medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Distinct patterns of cognitive outcome in young children with autism spectrum disorder receiving the Early Start Denver Model

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 significantly improves developmental outcome in young children with autism. Nonetheless, there is high interindividual heterogeneity in developmental trajectories during the therapy. It is currently not possible to predict a child's developmental outcome after receiving an early intervention based on his skills and symptoms at baseline. In this study, we analyze developmental trajectories of preschoolers with autism who received 2 years of intervention using the Early Start Denver Model (ESDM) approach in Geneva, Switzerland (n = 47). Our aim was to identify early predictors of response to treatment. We applied a cluster analysis to distinguish between 3 subgroups based on their cognitive level at intake, and rates of cognitive change over the course of treatment. The first subgroup of children only had a mild cognitive delay at intake, and nearly no cognitive delay by the end of treatment (Higher Cognitive at baseline: HC). The children in the two other subgroups all presented with severe cognitive delay at baseline, however, they had two very different patterns of response to treatment. The majority significantly improved developmental scores within their first year of intervention (Optimal Responders: OptR). Whereas a minority of children showed little to no improvement over the course of treatment (Minimal Responders: MinR). Our findings suggest that children who end up having an optimal two-year treatment outcome show rapid developmental improvement. Inversely, less significant progress by the sixth month of intervention may predict a less optimal response to treatment.

Keywords: Autism Spectrum Disorders, Early intervention, Predictors, Response to treatment, Heterogeneity, Minimal responder

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Introduction

Autism spectrum disorder (ASD) is characterized by impairments in communication and social interactions, along with restricted and repetitive behaviors (1). Over the last three decades, several comprehensive, evidence-based early intensive behavioral intervention (EIBI) approaches have been developed for young children with ASD, with the aim to improve their social communication, cognitive functioning, and adaptive skills (2–7). As the name suggests, the principles of EIBI comprise a significant number of hours (usually more than 15 hours per week) as well as an early age of onset (usually younger than 4 or 5 years old) (8). Systematic reviews and meta-analyses showed positive effects of EIBI on cognition, adaptive skills and communication at the group level (9,10). Nevertheless, many EIBI studies have reported a relatively heterogeneous response to these interventions, where most children show significant improvements, while others make smaller gains (11,12). Despite important efforts to better understand variables affecting treatment response, it is currently not possible to predict to what extent a child will respond to intervention based on his or her behavioral characteristics at intake (13). In the current therapeutic context and in the absence of additional knowledge about individual predictors of outcome, many authors suggest that intensive early intervention should be an intervention of choice for young children diagnosed with ASD (14,15) regardless of their specific behavioral or symptom profile. Yet, in the global framework of precision medicine (16), there is an urge to develop more individualized guidelines for intervention in ASD. Given the importance of providing effective programming for children with ASD as early as possible, and because of the costs and parental investment associated with early intervention, it is crucial that we move away from a "one size fits all" service provision model, and find ways to tailor a child's intervention to their specific needs, choosing therapy approaches based on the child's individual profile at diagnosis (17–19).

During the last decade, Early Start Denver Model (ESDM) has emerged as a promising EIBI approach in the framework of the Naturalistic Developmental Behavioral Interventions (NDBI, 7), notably because of its overall effectiveness, its emphasis on natural environment teaching, comprehensive learning objectives and parental involvement. ESDM intervention has originally been implemented in an individualized setting (one therapist for one child), but other applications have been developed such as G-ESDM where one therapist works with a little group of children and P-ESDM where parents/caregivers actually provide the intervention under supervision. In their 2010 landmark randomized controlled trial (RCT, 2), Rogers and Dawson reported a mean increase of 18 IQ points in a sample of 24 toddlers with ASD receiving the ESDM intervention over two years. Subsequent studies have replicated these results (for a review see 18) and demonstrated the cost-effectiveness of ESDM intervention (18,19). Overall, the ESDM approach has been shown to significantly increase cognitive, communication and adaptive skills at the group level (20). However, the inter-individual variability in child response to treatment (RTT) is high, as with all types of EIBI (21).

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

To date, research about RTT in ESDM remains sparse and most studies focusing on homogeneous and individualized therapy settings comprised limited sample size. To better understand heterogeneity in treatment outcome, Vivanti et al. (22) attempted to identify predictors of RTT in children receiving G-ESDM intervention. 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 the functional use of objects at baseline. This study offered insight into how children with certain baseline competencies might progress faster in a G-ESDM setting. However, outcomes were assessed after only one year of intervention and baseline measures used in this study were based on original tasks (i.e., specially developed for this study and not available in the common practice), making its results poorly reproducible. In addition, its group setting makes its conclusion hardly generalizable to the canonical individualized setting of ESDM. To the best of our knowledge there are no identified child characteristics or symptom profiles based on standardized clinical assessments that are known to predict child response to the individualized ESDM intervention. It is therefore currently not possible to know whether a child will respond or not to the ESDM intervention when advising it. Again, given the considerable investment that intensive early intervention represents for therapists, families, and healthcare systems (19,23), the ability to predict whether it will work for a child would drastically change how early autism treatment programs are prescribed.

In the present study, we aim to identify early predictors of outcome after receiving a specific and homogeneous EIBI. To do so we explored the developmental trajectories of 47 preschoolers with ASD who completed two years of individualized and intensive (20 hours per week) ESDM intervention available in Geneva, Switzerland. We used a longitudinal single group design without a control population, similar to previous studies in the field (21,22,24). Indeed, because of ethical as well as logistic considerations, a random referencing to either the ESDM intervention program or any other community treatment was not achievable. We first investigated if our sample's outcome data, in terms of cognition, symptom severity and adaptive functioning, reflected findings described in the ESDM literature. We then parsed the heterogeneity in our sample's outcome by using cluster analysis (CA) and cognitive scores as the main outcome measure. CA highlighted three different groups based on cognitive outcome. We further explored baseline differences as well as early rates of change between the three groups to identify potential predictors of two-year treatment outcome.

Method

Participants

Our original sample included 52 participants who completed two 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 were thus based on the

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

data collected from 47 participants (see Table 1). No children were affected by either Fragile X, Rett, Phelan-McDermid syndromes or neurofibromatosis. No participant suffered from any major somatic disorders or epilepsy. 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 (1) criteria and Autism Diagnosis Observation Schedule-Generic (ADOS-G, 29) or 2nd edition (ADOS-2, 30) diagnosis cut-offs. Inclusion criteria to the intervention program and research study included a confirmed diagnosis of ASD, being able to participate for two full preschool years prior to age of school entry (which in Switzerland is at 4 years old by July 31st), and parental consent. Enrollment also depended on place availability at time of referral. The parents of all participants gave their written informed consent to the research protocol that was approved by the institutional review board of the University of Geneva. All participants were assessed in the context of the ongoing longitudinal Geneva Autism Cohort study. Twenty-two children from this same sample were already included in a previous study measuring outcome after one year of ESDM intervention (27). Baseline evaluations were completed at the start of the intervention and comprised behavioral measures that are detailed below. Parents also filled out questionnaires regarding medical history, as well as demographic information detailed below. Children were then assessed at 3 other time points at 6, 12 and 24 months of therapy, for a total of 4 assessments. Post-intervention data about subsequent school placement and support needs were collected. Children went onto either regular educational classrooms with varied levels of in-class paraprofessional support or special education classrooms.

