larger samples could achieve more fine-grained subgrouping on a similar population based on the same behavioral measures. Also, cluster analysis performed on larger samples could lead to bigger subgroups, in turn increasing the statistical power to detect differences at baseline between low-functioning clusters that we could not highlight.

Another limitation that is somehow a direct consequence of the previous one lies in the choice of the main outcome. Here, we chose parameters related to cognitive skills as the main clustering factors. However, it would have been possible to use other measures such as the level of ASD symptoms, the adaptive skills or even a combination of all these. The inclusion of more factors in the model could help in defining a larger number of clusters and therefore help in understanding the heterogeneity of ASD in a refined manner. However, this was not possible in the present study, once again because of the limited sample size. The addition of more factors in the model and the multiplication of clusters would have violated the cluster analyses assumptions, making its interpretation invalid. Studies with larger samples should include more clinical parameters in the model, and could also use parents suggested outcome parameters (64).

Conclusion

In this study, we applied a cluster analysis approach on the largest European sample of preschoolers with ASD who participated to an intensive ESDM program for 20 hours a week over a duration of two years. Overall, participants decreased their symptom severity and partially alleviated their cognitive delay over the intervention period. Furthermore, the cluster analysis suggested three main patterns of cognitive trajectories over time. First, children who only displayed mild cognitive and adaptative behavioral delays before the age of three exhibited a good prognosis in terms of general development as they partially compensated their delay in cognition and adaptative behavior over time. Second, children who presented with severe cognitive delays before the start of the therapeutic program exhibited two dramatically opposite patterns of developmental trajectories. About a third of the low-functioning children at baseline widened the gap in cognitive and adaptive functioning scores compared to typical development. The two remaining thirds of the children in the low-functioning group exhibited early and spectacular gains in cognition and adaptative behavior that were maintained throughout the duration of the intervention. We did not evidence any significant behavioral or demographic differences between these two low-functioning groups at baseline, suggesting that it is not yet possible to predict the gains that will be made by low-functioning preschoolers with ASD solely based on the clinical measures that are commonly used in the clinical practice. Nevertheless, our results suggest that we could have classified them with very high levels of accuracy after only 6 months of

therapy based on their early gains in global cognition and communicational adaptative behavior. These results advocate for the implementation of close monitoring using standardized cognitive and adaptative behavioral testing within the first six months of ESDM intervention, especially for children that exhibit an important cognitive delay at baseline. This could help the clinicians predict clinical outcome and adapt the intervention as needed. Still, future studies are warranted to evaluate the kind of adaptation that would benefit each ASD subgroup the most. Also, longitudinal studies that provide a long-term follow-up after the end of the intervention are needed to assess whether the cognitive patterns that we observed are stable over time. Overall, our results advocate for a more systematic use of subgroup phenotyping that includes longitudinal parameters when assessing the efficacy of an early intensive intervention to better decipher the great heterogeneity of behavioral dynamics in treatment response.

Supplementary Material
Supplementary table 1
Supplementary table 2

Declaration

Author's contribution
M.F, M.S., conceived and designed the study. M.F, M.G., N.K., and F.R. participated to the data acquisition. M.G. and F.R. prepared and analyzed the data under the supervision of M.S. All authors participated in interpretation of results. M.G. and F.R. wrote the manuscript with the inputs from all other authors. All authors read and approved the final manuscript.

Funding
This research was supported by the Swiss National Foundation Synapsy Grant No. (51NF40 – 185897) and the Swiss National Foundation for Scientific Research Grant (No. 323630-191227 to M.G., and #163859 & #190084 to M.S.). and by the "Fondation Pôle Autisme" (https://www.pole-autisme.ch). The funders were not involved in this study and had no role other than to provide financial support.

Availability of data and materials

The datasets generated in the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

Informed consent was obtained from the parents of all participants included in the study. Swissethics - Commission d'éthique Suisse relative à la recherche sur l'être humain approved this study (Protocole 12-163/Psy 12-014), referred under the number PB_2016-01880.

Consent for publication

Consent for publication was obtained from the parents of all participants. Consent for publication of the picture of the children presented in the task was obtained from their parents.

Competing interests

The authors have no competing interests to report.

