
















A global multicohort study to map subcortical brain development and cognition in infancy and early childhood

Received: 11 April 2022

Accepted: 16 October 2023

Published online: 23 November 2023

 Check for updates

Ann M. Alex ¹, Fernando Aguatero^{1,2}, Kelly Botteron³, Claudia Buss ^{4,5,6}, Yap-Seng Chong ^{7,8}, Stephen R. Dager⁹, Kirsten A. Donald ^{10,11}, Sonja Entringer^{4,5,6}, Damien A. Fair ¹², Marielle V. Fortier^{8,13}, Nadine Gaab¹⁴, John H. Gilmore ¹⁵, Jessica B. Girault ^{15,16}, Alice M. Graham¹⁷, Nynke A. Groenewold ^{10,18,19,20}, Heather Hazlett^{16,17}, Weili Lin ²¹, Michael J. Meaney⁹, Joseph Piven ^{16,17}, Anqi Qiu ^{22,23,24,25,26,27}, Jerod M. Rasmussen^{5,6}, Annerine Roos^{10,11,28}, Robert T. Schultz ²⁹, Michael A. Skeide³⁰, Dan J. Stein ^{19,28}, Martin Styner ^{16,31}, Paul M. Thompson³², Ted K. Turesky¹⁴, Pathik D. Wadhwa^{5,6,33}, Heather J. Zar^{18,20}, Lilla Zöllei³⁴, Gustavo de los Campos^{1,2,35}, Rebecca C. Knickmeyer ^{1,36} ✉ & the ENIGMA ORIGINS group*

The human brain grows quickly during infancy and early childhood, but factors influencing brain maturation in this period remain poorly understood. To address this gap, we harmonized data from eight diverse cohorts, creating one of the largest pediatric neuroimaging datasets to date focused on birth to 6 years of age. We mapped the developmental trajectory of intracranial and subcortical volumes in ~2,000 children and studied how sociodemographic factors and adverse birth outcomes influence brain structure and cognition. The amygdala was the first subcortical volume to mature, whereas the thalamus exhibited protracted development. Males had larger brain volumes than females, and children born preterm or with low birthweight showed catch-up growth with age. Socioeconomic factors exerted region- and time-specific effects. Regarding cognition, males scored lower than females; preterm birth affected all developmental areas tested, and socioeconomic factors affected visual reception and receptive language. Brain–cognition correlations revealed region-specific associations.

Early childhood (birth to 6 years) represents a dynamic and critical period in human brain development. At a cellular level, this period is marked by glial proliferation and migration, dendritic arborization, synaptogenesis, myelination, programmed cell death and synaptic and axonal elimination¹. At the cognitive and behavioral levels, several abilities, including language, memory, social cognition, emotional

regulation and executive function, emerge and elaborate. Sensory and motor skills also develop rapidly². Between the cellular and cognitive/behavioral levels lie macroscale brain properties that are best characterized by imaging-based phenotypes (for example, global, regional and subcortical volumes, cortical thickness, white matter diffusivity, functional connectivity and so on). Our understanding of

A full list of affiliations appears at the end of the paper. ✉ e-mail: knickmey@msu.edu

early development of these brain-related phenotypes has increased substantially in the past 20 years, driven by an expanding number of cross-sectional and longitudinal neuroimaging studies, particularly magnetic resonance imaging (MRI) studies^{3,4}. However, many knowledge gaps remain, including (1) very few large-scale studies have mapped the development of brain-related phenotypes from birth to 6 years of age in diverse global populations with dense data, (2) there is limited information on how sociodemographic factors and adverse birth outcomes shape neurodevelopmental trajectories, and (3) the neural correlates of variations in cognitive development are not fully understood. The present study seeks to address these gaps with a particular focus on the development of intracranial volume (ICV) and subcortical structures, including the thalamus, hippocampus, amygdala, caudate, putamen and globus pallidus.

Regarding the first knowledge gap, most studies on ICV and subcortical volume development focus on later childhood, adolescence and adulthood, with limited data below 6 years of age. Studies focusing specifically on early life generally have narrow time frames, particularly birth to age 2 (ref. 4). Furthermore, published research on subcortical volume trajectories in early childhood is often limited by cross-sectional study design, low sample size, discontinuous age ranges and/or a lack of diversity in participants⁵. Longitudinal studies are necessary to model individual developmental trajectories and to elucidate how parameters of such trajectories are influenced by birth outcomes and sociodemographic factors⁶. Also, small sample size (leading to underpowered studies) and limited racial and socioeconomic diversity reduce reproducibility and replicability of published results. The study by Marek et al. points out that brain-wide association studies require thousands of participants to accurately identify brain–phenotype associations⁷. Collaborative research overcomes these challenges by offering adequately powered study designs and recognizing differences across cultures and measurement methods. Addressing both inter- and intraindividual variability can improve the accuracy of growth curves and increase the generalizability of results^{8,9}. The recent work on brain charts was one of the largest studies to map normative brain growth across the lifespan¹⁰ and a step in the direction of inclusion of diverse global populations; the current study differs by focusing on specific subcortical structures, examining the influence of socioeconomic factors (SES) and adverse birth outcomes and examining brain–cognition associations.