Intervention

The 47 participants were enrolled in one of the 4 units of the *Centre d'Intervention Précoce en Autisme* (CIPA) in Geneva, Switzerland (Fondation Pôle Autisme (http://www.pole-autisme.ch) & Office Médico-Pédagogique), where they received 20 hours a week of daily, individual 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 (7,28). 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, and continued parent support sessions as needed throughout the two-year period. 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 (28). Today, the team consists of 20 credentialed ESDM therapists, and the program is overseen by an ESDM certified trainer.

Measures

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

The ADOS (which refers to the ADOS-G and its later version, the ADOS-2), is a standardized assessment which comprises a series of semi-structured social presses aimed to elicit and measure ASD symptoms (25,26). The schedule comprises 5 different modules, adapted to the person'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 (29,30). The ADOS were administered by a trained examiner and filmed. The members of the team who rated the video recordings 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 (31). It measures the child's development in five 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 cognitive, motor, and adaptive domains in children 2 to 7 years of age (32). These domains include EL, RL, FM and cognitive verbal and preverbal (CVP). The PEP-3 as well as the MSEL were administered by psychologists following the standard instructions of both evaluations.

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 2006 by Lord and colleagues (33). The composite DQ was computed by calculating the average of all four subdomains' developmental ages, then dividing by the chronological age and multiplying by 100. Similarly, DQ scores were computed for the subdomains of the PEP-3 that assess domains equivalent to those of the MSEL, namely EL, RL, CVP and FM. The PEP-3 composite DQ was derived using the same formula as described for the MSEL, and has already been used for the PEP-3 subdomains in previous studies (27). For our analyses of cognitive skills, we used the MSEL Early Learning Composite DQ. Since 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 scores by their equivalent DQ scores 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 stagnation but rather continued developmental progress. Also, a loss of DQ over time does not necessarily imply regression (a loss of skills) but rather slower skill acquisition, leading to a widening of the gap between the child's current abilities and what would be expected in typical development.

The Vineland Adaptive Behavior Scales – 2nd edition (VABS-II) is a semi-structured interview administered by a trained clinician that assesses a person's adaptive behavior (34). The

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

domains assessed comprise communication, socialization, daily living skills (DLS) and motor skills. An overall adaptive behavior composite score (ABC) of all these 4 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 participant socio-economic 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 (35). We then divided these categories into two groups: 1) elementary school or high school completed, and 2) college and/or graduate degree completed.

Rate of change

For all behavioral measures acquired longitudinally (ADOS, VABS-II and DQ), we computed an individual rate of change using the following Symmetrized Percent 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, limited sensitivity to outliers and equivalent consideration of bothB_x and B_y measures (36).

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

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

months, and 24 months after the start of the intervention services. Greenhouse-Geisser correction was applied whenever the assumption of sphericity was violated according to Mauchly test.

Until now, methodological strategies to identify intervention-specific predictors of EIBI outcome include whole-sample correlations between baseline and outcome measures (22,37), comparison between subgroups defined based on an arbitrary cut-off such as rapid vs slow learners (38) or best vs non-best outcome (39,40). A promising alternative relies on the identification of distinct phenotypic subgroups within ASD (41). Defining more homogeneous subgroups based on behavioral characteristics in a data-driven manner can be achieved by applying cluster analyses (CA), a strategy that has already been used in ASD preschool studies (for a review see 27). To date, CA has only been applied once on children with ASD participating in an EIBI program (Applied Behavioral Analysis: ABA) with a special focus on language development (43). We here performed a cluster analysis (CA) using cognition (assessed with the composite DQ measure) as our main outcome measure. There are several reasons why we chose DQ over other parameters. First, it is generally the main outcome measure reported in early intervention studies, and it displays the most variability (2,44,45). Second, cognition has been shown to be the domain that improves the most after early intervention (46). Third, studies investigating possible ASD subtypes within ASD have shown that the most salient group differences emerge when categorized by cognitive skills (47). We used a *k*-means clustering approach to identify subgroups in terms of DQ trajectories with a maximal number of iterations set to 10 (48). We chose two variables that capture individual DQ trajectories: the composite DQ at baseline and the composite DQ SPC over the two-year intervention period. To objectively determine the number of clusters *k* we used a two-step clustering approach as suggested by Kodinariya (49). We used the two-step clustering algorithm developed by Chiu et al. (50) 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. Finally, the Akaike's information criterion is used to objectively determine the optimal number of clusters.

The 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 therefore named "higher cognitive at baseline" (HC, n = 18). Its average DQ at baseline was 78.8 ± 10.9 with a range 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 between 64 and 124. The second "optimal responders" cluster (OptR, n = 20) was characterized by high rates of progress within the two-year program. DQ at baseline was 50.0 ± 10.8 with a range 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 between 33 and 130. The third "minimal

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

responders" cluster (MinR, n = 9) 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 between 32 and 60. The average loss was 8.5 for a final DQ of 37.5 ± 8.9 with a range between 27 and 58. The OptR and MinR subgroups did not differ in composite DQ at baseline, with an average of 50.0 and 46.0 respectively. Together, they form a group of children with lower cognitive scores (LC) at baseline. Cluster 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. HC: Higher Cognitive. MinR: Minimal Responders. OptR: Optimal Responders. 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 the 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-tests 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.

Finally, we focused on the two LC 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 OptR and MinR on post-hoc T-tests at baseline. Then, we performed a multivariate logistic regression using the selected measures. Whenever a composite score and 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

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

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 (both RRB and SA) were stable during the second year of intervention. In parallel, participants' developmental scores improved. This improvement was significant in all subdomains (i.e., VR, RL and EL) except for fine motor (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 adaptive behavior composite (ABC) scores over time. Nevertheless, participants made significant gains in the communication subdomain which occurred both during the first and the second year of intervention. Participants also increased their delay in the motor skills subdomain. 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 public regular education classroom with individual educational support. One participant (2.1%) joined a regular education classroom without any support. Four children (8.5%) entered a private school that provided in-class support in a regular education classroom. Finally, 13 participants (27.7%) entered special education program within the public-school system.

Table 1 Sample characteristics over the two years of ESDM intervention. Scores at 6 months (in italic) are indicative and were not used in the repeated measure (R.M.) ANOVA. Scores with a significant difference are highlighted in bold*; *p<.05 ** p<.01 ***p<.001.* G*: Greenhouse-Geisser correction applied, ESDM = Early Start Denver Model*

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

323

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*** **	.05
ADOS CSS SA	**7.7 (2.0)**		**5.32 (1.8)**	**5.8 (1.9)**	**<0.001*** (G)**	.365	**<0.001*** **	**<0.001*** **	.26
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*** **	**.00**
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** **	.17
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*** **	.2
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*** **	
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*** **	
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*** **	
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 %)								

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Parsing the heterogeneity in treatment response

Difference between the three subgroups at baseline

We found no differences between clusters for parental educational attainment 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 HC showed higher scores in all DQ subdomains compared to both other clusters. Considering adaptive behavior, HC exhibited higher scores in ABC as well as in the communication subdomain compared to both other clusters. HC also showed a higher score in socialization compared to MinR. 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. In the ADOS subdomains, we found that HC exhibited lower RRB compared to MinR. Besides, there was no difference in any clinical measure between MinR and OptR.