Acknowledgments

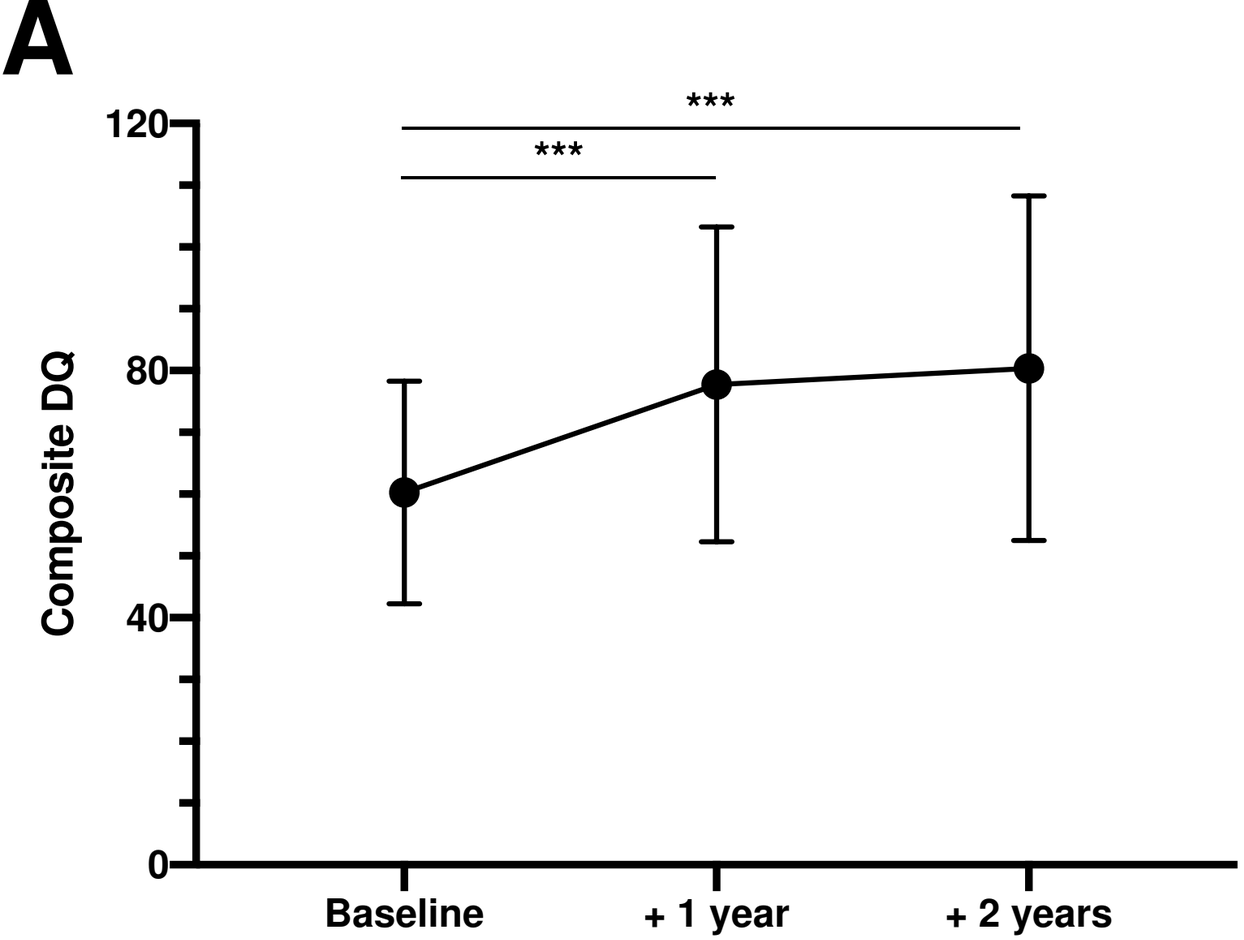
The authors would like to thank all the families who kindly participated in the study, all therapists at the Centre d'Intervention Précoce en Autisme in Geneva, as well as Alexandra Bastos, Stéphanie Baudoux, Lylia Ben Hadid, Aurélie Bochet, Léa Chambaz, Flore Couty, Sophie Diakonoff, Lisa Esposito, Constance Ferrat, Marie-Agnès Graf, Oriane Grosvernier, Kenza Latrèche, Sara Maglio, Matthieu Mansion, Eva Micol, Irene Pittet, Sonia Richetin, Laura Sallin, Stefania Solazzo, Myriam Speller, Chiara Usuelli and Ornella Vico Begara for their help with data collection.