Regarding the second knowledge gap, as with studies of age-related change in ICV and subcortical volumes, studies associating adverse birth outcome and sociodemographic factors with brain structure mainly focused on brain outcomes characterized in late childhood, adolescence and adulthood. Early childhood, the period when the brain is most malleable to environmental effects¹¹, has not been studied as thoroughly due to practical and technical challenges of conducting imaging studies in this age range. Ample studies show that sex^{12,13}, adverse birth outcomes (preterm birth and low birthweight¹⁴) and SES (family income and maternal education¹⁵) contribute to variation in structural brain development and cognition in late childhood and adolescence. Understanding the effects of these factors on neuroanatomical development in early childhood may reveal the earliest deviations from typical trajectories and provide guidance for interventions to prevent or reverse adverse neurodevelopment in children at risk¹⁶.

Regarding the third knowledge gap, given the dynamic nature of brain development in early childhood and concurrent development of cognitive and behavioral characteristics, it is important to understand the neural basis of cognitive abilities across this age range. ICV and subcortical structures play important roles in cognition^{17,18}, but our understanding of the neuroanatomical correlates of early cognitive development is limited. Cognitive functioning in childhood predicts later cognitive competence¹⁹, and a better understanding of its neuroanatomical correlates could help inform interventions to support early cognitive development.

The Organization for Imaging Genomics in Infancy (ORIGINS) was established to address the gaps discussed above and to facilitate large-scale genetics studies of brain structure and function during infancy and early childhood. ORIGINS is a working group within the Enhancing Neuroimaging Genetics through Meta-Analysis (ENIGMA) Consortium, a global network of greater than 2,025 scientists studying the human brain in health and disease²⁰. Here, we introduce one of the largest pediatric neuroimaging datasets spanning birth through age 6 and use it to map the trajectory of ICV, subcortical structures (thalamus, hippocampus, amygdala, caudate, putamen and pallidum) and cognitive development as indexed by the Mullen Scales of Early Learning (MSEL). We analyzed this unique dataset to investigate the effects of sex, gestational age, birthweight, maternal education and family income on trajectories of ICV and subcortical volumes and on cognitive development. Finally, we examined brain–cognition correlations in this age range. This dataset includes over 2,000 children across eight contributing sites in four countries (Germany, Singapore, South Africa and the United States). This study lays a strong foundation for understanding ICV and subcortical brain development and how it relates to early cognitive development encompassing children from diverse ethnic and socioeconomic backgrounds. This information is important as brain development in these early years establishes the path for lifelong cognitive development and psychiatric risk⁴.

Results

Developmental trajectories of ICV and subcortical structures

Demographic details of the whole sample and information regarding sample distributions for each cohort are given in Table 1.

To map longitudinal brain development, we fitted (mixed-effects) subject-specific nonlinear longitudinal growth curves to ICV and subcortical structures (thalamus, amygdala, hippocampus, caudate, putamen and pallidum). Our growth curve models have subject-specific intercepts (that is, volume at birth), asymptote and growth rate parameters. In our hierarchical model, we included effects of birth outcomes and sociodemographic factors on intercepts and asymptotes and random effects of cohort and subject. The fitted growth curves for male and female ICV are displayed in Fig. 1. Growth curves for ICV and subcortical structures as they relate to each predictor variable are found in Supplementary Figs. 2–6.

Our results show that ICV and subcortical structures follow nonlinear growth patterns from birth to age 6, with maturation ages differing across regions. We define maturation age as the age at which regional volume reaches 99% of its asymptotic value. We find that ICV matures by age 3.3–3.5 years. Among subcortical structures, the amygdala matures earliest (3.3 years), followed by the hippocampus (4.6 years), putamen (4.6 years), globus pallidus (5 years) and caudate (5.5 years). The thalamus has the most protracted trajectory, reaching maturation by 7.3–7.5 years.

Sex, adverse birth outcomes and SES influence global and regional developmental trajectories (Table 2). Males show significantly larger volume throughout the age range for all structures. Children born preterm had significantly lower ICV and amygdala, hippocampus and thalamus volumes at birth. However, the effect was no longer significant at asymptote, suggesting that volumes catch up as the child matures. Similarly, for children with low birthweight, all volumes were significantly lower at birth. This effect was also non-significant by asymptotic age.

Low maternal education was associated with lower ICV and caudate volume across the age range and with thalamus and amygdala volume at asymptote. For the putamen and globus pallidus, low maternal education was associated with larger volumes at birth and smaller volumes at asymptote. Children from lower-income families had significantly lower hippocampus, putamen and pallidum volumes throughout the age range and lower thalamus and caudate volumes only at the asymptote.