Figure 2: A. Statistically significant differences in DQ subdomains between subgroups at baseline. **B.** Statistically significant differences in VABS-II ABC and VABS-II subdomains between subgroups at baseline. *p<.05. **p<.01. ***p<.001. ABC: Adaptive Behavior Composite. DQ: Developmental Quotient. EL: Expressive Language. FM: Fine Motor. HC: Higher Cognitive. MinR: Minimal Responders. OptR: Optimal Responders. RL: Receptive Language. VABS-II: Vineland Adaptive Behavior Scale. VR: Visual Reception.

Differences between subgroups in rates of change over 6, 12 and 24 months of intervention

We found that over the two years of therapy, OptR exhibited higher rates of change compared to the other two subgroups in cognition (composite DQ as well as VR, FM and RL subdomains) as well as in adaptive behavior compared to MinR (VABS-II ABC, in socialization, communication and DLS subdomains) (see Supplementary table 2). Differences in rates of change for all these domains except for socialization, were already present within the 12 first months of therapy. Also, we found that MinR 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 in the rates of change were already present during the first 12 months of therapy, except for socialization, VR and EL.

Finally, we found that OptR already exhibited faster rates of change in composite DQ, and adaptive communication compared to MinR (Fig 3) after only 6 months of intervention. During the same early period, HC exhibited higher SPC in receptive language (RL) DQ compared to OptR.

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.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure 3: A. Statistically significant differences between subgroups in the rates of change of behavioral measures (DQ and VABS-II) within the first 6 months of intervention. **B.** Statistically significant differences between subgroups in the rates of change of behavioral measures (DQ and VABS-II) within the first 12 months of intervention. **p*<.05. ***p*<.01. ****p*<.001. ABC: Adaptive Behavior Composite. DLS: Daily Living Skills. DQ: Developmental Quotient. EL: Expressive Language. FM: Fine Motor. HC: Higher Cognitive. MinR: Minimal Responders. OptR: Optimal Responders. RL: Receptive Language. SPC: Symmetrized Percent Change. VABS-II: Vineland Adaptive Behavior Scale. VR: Visual Reception.

Subgroup classification based on early rates of change

Minimal responders (MinR) and optimal responders (OptR) showed no difference on the composite DQ at baseline and were both considered to have lower cognitive scores at baseline (LC). 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 differences between the two subgroups in any of the clinical measures at baseline, no logistic regression was performed on this time point. Within the first 6 months of therapy, MinR 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 MinR and OptR with a 95.5 % overall correct classification rate (see Fig 4A). All the 17 OptR included in the model and 4 out of the 5 MinR 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, OptR 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 OptR and MinR (see Fig 4B). All the 19 OptR included in the model and 8 out of the 9 MinR 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 a child as being an OptR or MinR with more than 95% of accuracy after six months of intervention based on the child's rates of change in the VABS-II communication and composite DQ, and after 12 months of intervention based on their rates of change in VABS-II ABC and composite DQ.

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

Discussion

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

In the present study, we analyzed data from one of the largest samples of children who underwent two years of intensive (20h per week) and individualized ESDM intervention to identify predictors of their developmental outcome. Overall, we observed that, preschoolers in our sample made significant cognitive progress and adaptive skill gains over the two years of intervention (see Fig 1A). Improvements in the current sample allowed 72.3% of the children to enter a regular education classroom post-intervention, which in Geneva requires the child to have near peer-level functioning. These results are consistent with those reported in other studies on ESDM-based intervention (2,51). More specifically, our sample exhibited an average change in DQ (+20 points) that is very close to the one described in the randomized controlled trial (RCT) study by Dawson et al. (2), which reported 18 points of cognitive gain, an average significantly greater than that of their control group. In parallel, a naturalistic study that explored developmental trajectories in preschoolers with ASD who were not enrolled in any specific therapeutic program reported an average DQ gain of only 6.3 points between 24 and 48 months of age (52). The average initial DQ in the cited study (63.6 ± 11.5) was similar to ours (60.3 ± 18.0). Considering similarities in the outcome between our results and previous ESDM studies as well as differences with naturalistic studies, one can infer that ESDM intervention in our study had a causal effect on the improvements observed at the whole group level. These results therefore highlight the possibility of implementing ESDM in Europe as effectively as in the US, despite differences in culture and health care system. They also support the cost-effectiveness of ESDM intervention, with improvements in cognition and adaptive behavior known to reduce subsequent school-based support needs, offsetting costs associated with early intensive intervention (17–19).

This study also aimed to determine whether preschoolers with ASD who participated in a two-year EIBI intervention program (here ESDM in an individual setting) could be separated into distinct subgroups based on their cognitive trajectories over time. To achieve this, 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 groups: ~38% (n = 18) of children with a mild cognitive delay at baseline that displayed a globally good outcome (Higher cognitive at baseline: HC), and two groups of children constituting the lower cognitive scores at baseline group (LC) that had very different outcomes. The first group of LC, which represented ~43% (n = 20) of the entire sample, underwent significant cognitive and adaptive skill improvements (Optimal responders: OptR) while the second group of LC, which represented ~19% (n = 9) of our sample, showed slower overall progress, and saw a widening of the developmental gap over time in cognition and adaptive behavior compared to same aged peers (Minimal responders: MinR). The clear distinction between toddlers with mild cognitive delay (HC) and those with the more severe cognitive delay at baseline (LC) observed in our sample is also reported in previous studies that applied CA to preschoolers with ASD (53–55). These studies identified at least two subgroups categorized as “high” and

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

"low-functioning" based on early cross-sectional cognitive measures. One of the main differences in the present study is that we included a longitudinal variable in our CA (i.e., the rate of cognitive change) and were therefore able to further define our subgroups of LC children based on individual cognitive trajectories over time that a cross-sectional CA would have failed to capture. In our second analysis, we aimed to uncover potential predictors of outcome by evaluating how we can predict a child's cluster membership. Amongst the LC subgroups (MinR and OptR), we found neither behavioral nor demographic differences at baseline (see Supplementary Table 1). However, we noted a significant difference between 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 rates of change predicted at 95% attrition to either the MinR or OptR subgroups. This means that despite the absence of differences at baseline that could predict outcome in LC children, early, rapid developmental progress by 6 months of intervention allowed an accurate classification of subsequent developmental pattern.