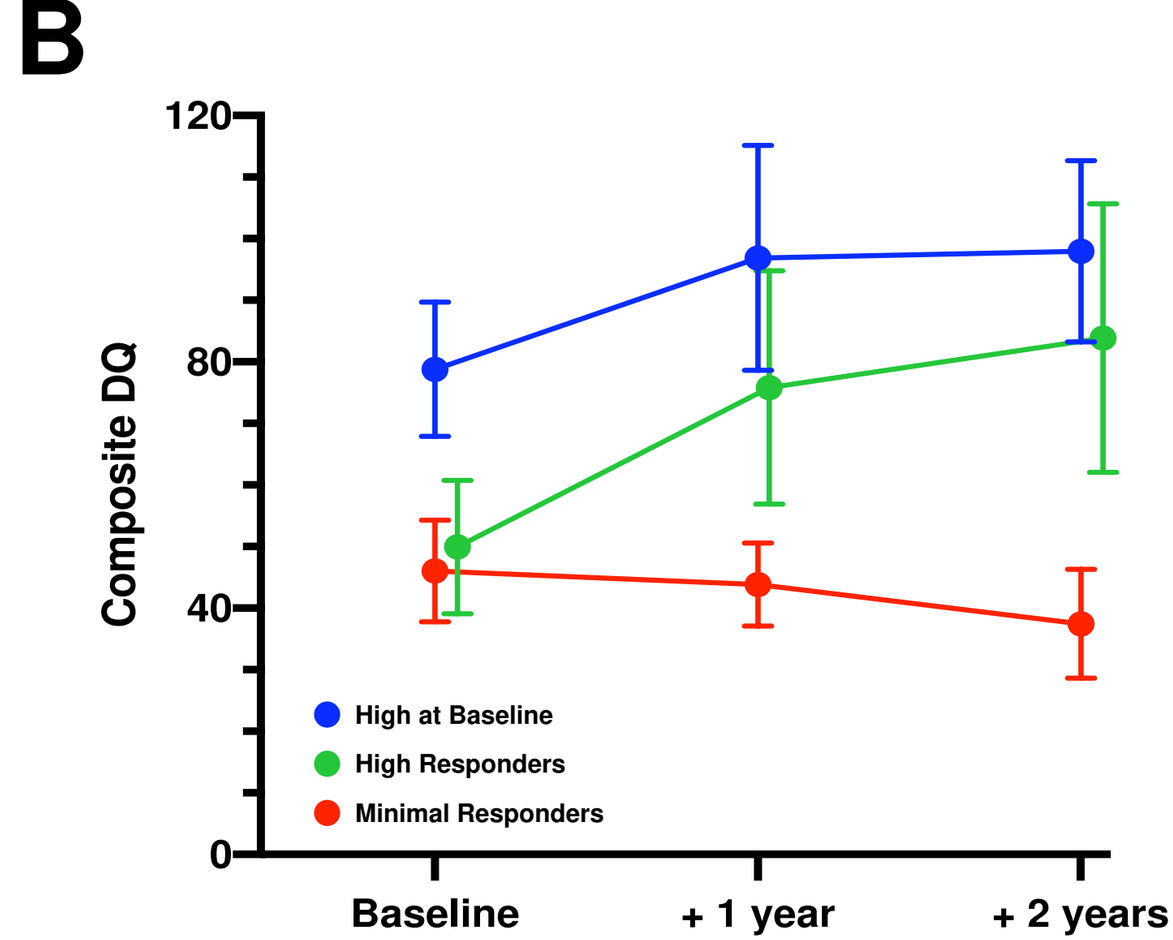
1. American Psychiatric Association, editor. Diagnostic and statistical manual of mental disorders: DSM-5. 5th ed. Washington, D.C: American Psychiatric Association; 2013. 947 p.
2. Dawson G, Rogers SJ, Munson J, Smith M, Winter J, Greenson J, et al. Randomized, Controlled Trial of an Intervention for Toddlers With Autism: The Early Start Denver Model. Pediatrics. 2010 Jan 1;125(1):e17–23.
3. Lovaas OI, Schreibman L, Koegel RL. A behavior modification approach to the treatment of autistic children. J Autism Dev Disord. 1974 Jun 1;4(2):111–29.
4. Freitag CM, Feineis-Matthews S, Valerian J, Teufel K, Wilker C. The Frankfurt early intervention program FFIP for preschool aged children with autism spectrum disorder: a pilot study. J Neural Transm. 2012 Sep 1;119(9):1011–21.
5. Koegel LK, Koegel RL, Harrower JK, Carter CM. Pivotal Response Intervention I: Overview of Approach. Journal of the Association for Persons with Severe Handicaps. 1999 Sep 1;24(3):174–85.
6. Prizant BM, Wetherby AM, Rubin E, Laurent AC, Rydell P. The Scerts Model: A Comprehensive Educational Approach for Children With Autism Spectrum Disorders. Baltimore, Md: Brookes Publishing Co; 2006. 774 p.
7. Schreibman L, Dawson G, Stahmer AC, Landa R, Rogers SJ, McGee GG, et al. Naturalistic Developmental Behavioral Interventions: Empirically Validated Treatments for Autism Spectrum Disorder. J Autism Dev Disord. 2015;45(8):2411–28.
8. Howlin P, Magiati I, Charman T. Systematic Review of Early Intensive Behavioral Interventions for Children With Autism. MacLean Jr William E, editor. American Journal on Intellectual and Developmental Disabilities. 2009 Jan 1;114(1):23–41.
9. Warren Z, McPheeters ML, Sathe N, Foss-Feig JH, Glasser A, Veenstra-VanderWeele J. A Systematic Review of Early Intensive Intervention for Autism Spectrum Disorders. Pediatrics. 2011 May 1;127(5):e1303–11.
10. Vivanti G, Prior M, Williams K, Dissanayake C. Predictors of Outcomes in Autism Early Intervention: Why Don't We Know More? Front Pediatr [Internet]. 2014 [cited 2017 Nov 27];2. Available from: https://www.frontiersin.org/articles/10.3389/fped.2014.00058/full
11. Zwaigenbaum L, Bauman ML, Choueiri R, Kasari C, Carter A, Granpeesheh D, et al. Early Intervention for Children With Autism Spectrum Disorder Under 3 Years of Age: Recommendations for Practice and Research. Pediatrics. 2015 Oct 1;136(Supplement 1):S60–81.
12. Fuller EA, Oliver K, Vejnoska SF, Rogers SJ. The Effects of the Early Start Denver Model for Children with Autism Spectrum Disorder: A Meta-Analysis. Brain Sci [Internet]. 2020 Jun 12 [cited 2021 Mar 24];10(6). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7349854/
13. Cidav Z, Munson J, Estes A, Dawson G, Rogers S, Mandell D. Cost Offset Associated With Early Start Denver Model for Children With Autism. Journal of the American Academy of Child & Adolescent Psychiatry. 2017 Sep;56(9):777–83.
14. Penner M, Rayar M, Bashir N, Roberts SW, Hancock-Howard RL, Coyte PC. Cost-Effectiveness Analysis Comparing Pre-diagnosis Autism Spectrum Disorder (ASD)-Targeted Intervention with Ontario's Autism Intervention Program. J Autism Dev Disord. 2015 Sep;45(9):2833–47.
15. Rogers SJ, Yoder P, Estes A, Warren Z, McEachin J, Munson J, et al. A Multisite Randomized Controlled Trial Comparing the Effects of Intervention Intensity and Intervention Style on Outcomes for Young Children With Autism. Journal of the American Academy of Child & Adolescent Psychiatry. 2020 Aug;S0890856720313502.