Table 1 | Demographic distribution of participants

	Sample size	Sex	Gestational age	Birthweight	Maternal education	Family income
	Individuals/ observations	Male/female	Preterm/full term	Low/normal	Primary–secondary/ tertiary	Low/medium-high/ missing
All	2,108/3,607	1,102/1,006	436/1,672	365/1,743	651/1,457	606/1,261/241
%		52.3/47.7	20.7/79.3	17.3/82.7	30.9/69.1	28.8/59.8/11.4
BCP	172/317	80/92	2/170	1/171	5/167	14/154/4
%	8.2/8.8	46.5/53.5	1.2/98.8	1/99	2.9/97.1	8.2/89.5/2.3
Boston (BCH)	130/181	66/64	3/127	6/124	3/127	9/108/13
%	6.2/5	50.8/49.2	2.3/97.7	4.6/95.4	2.3/97.7	6.9/83.1/10
Cape Town (DCHS)	135/135	70/65	9/126	5/130	128/7	45/90
%	6.4/3.7	51.9/48.1	6.7/93.3	3.7/96.3	94.8/5.2	33.3/66.7
EBDS	1,013/2,152	538/475	384/629	319/694	350/663	410/537/66
%	48/59.7	53.1/46.9	37.9/62.1	31.5/68.5	34.6/65.4	40.5/53/6.5
GUSTO	357/421	176/181	21/336	25/332	131/226	68/264/25
%	16.9/11.7	49.3/50.7	5.9/94.1	7/93	36.7/63.3	19/74/7
IBIS	86/186	47/39	1/85	1/85	0/86	12/72/2
%	4.1/5.2	54.7/45.3	1.2/98.8	1.2/98.8	0/100	14/83.7/2.3
Max Planck	127/127	73/54	10/117	4/123	5/122	0/0/127
%	6/3.5	57.5/42.5	7.9/92.1	3.1/96.9	3.9/96.1	0/0/100
UCI	88/88	52/36	6/82	4/84	29/59	48/36/4
%	4.2/2.4	59/41	6.8/93.2	4.5/95.5	33/67	54.6/40.9/4.5

BCP, Baby Connectome Project; Boston (BCH), Boston Children's Hospital; Cape Town (DCHS), Drakenstein Child Health Study; GUSTO, Growing Up in Singapore Towards Healthy Outcomes, Singapore; IBIS, Infant Brain Imaging Study Network; UCI, University of California, Irvine.

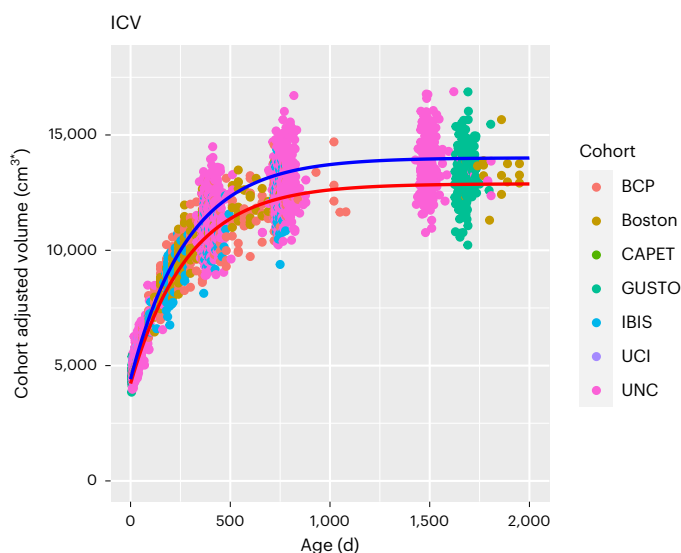


Fig. 1 | The effect of sex on the developmental trajectory of ICV. The lines represent the fitted growth curve for the male and female sex from the nonlinear mixed model regression (number of individuals = 1,835, number of observations = 3,168). Males (blue line) have significantly larger ICV than females (red line) throughout the age range studied (the *P* values derived from two-sided *t*-tests are 1.68×10^{-12} at intercept and 1.88×10^{-60} at asymptote; the Bonferroni-corrected *P* value for significance was <0.0001). Circles represent individual data points, and colors represent the cohorts; BCP, Baby Connectome Project; Boston, Boston Children's Hospital/Harvard Medical School; CAPET, Drakenstein Child Health Study, Cape Town; GUSTO, Growing Up in Singapore Towards Healthy Outcomes, Singapore; IBIS, Infant Brain Imaging Study Network; UCI, University of California, Irvine; UNC, University of North Carolina Early Brain Development Study.

Sensitivity analyses

We performed three sensitivity analyses and split-sample replication analyses to assess the robustness of predictive values of birth outcomes and SES on brain volume trajectories. First, we excluded family income from the model because of its high correlation with maternal education. Results obtained from this subset confirm those from the full model (Supplementary Table 1). However, hippocampal volume was significantly associated with maternal education in this model. Second, we controlled for interindividual variation in head size²¹ by dividing regional volumes by ICV. The only significant effect in this model was male–female difference on caudate volume, with males having smaller relative caudate volume than females (Supplementary Table 2). Finally, we excluded one twin/sibling from each pair in the University of North Carolina Early Brain Development Study (EBDS) cohort to check if relatedness between participants affected the model results. Results were like those of the main analysis (Supplementary Table 3).

Split-sample replication analysis revealed that sex and birth outcomes had robust associations with brain volumes, with most folds showing the same direction of effect and meeting the Bonferroni threshold for significance (Supplementary Table 4). Results for SES variables were more variable. Direction of effect was generally similar between folds but often did not reach the criteria for significance in both folds. This likely reflects reduced power when dividing the sample and the smaller effect size for SES–brain associations.