Our analyses of the HC subgroup suggest that a mild cognitive delay (78.8 ± 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 the course of treatment. In addition, children in the HC group exhibited higher levels of adaptive skills compared to other subgroups (84.9 ± 9.2 of ABC) at baseline, especially in the VABS-II domain of communication (85.7 ± 11.0). All HC children except for one were able to continue into a regular education classroom following the intervention. With a DQ of 64 at both the beginning and the end of the intervention, this child was the only one in the HC group with a DQ value lower than 80 at the end of the intervention. Overall, our HC subgroup results suggest a positive outcome in preschoolers with a mild delay in cognition and communication at baseline after receiving an individualized and intensive ESDM intervention. A recent review concluded that a higher cognitive level at baseline is a good predictor of positive outcome after various types of EIBI (56). Also, previous studies focusing on another type of intervention (Applied Behavioral Analysis: ABA) reported that higher abilities in adaptive behavior (38,57) as well as in language (58) constitute predictors of good outcome. This might suggest that mild delays in cognition and communication could represent a common predictor of good outcome among various EIBI approaches. These findings will need to be further explored with future RCT that assess the specific causality of ESDM intervention within these results. A practical implication of our findings concerning the HC subgroup is that clinicians who refer a toddler with a mild developmental delay at baseline to an ESDM program can be relatively confident that there will be a good outcome in cognition and adaptive behavior by the end of the intervention.

Apart from the HC subgroup, the rest of the sample included children presenting a severe cognitive delay at baseline (Lower Cognitive, or LC). These children presented drastically different cognitive trajectories of change over time and were attributed to two distinct

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

subgroups: OptR and MinR. Despite their severe cognitive delay at baseline (average DQ of 50.0 ± 10.8), the 15 children that composed the OptR subgroup greatly improved their cognitive and adaptive skills over time and 75% of them were able to join a regular education classroom with in-class support. On the other hand, the 9 children in the MinR subgroup had a similar level of cognitive impairment at baseline (average DQ of 46.0 ± 8.3), however their cognitive and adaptive functioning scores did not improve over time. Furthermore, the developmental gap between continued to widen, despite receiving intensive early intervention. Only 2 out of 9 (22.2%) MinR children joined a regular education classroom following the intervention. Here, we did not identify any differences (i.e., in cognition, adaptive behavior or autism severity) that could accurately discriminate between the OptR and MinR subgroups at baseline. There is currently no clear clinical marker within LC toddlers that could help clinicians predict how the child will respond to ESDM intervention. One clinical implication of our analyses of OptR and MinR at baseline is that the OptR constituted most of the LC children (~69%) thus supporting the *a priori* that most toddlers who present with lower cognitive scores at intake display a positive outcome after receiving an individual intensive ESDM intervention. Nonetheless, a better understanding of the factors (behavioral, biological, and environmental) that are associated with MinR remains necessary to develop more targeted clinical recommendations. For instance, future studies including more participants with comorbid conditions such as epilepsy should investigate whether they represent moderator of outcome. Furthermore, they could determine whether the additive effect of various genetic mutations may moderate the outcome (59). The observation of two distinct trajectories of change in children with larger cognitive impairments at baseline could shed new light on the inconsistencies that exist between various studies that measured cognitive response to EIBI within LC preschoolers with ASD. For instance, one previous study concluded that children with this kind of profile only improve in fine motor skills and receptive language but not in adaptive behavior after receiving an early and intensive ABA intervention (60). Other studies focusing in ESDM reported an association between lower cognitive level at baseline and high cognitive gains (61,62). One can hypothesize that the inter-individual heterogeneity of outcome reported by previous studies (60–62), as well as the differences in their results were potentially due to the existence of two latent subgroups (MinR and OptR) 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 EIBI to better describe the phenotypic heterogeneity within LC preschoolers with ASD.

Finally, our results suggest that despite the lack of significant differences at baseline between OptR and MinR children, the outcome after two years of intervention can be predicted by the end of the first year of intervention with high accuracy. The OptR group's rates of change appear to be significantly higher than those observed in the MinR group within the first year for cognitive and adaptive skills (especially in communication and daily

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

living skills). Our results are in line with those of Sallows et al. (38), who reported cognitive gain during the first year of an early and intensive ABA intervention as one of the best predictor of outcome at the end of the intervention. Furthermore, we found that based on the rates of change in cognition and communication within the first six months of intervention, we could infer the outcome after two years of intervention. Together, these conclusions might shed light on the timing of RTT, and when children can be considered as "*non responders*" as raised by Vivanti et al. (13). Indeed, our analyses suggest that the first 6 months to a year of intervention offers critical information about how a child will respond to ESDM intervention over time, and leads us to question whether an early, clear response to intervention can predict an optimal response overall. The emphasis on this early response to intervention as a predictor of long-term outcome has several clinical implications. One of them would be the importance of implementing regular, standardized follow-ups to measure children's cognition and adaptive behavior in the first 6 and 12 months, in addition to the systematic ESDM *Curriculum Checklist* that is currently used in the model. An alternative could lie in the development of a standardized way to use the ESDM-CC to track the rate of developmental change and ultimately the post-intervention outcome. This type of early standardized follow-up could potentially alert the clinician of difficulties a MinR child might face and give the child's team more time to consider alternative supports or enhanced service provision long before the end of the intervention. More research is needed to understand what supports or program enhancements would allow a child with a slower response to intervention to have a more optimal outcome.

In summary, our results show that despite the lack of reliable outcome predictors for children with ASD who present severe cognitive delays at baseline, the consideration of their early dynamic behavioral parameters may help predict their overall response to intervention. Further RCTs that explore the trajectories of subgroups similar to ours are needed to determine the precise effect of the ESDM on children with MinR and OptR profiles. More specifically, we need to understand whether ESDM helps OptR improve their outcome or if it prevents MinR from falling even further behind developmentally, or both. 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 OptR participants would have been MinR if they had not undergone an ESDM based therapy. Future research on the specific effects of ESDM on each subgroup could result in improved therapeutic guidelines that are more tailored to each child's individual developmental trajectory. Our study provides relevant variables that should be explored by future research at the beginning and during the very first months of an ESDM intervention.

Limitations

Despite being one of the largest samples of preschoolers who benefited from a two-year intensive and individualized ESDM program, the sample size of the present study limits the number as well as the size of subgroups that can be detected by a cluster analysis.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

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 (63,64). It is possible that studies performed on larger samples could achieve more fine-grained subgrouping on a similar population based on the same measures and could lead to bigger subgroups, in turn increasing the statistical power to detect differences at baseline between lower cognitive clusters that we could not highlight.

Another limitation that is a direct consequence of the previous one lies in the choice of the main outcome. We chose parameters related to cognitive skills as the main clustering factors. However, it would have been possible to use other measures such as level of ASD symptoms, adaptive skills or even a combination of these two. The inclusion of more variables in the model could help in defining a larger number of clusters and therefore increase our understanding of the heterogeneity of ASD in a refined manner. However, this was not possible in the present study, because of the limited sample size. The addition of more variables in the model and the multiplication of clusters would have violated the cluster analysis assumptions, making its interpretation invalid. Studies with larger samples should include more clinical parameters and could also use outcome variables suggested by parents (65).