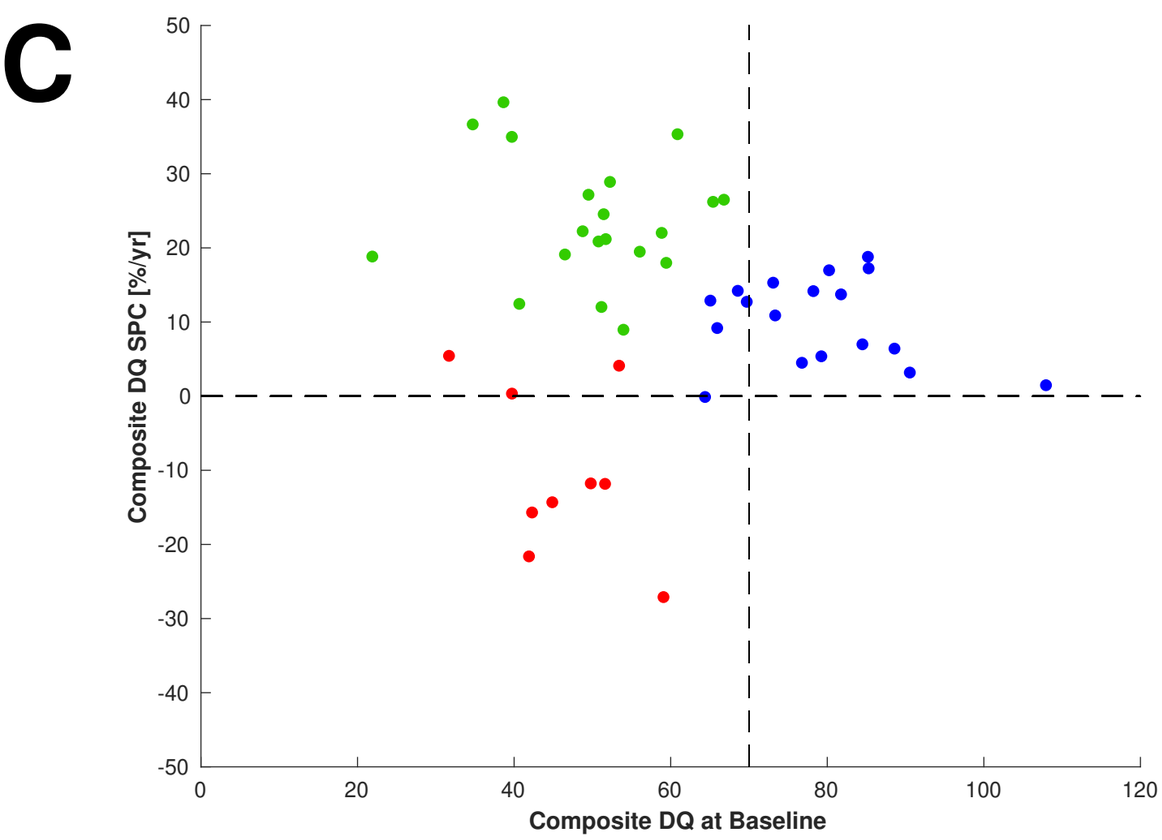
16. Contaldo A, Colombi C, Pierotti C, Masoni P, Muratori F. Outcomes and moderators of Early Start Denver Model intervention in young children with autism spectrum disorder delivered in a mixed individual and group setting. Autism. 2020 Apr;24(3):718–29.
17. Vivanti G, Dissanayake C, Zierhut C, Rogers SJ, Team VA. Brief Report: Predictors of Outcomes in the Early Start Denver Model Delivered in a Group Setting. J Autism Dev Disord. 2013 Jul 1;43(7):1717–24.
18. Estes A, Swain DM, MacDuffie KE. The effects of early autism intervention on parents and family adaptive functioning. Pediatr Med. 2019 Jun;2:21–21.
19. König IR, Fuchs O, Hansen G, Mutius E von, Kopp MV. What is precision medicine? European Respiratory Journal [Internet]. 2017 Oct 1 [cited 2021 Mar 25];50(4). Available from: https://erj.ersjournals.com/content/50/4/1700391
20. Klintwall L, Eikeseth S. Number and controllability of reinforcers as predictors of individual outcome for children with autism receiving early and intensive behavioral intervention: A preliminary study. Research in Autism Spectrum Disorders. 2012 Jan;6(1):493–9.
21. Victorian ASELCC Team, Vivanti G, Dissanayake C, Zierhut C, Rogers SJ. Brief Report: Predictors of Outcomes in the Early Start Denver Model Delivered in a Group Setting. J Autism Dev Disord. 2013 Jul;43(7):1717–24.
22. Sallows GO, Graupner TD. Intensive Behavioral Treatment for Children With Autism: Four-Year Outcome and Predictors. Am J Mental Retard. 2005;110(6):417.
23. Sallows GO, Graupner TD. Replicating Lovaas' treatment and findings: Preliminary results. InPEACH: Putting research into Practice Conference. 1999;
24. Perry A, Cummings A, Dunn Geier J, Freeman NL, Hughes S, LaRose L, et al. Effectiveness of Intensive Behavioral Intervention in a large, community-based program. Research in Autism Spectrum Disorders. 2008 Oct;2(4):621–42.
25. Ousley O, Cermak T. Autism Spectrum Disorder: Defining Dimensions and Subgroups. Curr Dev Disord Rep. 2014 Mar;1(1):20–8.
26. Syriopoulou- Delli CK, Papaefstathiou E. Review of cluster analysis of phenotypic data in Autism Spectrum Disorders: distinct subtypes or a severity gradient model? International Journal of Developmental Disabilities. 2020 Jan 1;66(1):13–21.
27. Frazier TW, Klingemier EW, Anderson CJ, Gengoux GW, Youngstrom EA, Hardan AY. A Longitudinal Study of Language Trajectories and Treatment Outcomes of Early Intensive Behavioral Intervention for Autism. J Autism Dev Disord. 2021 Feb 8;1–17.
28. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders [Internet]. Fifth Edition. American Psychiatric Association; 2013 [cited 2018 Nov 4]. Available from: https://psychiatryonline.org/doi/book/10.1176/appi.books.9780890425596
29. Rutter M, DiLavore PC, Risi S, Gotham K, ishop S. Autism diagnostic observation schedule: ADOS-2. Torrance, CA: Western Psychological Services.; 2012.
30. Lord C, Rutter M, DiLavore PC, Risi S, Gotham K, Bishop SL. Autism Diagnostic Observation Schedule, 2nd edition (ADOS-2). Los Angeles, CA: Western Psychological Services; 2012.
31. Franchini M, Zöller D, Gentaz E, Glaser B, Wood de Wilde H, Kojovic N, et al. Early Adaptive Functioning Trajectories in Preschoolers With Autism Spectrum Disorders. Journal of Pediatric Psychology. 2018 Aug 1;43(7):800–13.
32. Rogers SJ, Dawson G. Early Start Denver Model for young children with autism: promoting language, learning, and engagement. New York: Guilford Press; 2010. 297 p.