Cognitive and motor development

Next, we examined patterns of cognitive and motor development in these children by fitting linear mixed-effects models with fixed effects of covariates and of cohort and participants' random effects. To account for possible heterogeneity in scales, we also fitted cohort-specific error variances. As displayed in Table 3, we observed a linear increase (slope *P* values <0.01) in cognitive scores by age in days. Gross motor scores

Table 2 | Birth outcomes and sociodemographic factors have significant effects on structural brain development^a

			ICV	Thalamus	Hippocampus	Amygdala	Caudate	Putamen	Globus pallidus
Reference curve									
Reference ^b	Intercept	Effect	402,952.30	8,924.17	2,190.29	398.78	3,711.22	3,345.64	1,141.27
		(s.e.)	(2,676.98)	(42.15)	(13.16)	(5.80)	(22.54)	(51.25)	(16.79)
	Asymptote	Effect	1,285,208.00	13,292.35	5,159.63	2,082.93	6,947.72	8,535.24	2,114.39
		(s.e.)	(6,309.91)	(80.24)	(31.85)	(13.30)	(53.36)	(55.41)	(15.88)
	Growth Rate	Effect	-5.66	-6.64	-6.02	-5.60	-6.26	-6.01	-6.16
		(s.e.)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.03)
Main effects									
Male sex	Intercept	Effect	19,177.04**	423.13**	80.27**	19.84**	114.07**	226.31**	40.48*
		(s.e.)	(2,690.45)	(42.57)	(13.34)	(5.48)	(21.93)	(50.16)	(13.33)
		%	4.76	4.74	3.66	4.98	3.07	6.76	3.55
	Asymptote	Effect	112,624.10**	990.70**	365.18**	190.05**	325.09**	671.36**	182.94**
		(s.e.)	(6,517.97)	(74.91)	(32.32)	(13.37)	(53.41)	(59.57)	(16.34)
		%	8.76	7.35	7.07	9.12	4.67	7.86	8.63
Preterm birth	Intercept	Effect	-48,142.85**	-282.32**	-179.79**	-109.84**	-82.43*	-218.48*	-71.87*
		(s.e.)	(4,056.34)	(67.89)	(20.14)	(8.86)	(34.63)	(87.78)	(22.75)
		%	-11.95	-3.16	-8.2	-27.54	-2.22	-6.53	-6.3
	Asymptote	Effect	17,063.40	-184.55	-93.51	-41.81*	-143.08	-56.33	-34.43
		(s.e.)	(9,981.36)	(112.09)	(47.82)	(19.84)	(79.84)	(89.71)	(23.57)
		%	1.33	-1.46	-1.83	-2	-2.1	-0.67	-1.65
Low birthweight	Intercept	Effect	-45,116.71**	-295.56**	-176.08**	-74.96**	-185.32**	-385.05**	-96.05**
		(s.e.)	(4,263.44)	(69.66)	(20.97)	(9.16)	(35.98)	(90.67)	(24.03)
		%	-11.2	-3.31	-8.04	-18.8	-4.99	-11.5	-8.42
	Asymptote	Effect	-26,109.07*	-110.19	-87.64	-24.73	-146.68	-165.53	-69.30*
		(s.e.)	(10,355.73)	(115.98)	(49.65)	(20.63)	(82.80)	(93.09)	(24.47)
		%	-2.03	-0.93	-1.71	-1.19	-2.14	-1.96	-3.3
Low maternal education	Intercept	Effect	-16,269.64**	-186.31*	-4.71	7.27	-120.10**	383.86**	205.47**
		(s.e.)	(3,203.33)	(55.66)	(16.63)	(6.52)	(24.50)	(70.81)	(17.43)
		%	-4.04	-2.09	-0.22	1.82	-3.24	11.47	18
	Asymptote	Effect	-45,624.15**	-394.91**	-106.46	-88.94**	-298.09**	-260.66*	-92.64**
		(s.e.)	(8,944.40)	(101.29)	(44.70)	(18.62)	(71.94)	(84.23)	(22.38)
		%	-3.55	-2.94	-2.06	-4.27	-4.28	-3.03	-4.26
Low-income family	Intercept	Effect	4,526.54	-92.39	-58.86**	-13.37*	10.61	-254.62**	-51.61**
		(s.e.)	(3,174.30)	(53.61)	(16.10)	(6.46)	(25.67)	(60.30)	(14.59)
		%	1.12	-1.04	-2.69	-3.35	0.29	-7.61	-4.52
	Asymptote	Effect	-14,994.01	-398.43**	-162.85**	-45.89*	-233.97**	-336.27**	-85.68**
		(s.e.)	(8,452.62)	(95.18)	(41.57)	(17.25)	(68.13)	(77.43)	(20.71)
		%	-1.17	-2.92	-3.16	-2.21	-3.34	-3.95	-4.1
Number of observations		3,168	3,276	3,242	3,257	3,270	2,245	2,254	
Number of individuals		1,835	1,865	1,857	1,861	1,865	1,469	1,469	

^a $P < 0.05$; ^{**} $P < 0.0007$. ^bData were analyzed by two-tailed t-test with a Bonferroni post hoc test for P value significance. Exact P values are provided in Extended Data Table 1. ^bReference: female, full term, normal birthweight, tertiary maternal education and medium high-income family.

were significantly lower in children born preterm. Visual reception scores were significantly impacted by all predictor variables, with lower scores observed in males, children born preterm, children of mothers with lower education and children from lower-income families. Fine motor scores were lower in males and in children born preterm. Children born preterm or from lower-income families had significantly

lower receptive language scores. Preterm birth was a significant predictor of lower expressive language scores. Further, we tested for an interaction effect of predictor variables with age on cognitive scores. The interaction between age and preterm birth was significant for visual reception and receptive language scores. For visual reception, children born full term and preterm were similar until around 1.5 years