A last limitation here lies in a lack of a non-ASD lower cognitive (LC) control group making difficult to evaluate the causality of ESDM intervention in the observed outcome of this specific population. Nevertheless, Hedval et al. 2014 reported that 87.7% of the preschoolers with ASD and LC at baseline (<70 of DQ) still had a DQ lower than 70 when assessed after two years without receiving any EIBI (66). Moreover, their delay in adaptive functioning worsened in all the VABS-II subdomains except for communication at the group level. In contrast, in the present study LC children with similar developmental pattern (MinR) only constituted 31% of our LC group, while the other LC participants (OptR) exhibited large improvements in DQ as well as in adaptive behaviors. Considering these results, one can infer a causal effect of ESDM in the progress made by children with important cognitive delay at start. The specific effects of ESDM compared to other types of EIBI still needs to be addressed with future RCT.

Conclusion

In this study, we applied a cluster analysis to the largest European sample of preschoolers with ASD who participated in an ESDM program for 20 hours a week over a two-year period. Overall, we found that ASD symptom severity decreased, and cognitive delay improved over the intervention period. Furthermore, the cluster analysis suggested three main patterns of cognitive trajectories over time. First, children who displayed mild cognitive and adaptive behavioral delays at baseline tended to have a good developmental prognosis, finishing their two years of early intervention with cognitive and adaptive behavior scores within the

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

normal range. Second, children who presented with severe cognitive delays at the start of their early intervention exhibited two dramatically different patterns of developmental trajectories. About a third of these children continued to fall behind developmentally, despite intensive therapy services. The two remaining thirds of the children, who presented with lower cognitive and adaptive behavior scores at the beginning of treatment, exhibited early and important gains in cognition and adaptive behavior which continued for the duration of the 2 years of intervention. We did not find any significant demographic or behavioral differences between these two lower cognitive subgroups at baseline, suggesting that it is not yet possible to predict how a child with a more severe presentation of ASD will respond to treatment based solely on developmental assessment measures commonly used at intake in clinical practice. Nevertheless, our results suggest that it may be possible to predict, after only 6 months of early intervention, and with very high levels of accuracy, whether a child will have an overall minimal or optimal response to treatment, based on their early gains in cognition and communicative adaptive behavior. These results advocate for close monitoring using standardized cognitive and adaptive behavioral testing during the first six months of intervention, especially for children that exhibit a clinically significant cognitive delay at baseline. Having an understanding early-on of how a child is responding to early intervention could alert clinicians and parents to the need to adapt and enhance the child's treatment plan Future studies are needed to replicate these findings, and to evaluate the kinds of treatment adaptations that would optimize child outcome for each ASD subgroup. Also, there is a need for longitudinal studies that provide a long-term follow-up in the years following the end of early intervention, to be able to assess whether the patterns of cognitive profiles and response to treatment observed remain 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.S. conceived and designed the study. M.F, M.G., N.K., and F.R. participated in 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

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

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 and accepted on september the 25th, 2012.

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, Irène Pittet, Sonia Richetin, Laura Sallin, Stefania Solazzo, Myriam Speller, Chiara Usuelli and Ornella Vico Begara for their help with data collection.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

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. Corsello CM. Early Intervention in Autism: Infants & Young Children. 2005 Apr;18(2):74–85.
9. Eldevik S, Hastings RP, Hughes JC, Jahr E, Eikeseth S, Cross S. Meta-Analysis of Early Intensive Behavioral Intervention for Children With Autism. Journal of Clinical Child & Adolescent Psychology. 2009 May 19;38(3):439–50.
10. Fuller EA, Kaiser AP. The Effects of Early Intervention on Social Communication Outcomes for Children with Autism Spectrum Disorder: A Meta-analysis. J Autism Dev Disord. 2020 May;50(5):1683–700.
11. 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.
12. 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.
13. 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
14. 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.
15. Fuentes J, Hervás A, Howlin P, (ESCAP ASD Working Party). ESCAP practice guidance for autism: a summary of evidence-based recommendations for diagnosis and treatment. Eur Child Adolesc Psychiatry. 2021 Jun;30(6):961–84.
16. 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
17. 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 1;33(6):1763–72.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

18. 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.
19. Cidav Z, Munson J, Estes A, Dawson G, Rogers SJ, Mandell D. Cost Offset Associated With Early Start Denver Model for Children With Autism. J Am Acad Child Adolesc Psychiatry. 2017 Sep;56(9):777–83.
20. 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/
21. 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.
22. 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.
23. Estes A, Swain DM, MacDuffie KE. The effects of early autism intervention on parents and family adaptive functioning. Pediatr Med. 2019 Jun;2.
24. Lombardo MV, Busuoli EM, Schreibman L, Stahmer AC, Pramparo T, Landi I, et al. Pre-treatment clinical and gene expression patterns predict developmental change in early intervention in autism. Mol Psychiatry [Internet]. 2021 Aug 2 [cited 2021 Oct 3]; Available from: https://www.nature.com/articles/s41380-021-01239-2
25. 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.
26. 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.
27. 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. J Pediatr Psychol. 2018 Apr 25;
28. 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.
29. 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.
30. 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.
31. Mullen EM. Mullen Scales of Early Learning. AGS ed. MN: American Guidance Service Inc.; 1995. (Circle Pines).
32. Schopler E, Lansing MD, Reichler RJ, Marcus LM. Psychoeducational Profile: TEACCH individualized assessment for children with autism spectrum disorders (3rd ed.). Austin, TX: Pro-Ed, Inc.; 2005.
33. 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;63(6):694–701.
34. Sparrow SS, Balla DA, Cicchetti DV. Vineland II: Vineland adaptive behavior scales. American Guidance Service. 2005.
35. Hollingshead AA. Four-factor index of social status. Yale University.

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

1975;Unpublished manuscript.
36. Berry DA, Ayers GD. Symmetrized Percent Change for Treatment Comparisons. The American Statistician. 2006;60(1):27–31.
37. 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.
38. 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.
39. Sallows GO, Graupner TD. Replicating Lovaas ' Treatment and Findings: Preliminary Results 1. 1999.
40. Perry A, Cummings A, Geier J, Freeman N, Hughes S, LaRose L, et al. Effectiveness of Intensive Behavioral Intervention in a large, community-based program. Research in Autism Spectrum Disorders. 2008 Oct 1;2:621–42.
41. Ousley O, Cermak T. Autism Spectrum Disorder: Defining Dimensions and Subgroups. Curr Dev Disord Rep. 2014 Mar 1;1(1):20–8.
42. 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.
43. 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.
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. 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.
46. 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 May 1;25(4):1137–53.
47. Witwer AN, Lecavalier L. Examining the validity of autism spectrum disorder subtypes. J Autism Dev Disord. 2008 Oct;38(9):1611–24.
48. Hartigan JA, Wong MA. Algorithm AS 136: A K-Means Clustering Algorithm. Applied Statistics. 1979;28(1):100.
49. Kodinariya T, Makwana PR. Review on determining number of Cluster in K-Means Clustering. undefined [Internet]. 2013 [cited 2021 May 26]; Available from: /paper/Review-on-determining-number-of-Cluster-in-K-Means-Kodinariya-Makwana/1a34936bffe558a380168b790dc37956813514ba
50. Chiu T, Fang D, Chen J, Wang Y, Jeris C. A robust and scalable clustering algorithm for mixed type attributes in large database environment. KDD '01 [Internet]. 2001 [cited 2021 May 26];Proceedings of the seventh QCM SIGKDD internqtionql conference on Knowledge discovery and data mining. Available from: /paper/A-robust-and-scalable-clustering-algorithm-for-type-Chiu-Fang/a1ddee7a0c4885b46891d5d8f873e9388d42a4c1
51. 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.
52. Clark MLE, Barbaro J, Dissanayake C. Continuity and Change in Cognition and

medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Autism Severity from Toddlerhood to School Age. J Autism Dev Disord. 2017 Feb;47(2):328–39.