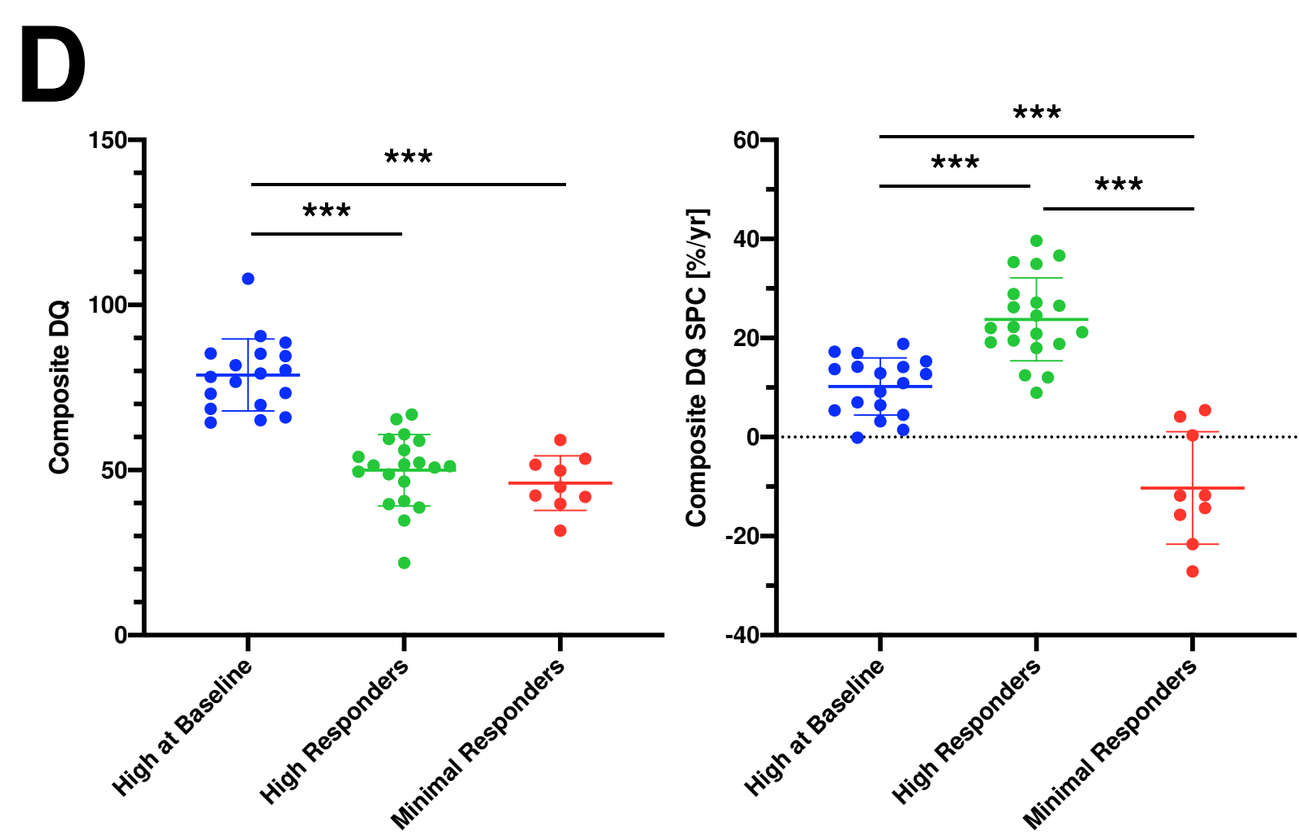
33. Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, DiLavore PC, et al. The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. J Autism Dev Disord. 2000 Jun;30(3):205–23.
34. Hus V, Gotham K, Lord C. Standardizing ADOS Domain Scores: Separating Severity of Social Affect and Restricted and Repetitive Behaviors. J Autism Dev Disord. 2014 Oct;44(10):2400–12.
35. Gotham K, Pickles A, Lord C. Standardizing ADOS scores for a measure of severity in autism spectrum disorders. J Autism Dev Disord. 2009 May;39(5):693–705.
36. Mullen EM. Mullen scales of early learning. Circle Pines, MN: American Guidance Service; 1995.
37. E. Schopler, Lansing M, Reichler R, Marcus L. Psychoeducational Profile: TEACCH individualized assessment for chil- dren with autism spectrum disorders (3rd ed.). Pro-Ed Inc. Austin; 2005.
38. Lord C, Risi S, DiLavore PS, Shulman C, Thurm A, Pickles A. Autism From 2 to 9 Years of Age. Arch Gen Psychiatry. 2006 Jun 1;63(6):694.
39. Sparrow S, Balla D, Cicchetti D. Vineland adaptive behavior scales (2nd ed.). Circle Pines, MN: American Guidance Service. 2005.
40. Hollingshead AB. Four factor index of social status. 1975;
41. Berry DA, Ayers GD. Symmetrized Percent Change for Treatment Comparisons. The American Statistician. 2006 Feb;60(1):27–31.
42. Smith T, Groen AD, Wynn JW. Randomized trial of intensive early intervention for children with pervasive developmental disorder. Am J Ment Retard. 2000 Jul;105(4):269–85.
43. Dawson G, Rogers S, Munson J, Smith M, Winter J, Greenson J, et al. Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model. Pediatrics. 2010 Jan;125(1):e17-23.
44. Landa RJ. Efficacy of early interventions for infants and young children with, and at risk for, autism spectrum disorders. International Review of Psychiatry. 2018 Jan 2;30(1):25–39.
45. Rodgers M, Simmonds M, Marshall D, Hodgson R, Stewart LA, Rai D, et al. Intensive behavioural interventions based on applied behaviour analysis for young children with autism: An international collaborative individual participant data meta-analysis. Autism. 2021 Jan 22;136236132098568.
46. Witwer AN, Lecavalier L. Examining the validity of autism spectrum disorder subtypes. J Autism Dev Disord. 2008 Oct;38(9):1611–24.
47. Hartigan JA, Wong MA. Algorithm AS 136: A K-Means Clustering Algorithm. Applied Statistics. 1979;28(1):100.
48. Trupti M. Kodinariya PRM. Review on determining number of Cluster in K-Means Clustering. International Journal of Advance Research in Computer Science and Management Studies. 2013;1(6):90.
49. Chiu T, Fang D, Chen J, Wang Y, Jeris C. A robust and scalable clustering algorithm for mixed type attributes in large database environment. In: Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '01 [Internet]. San Francisco, California: ACM Press; 2001 [cited 2021 Feb 10]. p. 263–8. Available from: http://portal.acm.org/citation.cfm?doid=502512.502549
50. Rogers SJ, Estes A, Lord C, Vismara L, Winter J, Fitzpatrick A, et al. Effects of a Brief Early Start Denver Model (ESDM)–Based Parent Intervention on Toddlers at Risk for Autism