Table 3 | Demographics, birth outcomes and SES significantly influence cognitive and motor development^a

	Gross motor	Visual reception	Fine motor	Receptive language	Expressive language
	Estimate (s.e.)	Estimate (s.e.)	Estimate (s.e.)	Estimate (s.e.)	Estimate (s.e.)
Reference growth model					
Intercept	7.65 (0.18)	6.13 (0.2)	7.5 (0.15)	5.76 (0.2)	5.52 (0.2)
Age	0.03** (<0.01)	0.03** (<0.01)	0.03** (<0.01)	0.03** (<0.01)	0.03** (<0.01)
Main effects					
Male sex	-0.19 (0.12)	-0.51** (0.13)	-0.3** (0.1)	-0.4* (0.15)	-0.4* (0.15)
Preterm birth	-0.92** (0.19)	-0.88** (0.22)	-0.91** (0.16)	-1.32** (0.25)	-1.28** (0.25)
Low birthweight	-0.26 (0.2)	-0.64* (0.23)	-0.52* (0.17)	-0.06 (0.26)	-0.18 (0.26)
Low maternal education	0.16 (0.2)	-0.74** (0.22)	-0.34* (0.16)	-0.46 (0.26)	-0.27 (0.26)
Low family income	0.01 (0.16)	-0.6** (0.18)	-0.11 (0.13)	-0.8** (0.21)	-0.53* (0.21)
Number of observations	2,384	2,404	2,410	2,472	2,477
Number of individuals	1,209	1,203	1,205	1,222	1,223
Interindividual variation	21.5%	10.37%	6.07%	13.12%	19.05%

^a $P < 0.05$; ** $P < 0.002$. ^aData were analyzed by two-tailed *t*-test with a Bonferroni post hoc test for *P* value significance. Exact *P* values are provided in Extended Data Table 2.

of age; subsequently, there was an increasing gap between the two, with children born full term scoring higher. For receptive language, children born preterm scored lower in early life but overtook children born full term around 3.5 years of age. The interaction of maternal education and family income with age was significant for visual reception scores. In both cases, a widening gap between children from lower-SES families and children from higher SES emerged around 1.5 years of age. For gross motor scores, interaction between family income and age was significant. Children from lower-income families had higher scores in early infancy than children from higher-income families and lower scores in late toddlerhood and early childhood (Supplementary Table 5 and Supplementary Figs. 7–9). To test if family relatedness affected our results, we performed a sensitivity analysis removing a twin/sibling from the data and reanalyzed the data using the main model. The results were like those of the main analysis (Supplementary Table 6). Split-sample replication analysis revealed non-significant associations but with similar direction of effect for all significant associations in the main analysis (Supplementary Table 7).

Correlations between brain volumes and cognitive scores

To assess brain–cognition correlations, we used Pearson's correlations between predicted brain volumes (ICV and subcortical structures) and cognitive scores at 2 years of age. In the full sample, volumes and cognitive scores showed a slight overall positive correlation. But the direction of effects varied across individual cohorts. Individual cohorts showing positive correlations had a greater representation of children born preterm, whereas those showing negative correlations included few or no children born preterm (Supplementary Table 8). Correlations were then analyzed separately for two subgroups, full-term children and preterm children. For full-term children (Fig. 2a), several brain volumes were significantly correlated with cognitive scores (Supplementary Table 9). ICV was negatively correlated with gross motor and fine motor scores. Amygdala volume was negatively correlated with gross motor

scores. Globus pallidus and caudate volumes were positively correlated with visual reception scores. Results were similar when excluding one twin/sibling from each pair (Supplementary Table 10). Split-sample replication analyses revealed that the direction of effect was highly consistent for these associations but seldom met the criteria for significance in both folds (Supplementary Table 11).

For children born preterm (Fig. 2b), most brain volumes were correlated with cognitive scores, with larger volumes associated with higher scores (Supplementary Table 12). However, the only significant relationship was between hippocampal volume and visual reception. This relationship did not meet the significance threshold when excluding one twin/sibling from each pair (Supplementary Table 13). Direction of effect was highly similar across folds in the split-sample analysis but was never significant in both folds (Supplementary Table 14).

To test whether correlations differed between children born full term and those born preterm more than would be expected from sample variability, we computed confidence intervals for the difference in correlation coefficients using a bootstrap method (Supplementary Table 15). Significant differences in correlation coefficients between full-term and preterm children were observed for gross motor scores with thalamus and pallidum volume, fine motor scores with thalamus, caudate, hippocampus, amygdala and ICV and visual reception and expressive language scores with ICV.