53. Fein D, Stevens M, Dunn M, Waterhouse L, Allen D, Rapin I, et al. Subtypes of Pervasive Developmental Disorder: Clinical Characteristics. Child Neuropsychology - CHILD NEUROPSYCHOL. 1999 Mar 1;5:1–23.

54. 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. J Am Acad Child Adolesc Psychiatry. 2000 Mar;39(3):346–52.

55. 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;13(5):796–809.

56. Zachor DA, Ben-Itzchak E. Variables Affecting Outcome of Early Intervention in Autism Spectrum Disorder. Journal of Pediatric Neurology. 2017 Jun;15(03):129–33.

57. 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. Am J Intellect Dev Disabil. 2010 Sep;115(5):381–405.

58. 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 1;5(3):1016–27.

59. Warrier V, Zhang X, Reed P, Havdahl A, Moore TM, Cliquet F, et al. Genetic correlates of phenotypic heterogeneity in autism [Internet]. Genetic and Genomic Medicine; 2020 Jul [cited 2021 Oct 3]. Available from: http://medrxiv.org/lookup/doi/10.1101/2020.07.21.20159228

60. 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 1;44(9):2221–9.

61. 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. Neuropsychiatr Dis Treat. 2016;12:1379–84.

62. 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;

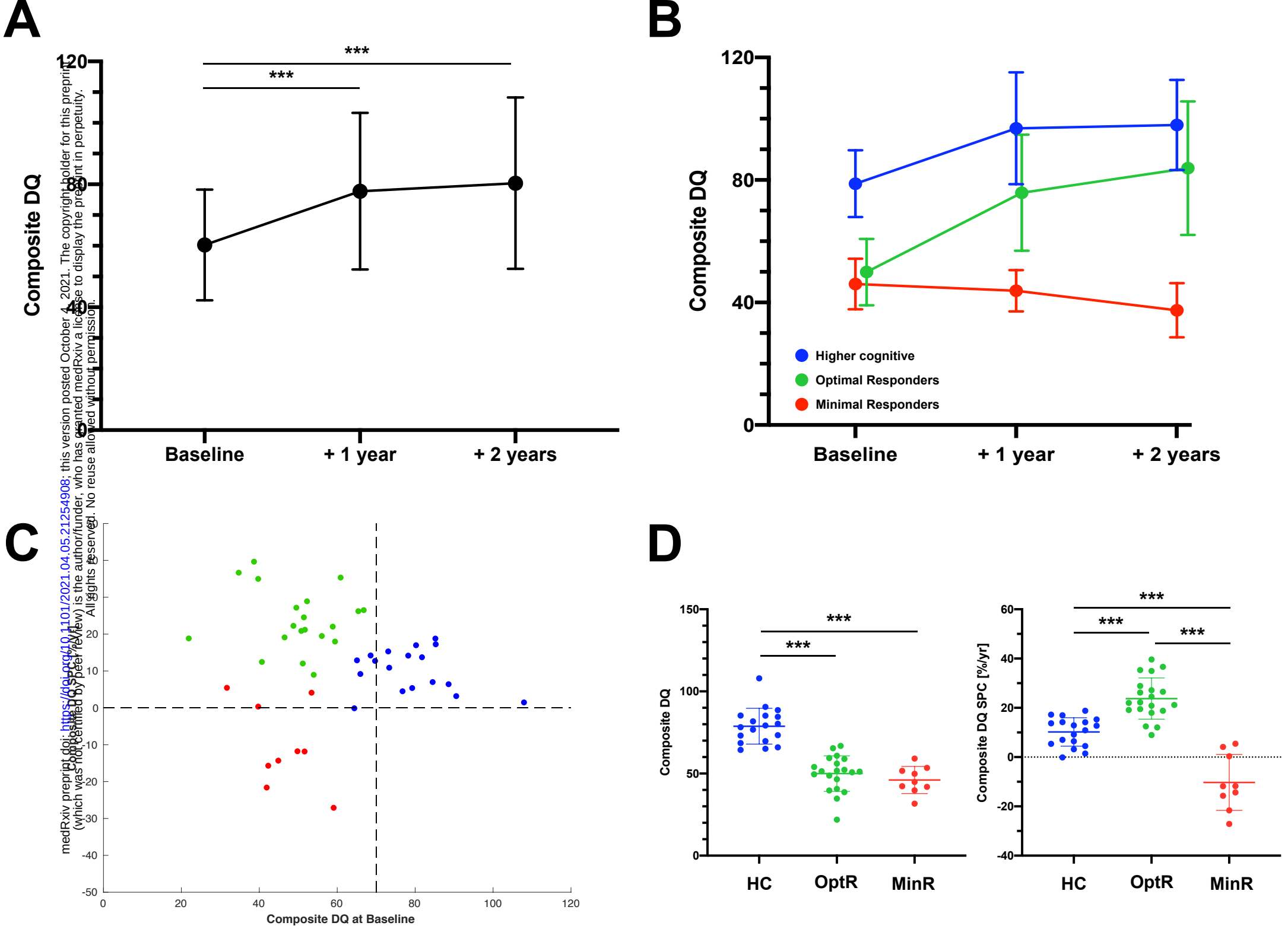
63. Formann AK. Die Latent-Class-Analyse: Einführung in die Theorie und Anwendung [Latent class analysis: Introduction to theory and application]. Weinheim. Beltz; 1984.

64. Qiu W, Joe H. Generation of Random Clusters with Specified Degree of Separation. Journal of Classification. 2006 Feb 1;23:315–34.