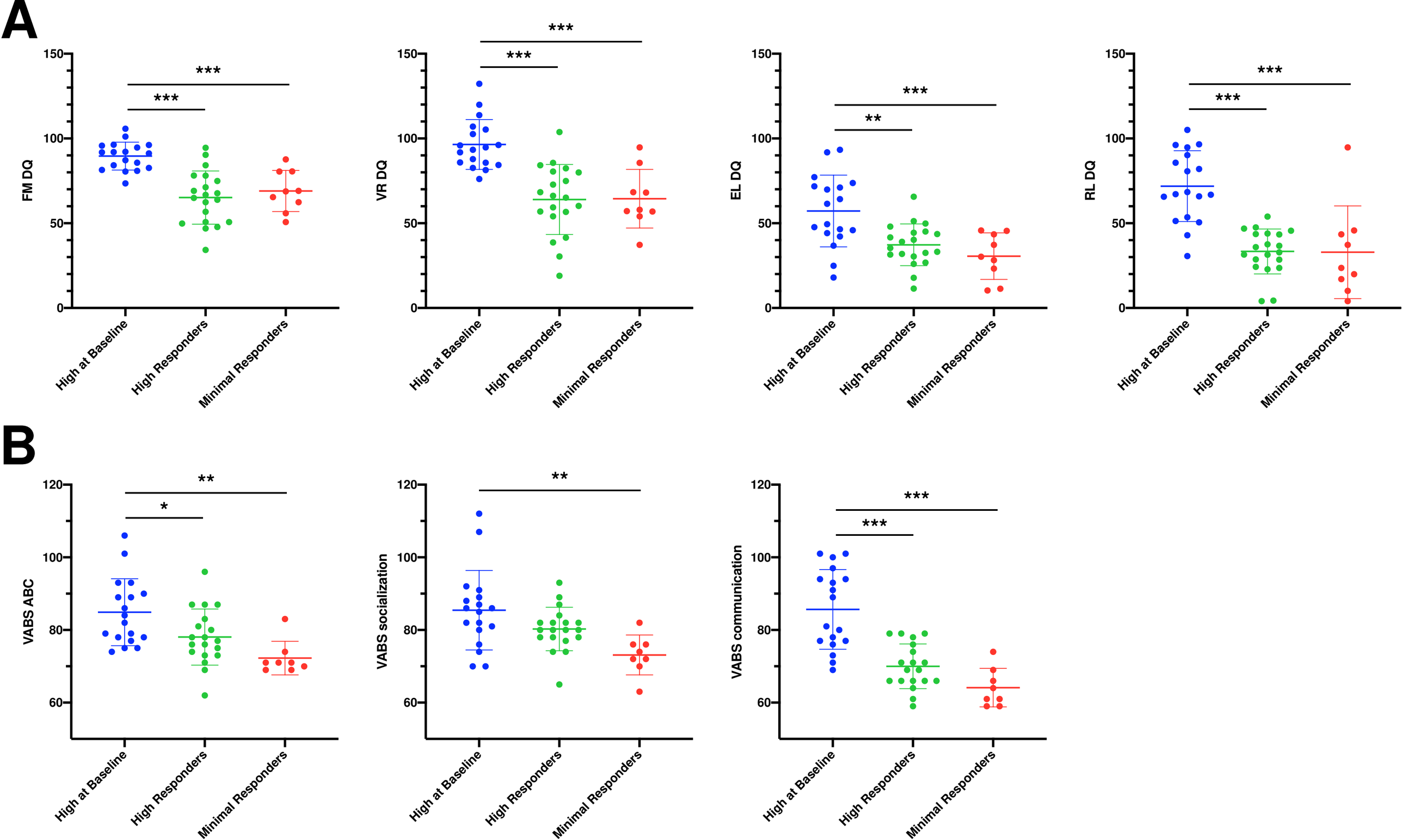
Spectrum Disorders: A Randomized Controlled Trial. Journal of the American Academy of Child & Adolescent Psychiatry. 2012 Oct 1;51(10):1052–65.
51. Peters-Scheffer N, Didden R, Korzilius H, Matson J. Cost comparison of early intensive behavioral intervention and treatment as usual for children with autism spectrum disorder in the Netherlands. Research in Developmental Disabilities. 2012 Nov;33(6):1763–72.
52. Fein D, Stevens M, Dunn M, Waterhouse L, Allen D, Rapin I, et al. Subtypes of Pervasive Developmental Disorder: Clinical Characteristics. Child Neuropsychology. 1999 Mar;5(1):1–23.
53. Stevens MC, Fein DA, Dunn M, Allen D, Waterhouse LH, Feinstein C, et al. Subgroups of Children With Autism by Cluster Analysis: A Longitudinal Examination. Journal of the American Academy of Child & Adolescent Psychiatry. 2000 Mar;39(3):346–52.
54. Zheng S, Hume KA, Able H, Bishop SL, Boyd BA. Exploring Developmental and Behavioral Heterogeneity among Preschoolers with ASD: A Cluster Analysis on Principal Components. Autism Research. 2020 May;13(5):796–809.
55. Zachor D, Ben-Itzchak E. Variables Affecting Outcome of Early Intervention in Autism Spectrum Disorder. J Pediatr Neurol. 2017 Jun;15(03):129–33.
56. Eldevik S, Hastings RP, Hughes JC, Jahr E, Eikeseth S, Cross S. Using Participant Data to Extend the Evidence Base for Intensive Behavioral Intervention for Children With Autism. American Journal on Intellectual and Developmental Disabilities. 2010 Sep 1;115(5):381–405.
57. Magiati I, Moss J, Charman T, Howlin P. Patterns of change in children with Autism Spectrum Disorders who received community based comprehensive interventions in their pre-school years: A seven year follow-up study. Research in Autism Spectrum Disorders. 2011 Jul;5(3):1016–27.
58. Ben-Itzchak E, Watson LR, Zachor DA. Cognitive Ability is Associated with Different Outcome Trajectories in Autism Spectrum Disorders. J Autism Dev Disord. 2014 Sep;44(9):2221–9.
59. Devescovi R, Monasta L, Mancini A, Bin M, Vellante V, Carrozzi M, et al. Early diagnosis and Early Start Denver Model intervention in autism spectrum disorders delivered in an Italian Public Health System service. NDT. 2016 Jun;1379.
60. Robain F, Franchini M, Kojovic N, Wood de Wilde H, Schaer M. Predictors of Treatment Outcome in Preschoolers with Autism Spectrum Disorder: An Observational Study in the Greater Geneva Area, Switzerland. J Autism Dev Disord. 2020 Mar 12;
61. Sallows GO, Graupner TD. Intensive Behavioral Treatment for Children With Autism: Four-Year Outcome and Predictors. MacLean Jr. WE, editor. American Journal on Mental Retardation. 2005 Nov 1;110(6):417–38.
62. Formann AK. Die Latent-Class-Analyse: Einführung in die Theorie und Anwendung [Latent class analysis: Introduction to theory and application]. Weinheim. Beltz; 1984.
63. Qiu W, Joe H. Generation of Random Clusters with Specified Degree of Separation. Journal of Classification. 2006 Feb 1;23:315–34.
64. McConachie H, Livingstone N, Morris C, Beresford B, Couteur AL, Gringras P, et al. Parents Suggest Which Indicators of Progress and Outcomes Should be Measured in Young Children with Autism Spectrum Disorder. J Autism Dev Disord. 2018 Apr 1;48(4):1041–51.