To explore if the relationships between sociodemographic factors and cognitive and motor development were mediated by brain volumes, we tested the mediation effect on significant brain–cognition correlations using predicted values at age 2 (Supplementary Table 16). In full-term children, the mediation effects by caudate and pallidum volumes on the influence of sex, family income and maternal education on visual reception scores were found to be significant (Fig. 3). In preterm children, the mediation effects of hippocampal volume on the relationship between sociodemographic factors (sex and maternal education) and visual reception scores were found to be significant

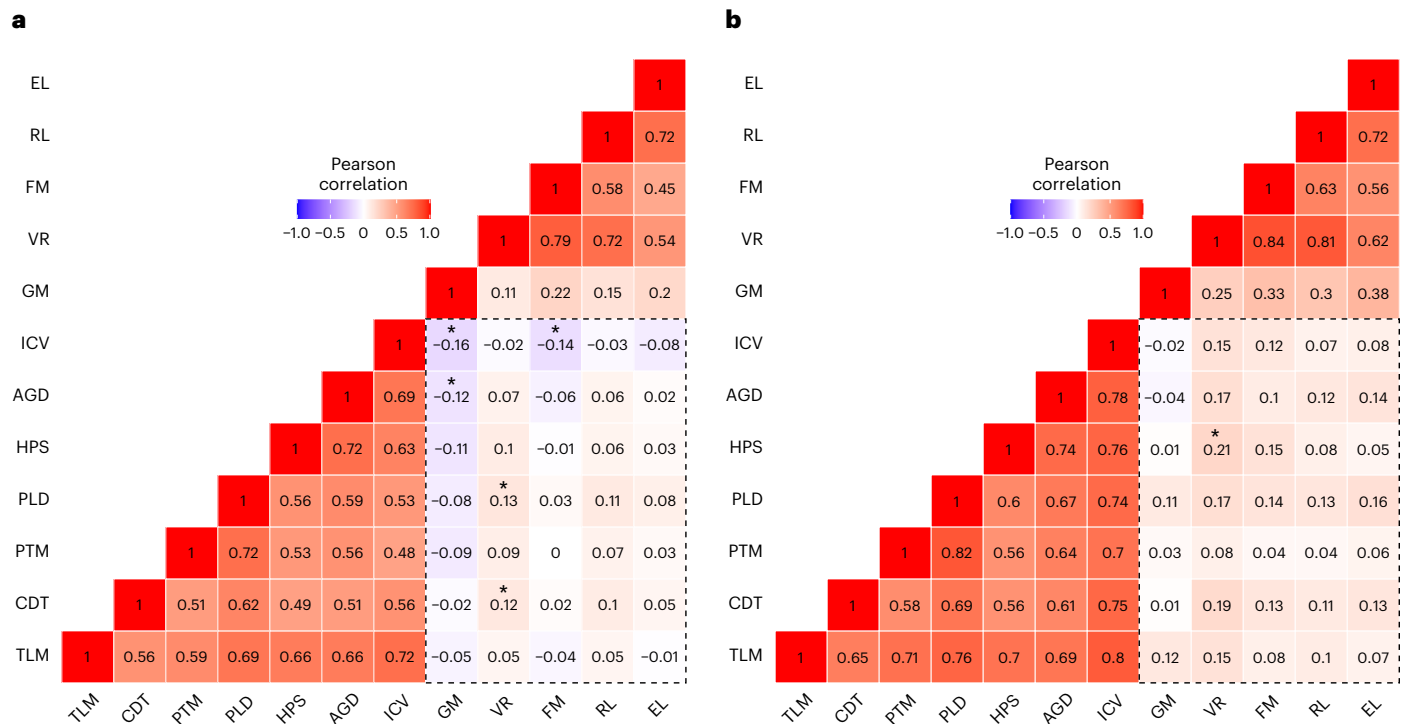


Fig. 2 | Heat map of the correlation between brain volumes and cognitive scores. a, b, Correlation assessed between the predicted values for brain volumes and cognitive and motor scores at age 2 using Pearson’s correlations in children born full term (a) and preterm (b). In full-term children, the significant correlations are ICV–gross motor score ($P = 0.0000$), amygdala–gross motor score ($P = 0.001$), ICV–fine motor score ($P = 0.0002$), caudate–visual reception score ($P = 0.001$) and globus pallidus–visual reception score ($P = 0.0005$). In preterm children, the significant correlation is hippocampus–visual reception

score ($P = 0.0007$). Brain volume–cognitive score correlations are highlighted in the dashed rectangle. The squares marked with an asterisk (*) represent significant correlations ($P < 0.0015$, which is the threshold for significance after Bonferroni correction for multiple comparisons). Data were analyzed by two-tailed t -tests; TLM, thalamus; CDT, caudate; PTM, putamen; PLD, globus pallidus; HPS, hippocampus; AGD, amygdala; GM, gross motor score; VR, visual reception score; FM, fine motor score; RL, receptive language score; EL, expressive language score.