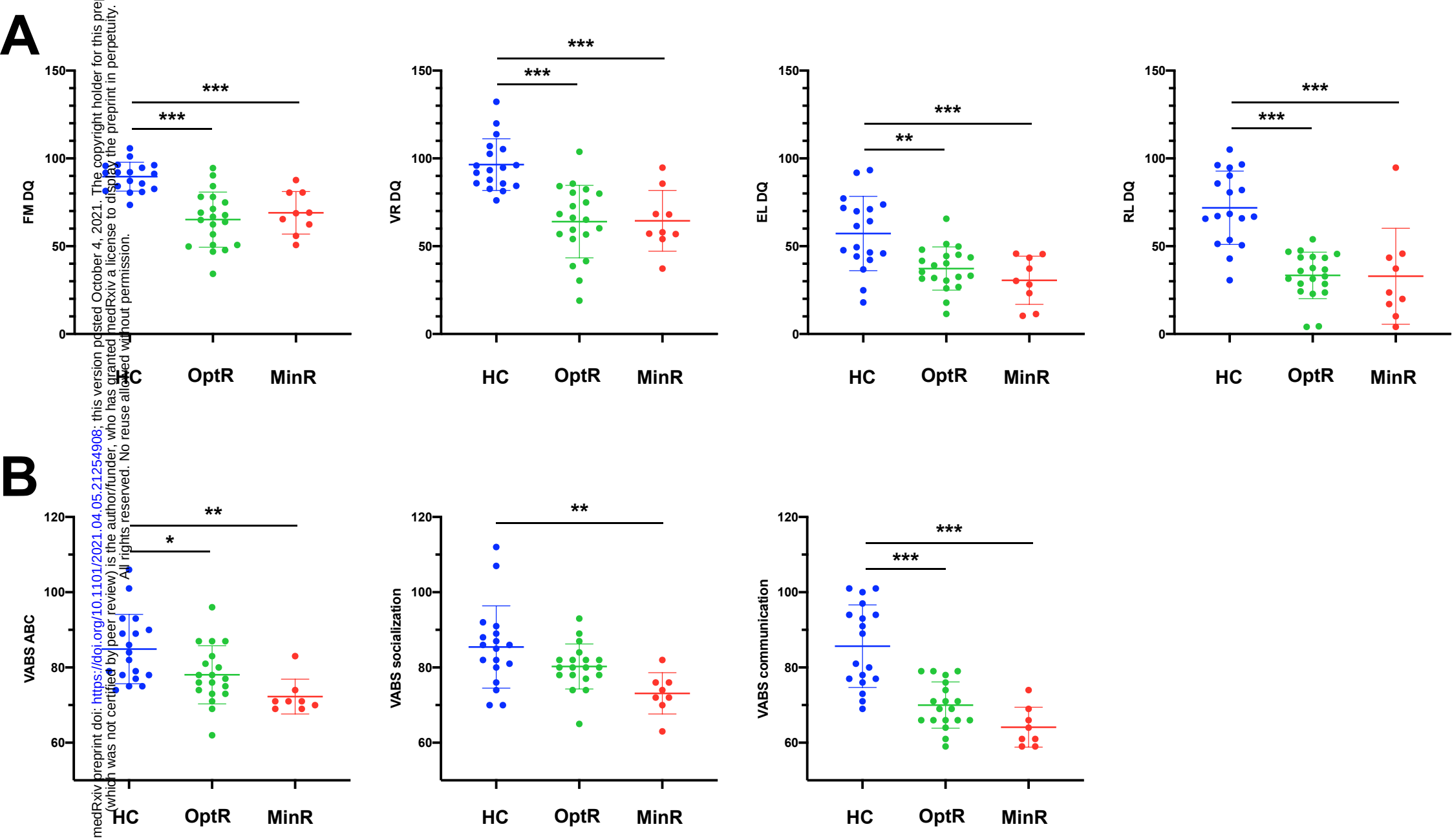
65. McConachie H, Livingstone N, Morris C, Beresford B, Couteur AL, Gringras P, et al. Parents Suggest Which Indicators of Progress and Outcomes Should be Measured in Young Children with Autism Spectrum Disorder. J Autism Dev Disord. 2018 Apr 1;48(4):1041–51.

66. Hedvall Å, Westerlund J, Fernell E, Holm A, Gillberg C, Billstedt E. Autism and developmental profiles in preschoolers: stability and change over time. Acta Paediatr. 2014 Feb;103(2):174–81.

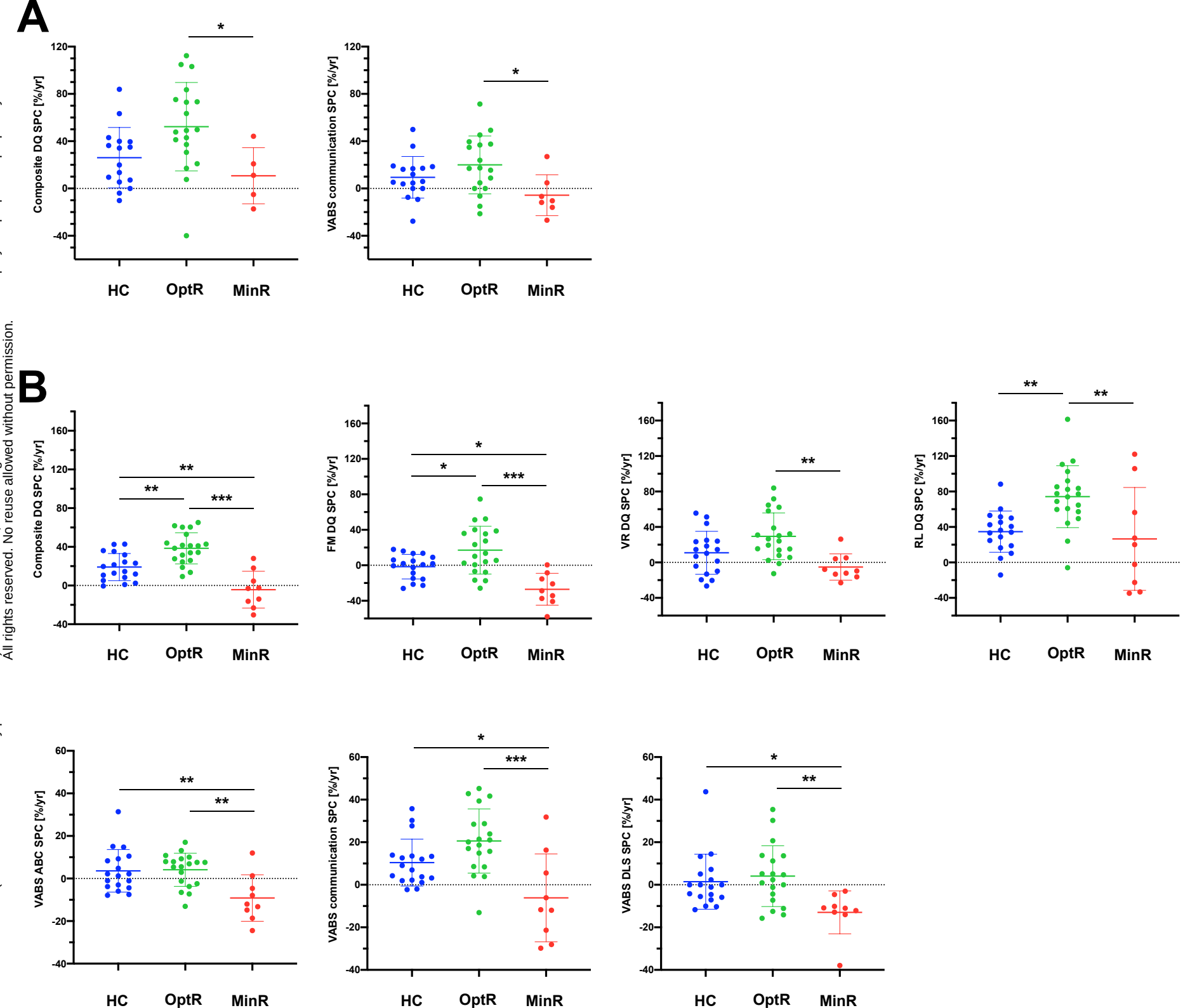
medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