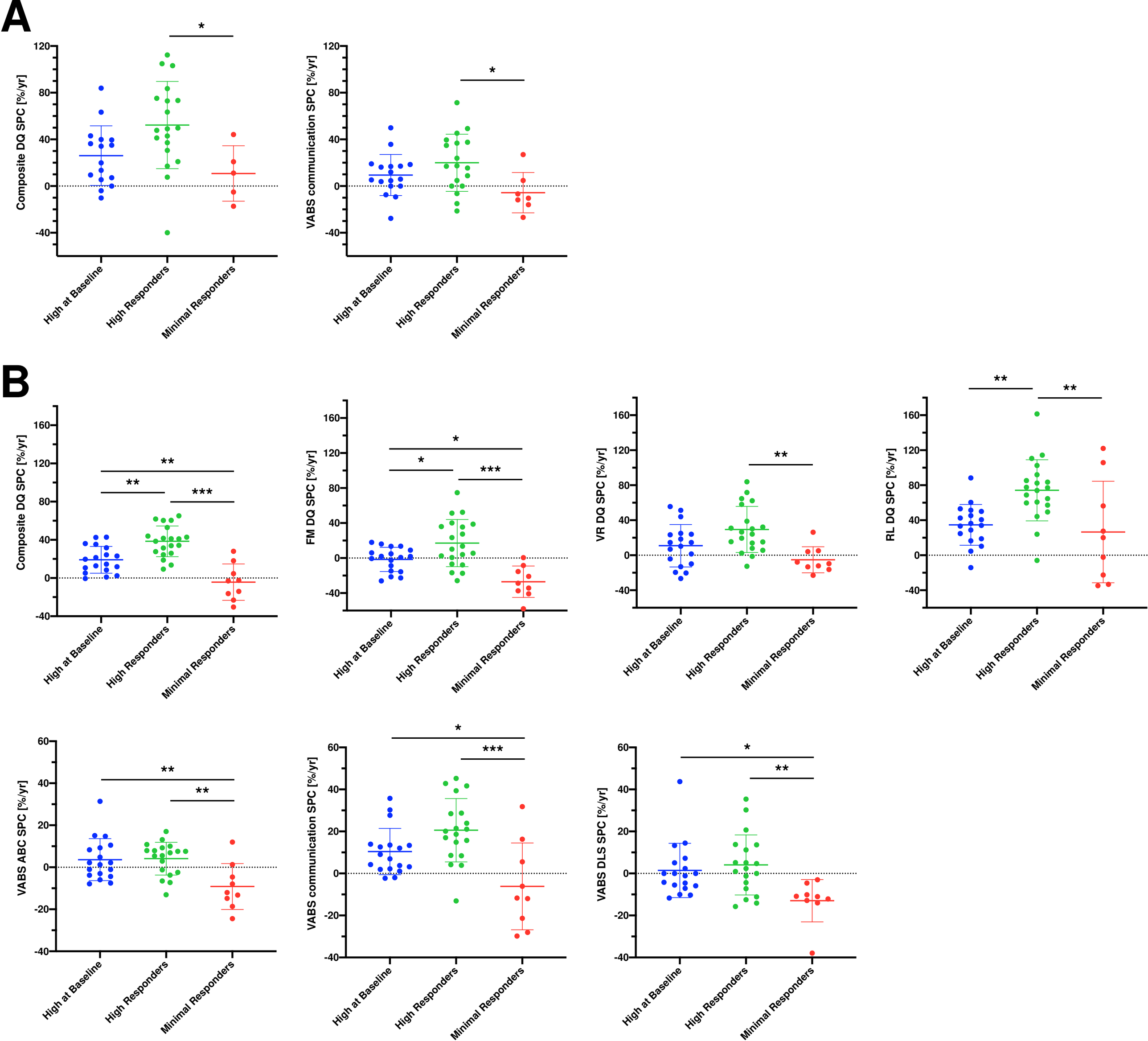

A
***
***
120
80
40
0
Composite DQ
Baseline
+ 1 year
+ 2 years

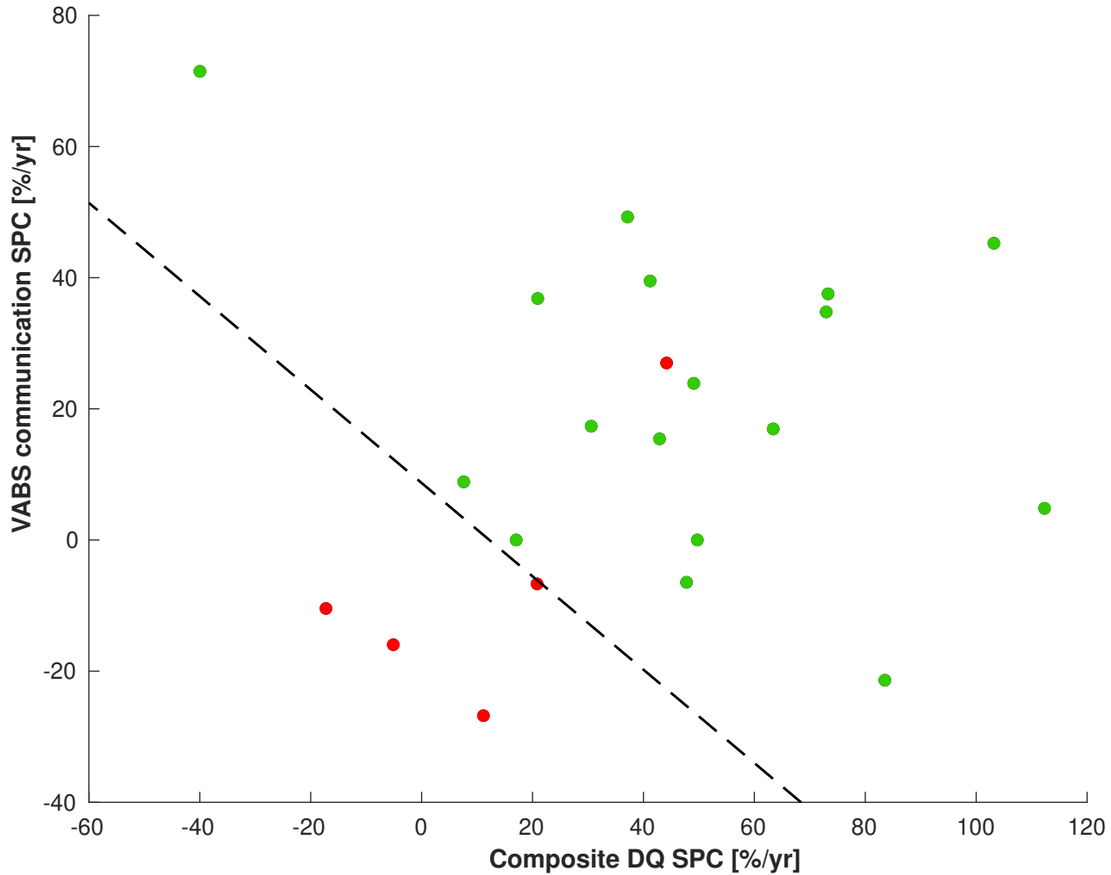

B
120
80
40
0
Composite DQ
High at Baseline
High Responders
Minimal Responders
Baseline
+ 1 year
+ 2 years

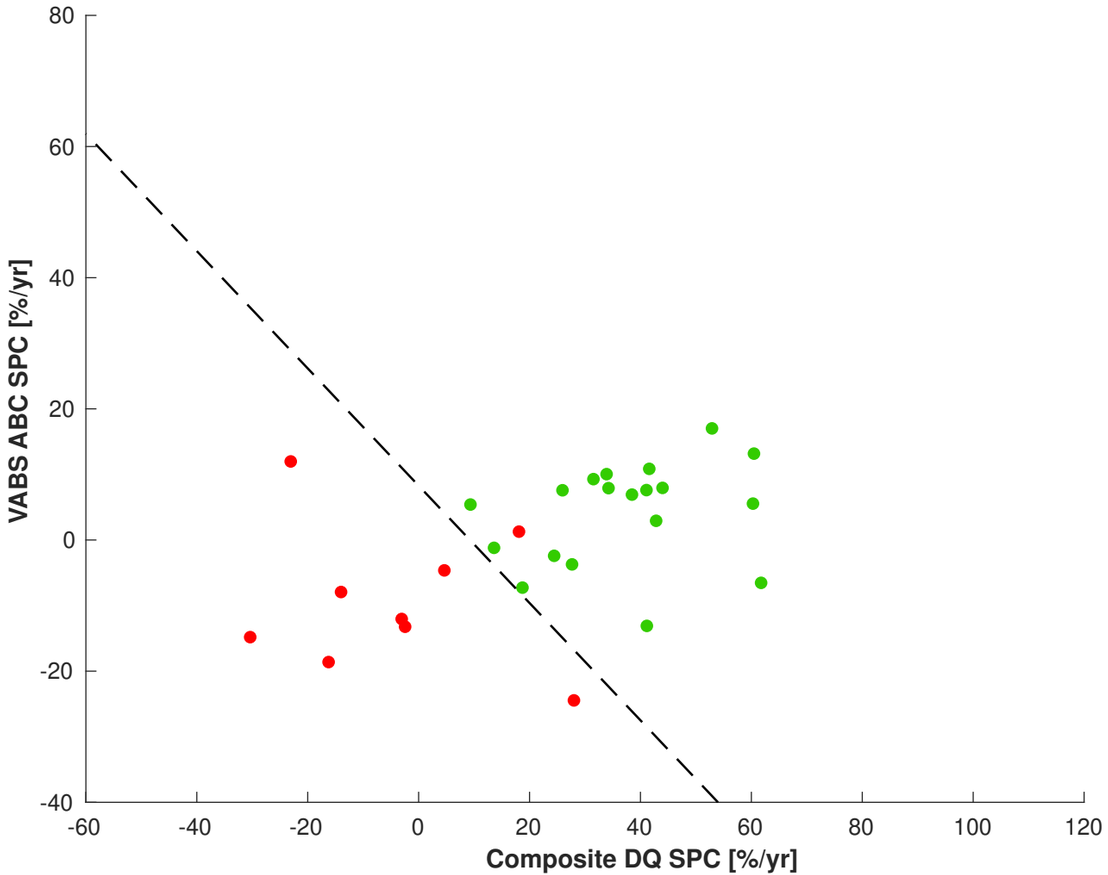

C
50
40
30
20
10
0
-10
-20
-30
-40
-50
Composite DQ SPC [%/yr]
0
20
40
60
80
100
120
Composite DQ at Baseline


D
150
100
50
0
Composite DQ
***
***
High at Baseline
High Responders
Minimal Responders
60
40
20
0
-20
-40
Composite DQ SPC [%/yr]
***
***
***
High at Baseline
High Responders
Minimal Responders


A
FM DQ
VR DQ
EL DQ
RL DQ
150
100
50
0
***
**
High at Baseline
High Responders
Minimal Responders
B
VABS ABC
VABS socialization
VABS communication
120
80
60
*


A
Composite DQ SPC [%/yr]
VABS communication SPC [%/yr]
High at Baseline
High Responders
Minimal Responders
B
Composite DQ SPC [%/yr]
FM DQ SPC [%/yr]
VR DQ SPC [%/yr]
RL DQ SPC [%/yr]
VABS ABC SPC [%/yr]
VABS communication SPC [%/yr]
VABS DLS SPC [%/yr]


VABS communication SPC [%/yr]
Composite DQ SPC [%/yr]
80
60
40
20
0
-20
-40
-60
-40
-20
0
20
40
60
80
100
120


VABS ABC SPC [%/yr]
Composite DQ SPC [%/yr]
80
60
40
20
0
-20
-40
-60
-40
-20
0
20
40
60
80
100
120