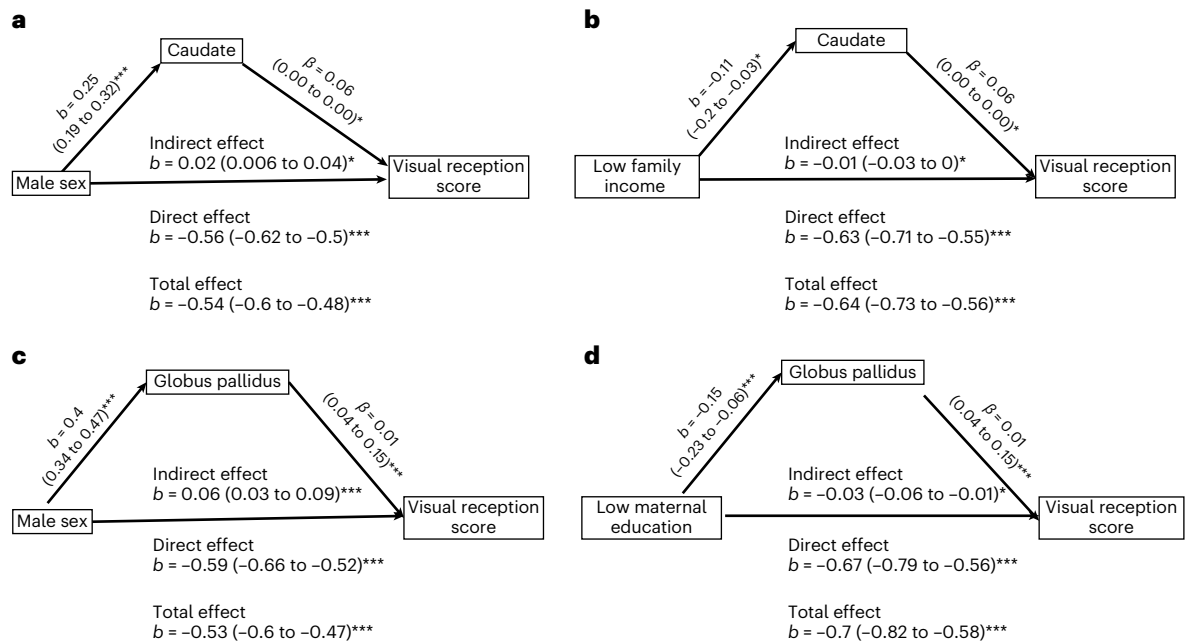


Fig. 3 | Causal mediation analyses for children born full term. a–d, Causal mediation analysis reveals partial mediation effects of the caudate and globus pallidus on the association between sociodemographic factors (sex, maternal education and family income) and visual reception scores in children born full term. a, b, Mediation by the caudate on the effect of sex (a; $P = 0.0052$) and low

family income (b; $P = 0.015$) on visual reception scores ($N = 760$). c, d, Mediation by the globus pallidus on the effect of sex (c; $P = 0.0004$) and low maternal education (d; $P = 0.0012$) on visual reception scores ($N = 659$). Bootstrapping was used to estimate P values; * $P < 0.05$; *** $P < 0.001$. No correction for multiple comparisons was performed.

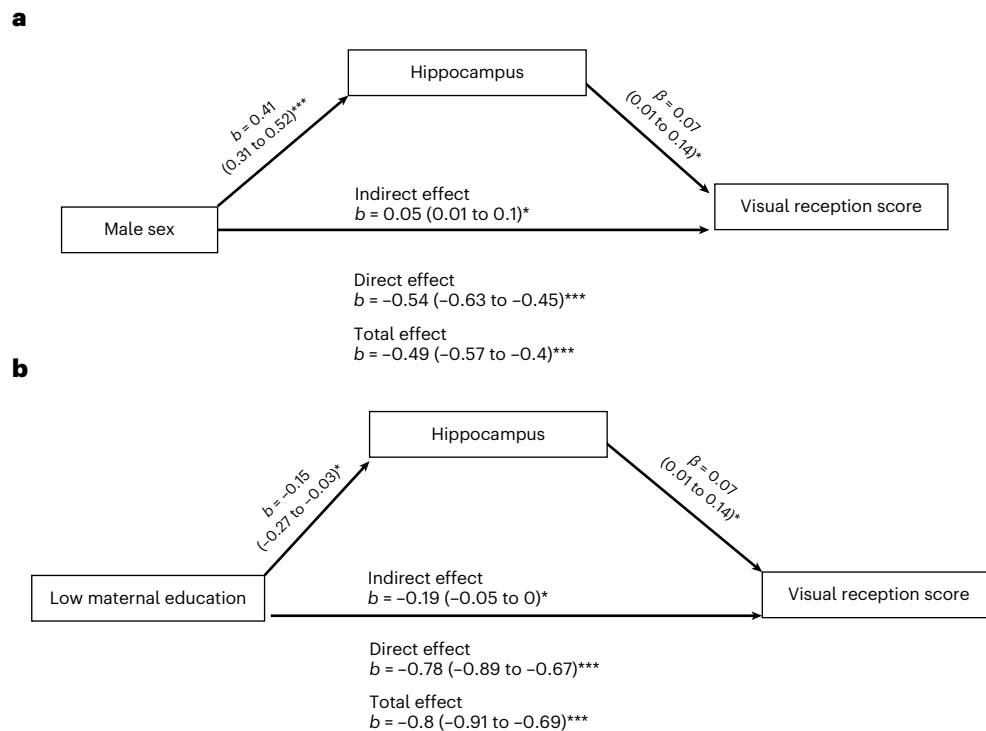


Fig. 4 | Causal mediation analyses for children born preterm. a, b, Causal mediation analysis reveals the partial mediation effects of hippocampal volume on the association between sociodemographic factors (sex and maternal education) and visual reception scores in children born preterm. Mediation by

the hippocampus on the effect of sex (**a**; $P = 0.01$) and low maternal education (**b**; $P = 0.041$) on visual reception scores ($N = 269$). Bootstrapping was used to estimate P values; * $P < 0.05$; *** $P < 0.001$. No correction for multiple comparisons was performed.

(Fig. 4). However, the magnitude of mediation effects (indirect effect) was smaller than that of direct effects. Because the effect sizes were so small, we did not conduct split-sample analyses for the mediation models.

