

Supplementary Online Content

Harstad E, Hanson E, Brewster SJ, et al. Persistence of autism spectrum disorder from early childhood through school age. *JAMA Pediatr*. Published online October 2, 2023. doi:10.1001/jamapediatrics.2023.4003

eTable. Differences Between Study Participants and a Subset of Eligible Nonparticipants

eMethods. Description of Imputation Methods

eFigure 1. IQ Distributions for Children With Nonpersistent ASD and Persistent ASD, Separately

eFigure 2. Probability of Persistent ASD at Research Assessment as a Function of Baseline Vineland Adaptive Scores

This supplementary material has been provided by the authors to give readers additional information about their work.

eTable 1. Differences between Study Participants and a Subset of Eligible Non-Participants

Variable Mean (SD) or %	Study Sample (N=213)	Eligible Non-Participants (N=68) ^a	t-test/Chi- square	E.S. Cohen's D ^b
Age at Clinical ASD Diagnosis	24.6 (3.8)	24.1 (4.8)	0.75	0.11
Bayley Cognitive	81.4 (15.0)	79.1 (14.6)	-1.38	0.15
Bayley Language	65.2 (15.7)	59.4 (12.9)	2.43*	0.39
Vineland Adaptive Behavior Composite	73.5 (10.2)	67.0 (18.0)	3.44*	0.53
Vineland Communication	68.6 (16.5)	59.2 (23.5)	3.41*	0.51
Sex, male	83.1%	75%	2.21	
Race ^c			7.96	
Asian	3.9%	5.9%		
Black	7.2%	11.8%		
Native American	1.7%	0%		
Other	7.2%	17.6%		
White	80.2%	64.7%		
Income ^d			4.46	
Low	4.2%	10.3%		
Moderate	28.2%	30.9%		
Middle	55.4%	47.1%		
Upper	11.7%	11.8%		
Ethnicity ^c			9.14*	
Hispanic	9.6%	27.5%		
Not Hispanic	90.4%	72.5%		

^aWe abstracted baseline medical record data on a random subset of eligible non-participants drawn from the initial population of 1,120 eligible children (n=68, determined to allow detection of medium level standardized differences of a 0.5 standard deviation when contrasting two means) to compare them with study participants.⁴²

^bEffect size conventions are as follows: .02(small), .5(medium), and .8(large) based on Cohen, (1992). Note that these analyses were run with excessive levels of power. Specifically, for a medium effect size, power for n=68 and n=213 of a t-test was equal to 94.5%, which is excessive. Thus, we supplemented these results with indices of effect size to aid interpretation of the two groups' differences. Significance was corrected for experimenter-wise error using the Benjamini-Hochberg correction with a false discovery rate equal to the $\alpha=5\%$.

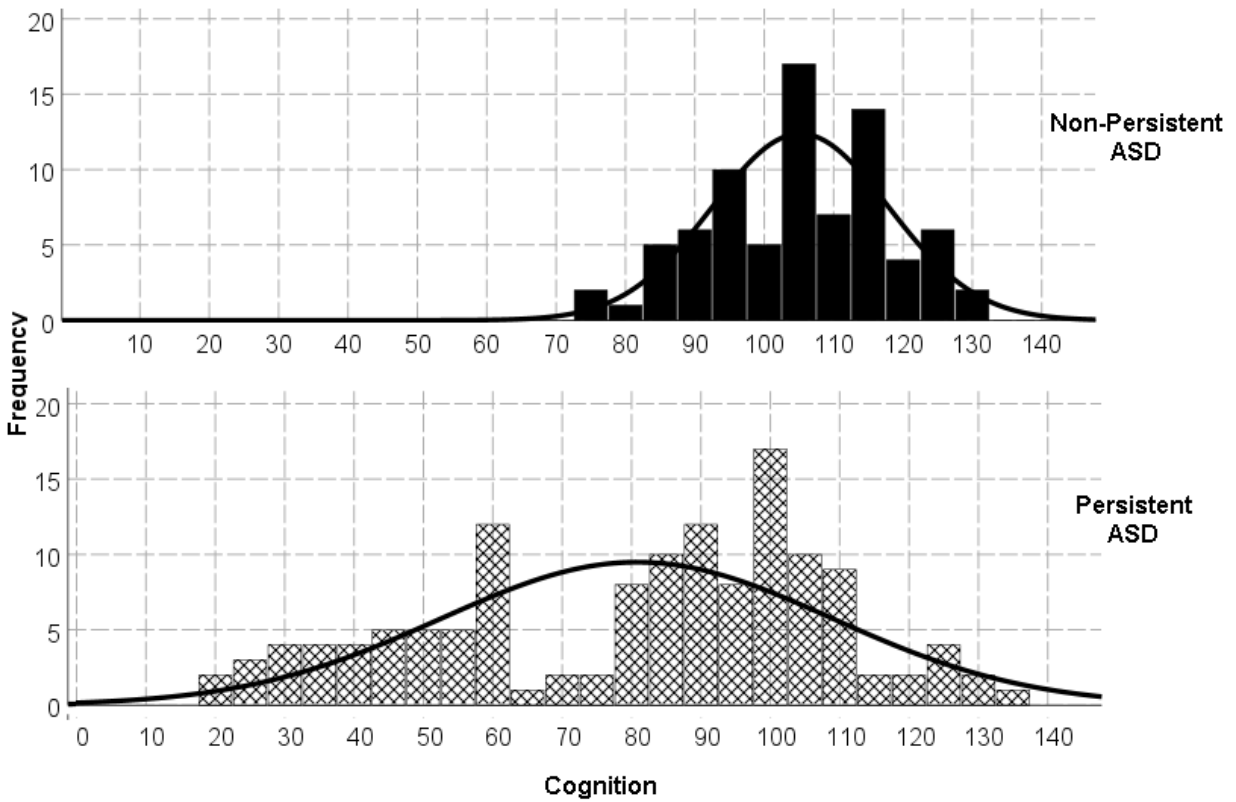
^cNote that the race and ethnicity data for this comparison were drawn from medical record data, as that was all that was available for the Non-participant group.

^dIncome was estimated via census tracts based zip code at ASD clinical diagnosis.

* $p<0.05$

eMethods Description of Imputation Methods: A multiple imputation procedure was implemented to account for the missing data in baseline cognitive, language and adaptive functioning. Specifically, the percent of missing data was 0.005% in cognition, 19.7% in language, and 12.7% in adaptive functioning. The multiple imputation procedure involved creating 50 imputed datasets with several auxiliary variables that were not part of the predictive model (e.g., PLS scores). Subsequently, logistic models were fitted for each imputed dataset and estimates were pooled across replications. The method of imputation involved maximum likelihood (ML) and the method of analysis maximum likelihood with robust standard errors (MLR) so that non-normality of the predictors would be accounted for.

eFigure 1. IQ Distributions for Children with Non-Persistent ASD and Persistent ASD, Separately



eFigure 2. Probability of Persistent ASD at Research Assessment as a Function of Baseline Vineland Adaptive Scores