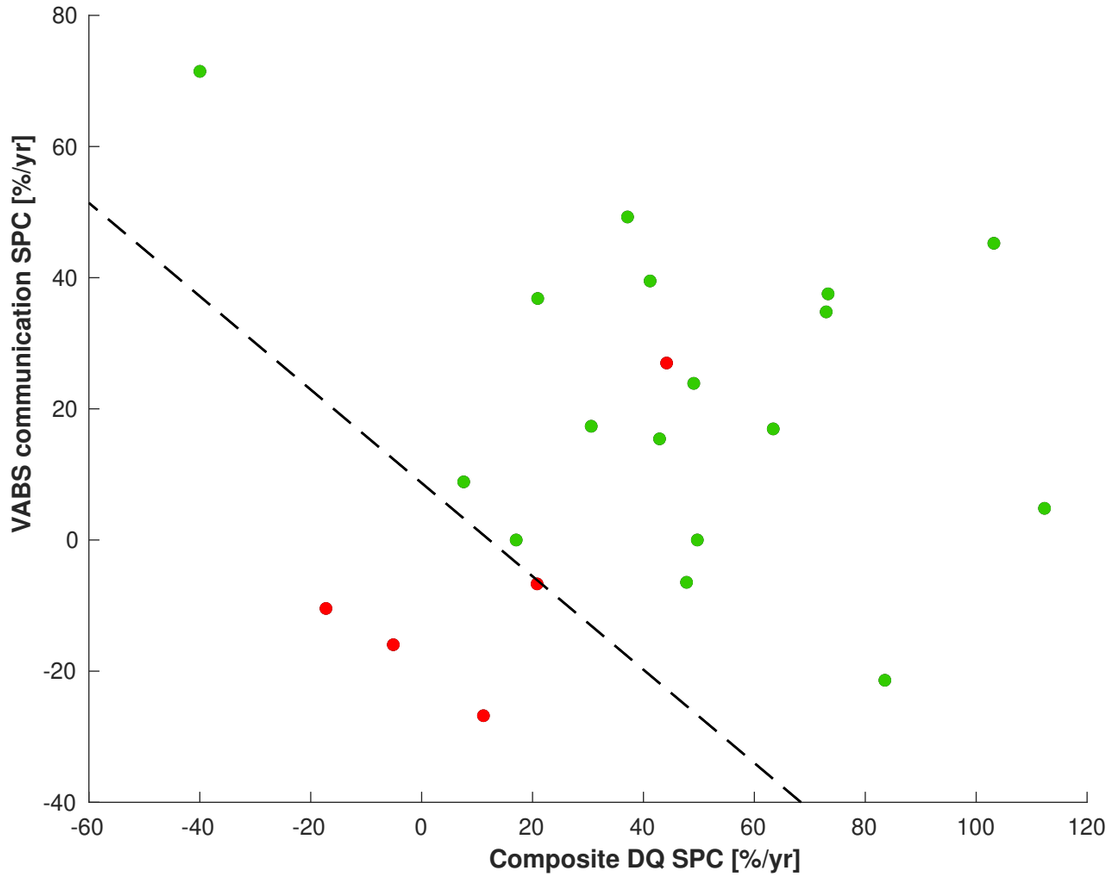


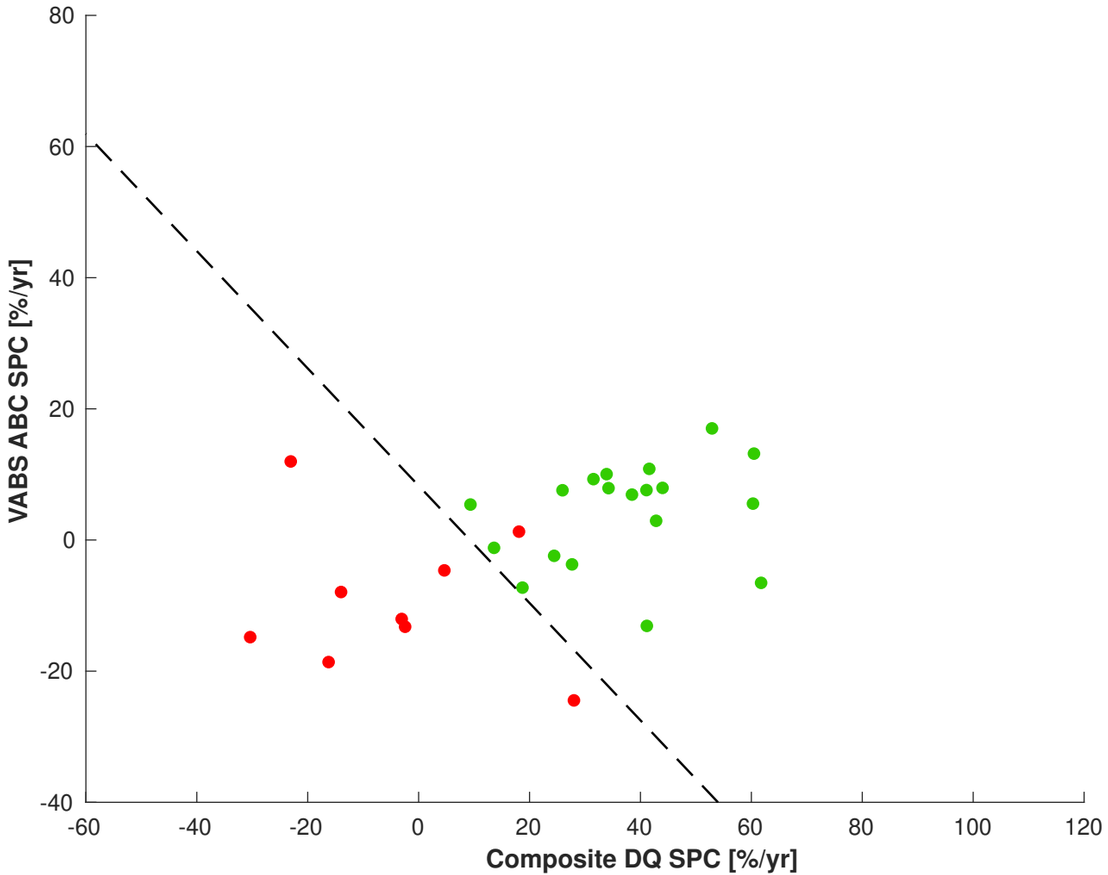
medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.



medRxiv preprint doi: https://doi.org/10.1101/2021.04.05.21254908; this version posted October 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.




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