Discussion

Using one of the largest pediatric neuroimaging datasets to date focused on birth to 6 years of age, we addressed three critical questions: (1) What is the trajectory of ICV and subcortical volume development? (2) How do sociodemographic factors and birth outcomes shape neurodevelopmental trajectories? (3) What are the neural correlates of cognitive development? ICV and subcortical structures followed nonlinear growth patterns in early childhood with considerable regional heterogeneity. Developmental trajectories were significantly associated with sex, adverse birth outcomes (preterm birth and low birthweight) and SES (maternal education and family income), and predicted ICV and volumes of the amygdala, hippocampus, caudate and pallidum correlate with different cognitive measures at 2 years of age in children born full term and preterm. Furthermore, associations between cognitive development and sociodemographic factors were partially mediated by subcortical brain volumes with small, but significant, effects.

Regarding the trajectory of ICV and subcortical volumes, we observed a faster growth rate in the first 1,000 postnatal days (~3 years), followed by a slower growth rate for all volumes. Extremely rapid growth in the first few years of life may make the brains of infants and young children especially vulnerable to environmental insults, such as poverty and preterm birth, and also especially responsive to interventions. As in prior studies of regional brain volumes performed from childhood to adolescence⁵, trajectories of ICV and subcortical structures followed a nonlinear pattern across early childhood, with different structures attaining maturation at different ages. This

finding suggests that different regions will have different windows of vulnerability/opportunity.

Among the subcortical structures, maturation age for the amygdala was around 3.3 years. Early structural maturation of the amygdala is consistent with other studies reporting early structural and functional development of the amygdala in primates including humans and suggesting late infancy to childhood as a sensitive period for its development²². The amygdala serves important roles in sensing the environment to evaluate potential dangers, process fear and mount an appropriate response²². Given the importance of threat stimuli for survival, natural selection may have promoted early development of neural circuits for processing and responding to threats. Hippocampal volume matured around 4.5 years of age. The hippocampus is involved in many skills, including memory, language and spatial cognition²³. The improvement of spatial performance observed in children aged 4–5 years could reflect hippocampal maturation²⁴. This period also coincides with the offset of infantile amnesia, which has been linked to an immature hippocampus²³. Prior studies differed on age at which the volumes of different basal ganglia structures peak, from late childhood to adolescence^{25,26}. Our results show that basal ganglia volumes (caudate, putamen and pallidum) asymptote around 5 years of age. In our study, thalamus volume was predicted to reach maturation around 7.5 years. The caveat of this observation is that we do not have any observations at that period. Reports on normative developmental trajectory of the thalamus are inconsistent, with peak age varying from 4 to 13 years (refs. 25,27). Inconsistencies between our study and previous reports could be due to differences in study design, sample size and population demographics.

Regarding the effects of demographic factors and birth outcomes, sex, adverse birth outcomes and SES influenced ICV and subcortical volumes. ICV was significantly larger in males than in females throughout the age range. This result is consistent with existing literature on sex

differences in neuroanatomy, which finds that male brains are larger, on average, than female brains throughout the lifespan¹². Prior reports on sex differences in subcortical volumes in late childhood through adulthood are inconsistent¹², perhaps due to differences in study design and cohort age. In our study, all subcortical volumes were larger in males than in females from birth to age 6. This association was robust and remained significant in the replication analysis. However, males had relatively smaller caudate volumes than females when adjusting for ICV. The relative lack of sex differences when adjusting for ICV is unsurprising given collinearity between ICV and individual subcortical volumes and is consistent with prior research (see Methods). The relatively smaller caudate volume in males could relate to sex differences in striatal-associated functions, including motivated behaviors and responsiveness to drugs of abuse, but ought to be considered with caution as the proportions method we used can result in misassignment of structures as larger or smaller than their actual size²⁸. We also found that sex plays a significant role in cognitive development, with males having lower visual reception and fine motor scores than females across the age range. There are many factors that could potentially influence sex differences in brain development, such as prenatal and postnatal exposure to gonadal steroids, sex chromosome genes and environmental effects²⁸.

Preterm birth and low birthweight were negatively associated with all volumes, but only at intercept. This possibly reflects delayed maturation of regions in early infancy with subsequent catching up in later development. Several studies in children and adults born very preterm with very low birthweight report persistent lower volumes in subcortical structures¹⁴. However, most children born preterm fall into the category of moderate (32–34 weeks) to late (34–37 weeks) preterm. Worldwide, of all children born preterm, around 84.7% are moderate to late preterm²⁹. This group is underrepresented in most studies, and there is a paucity of information on their developmental outcomes. The current study is more representative of these moderate-late preterm children. Associations with ICV, amygdala and hippocampus are robust as evidenced in the replication analysis. During cognitive development, we see significant negative associations with gestational age at birth on all developmental domains tested in MSEL. For visual reception, deficits became more pronounced with age. By contrast, receptive language developed more quickly in preterm children, mirroring our brain volume findings. In general, our results are in keeping with reported deficits in cognitive, language and motor skills in children born preterm¹⁴.

Maternal education and family income exhibited different patterns of association with our developmental imaging phenotypes, suggesting that parental education and family income represent distinct resources that influence children's environments and development in different ways³⁰. Maternal education could influence the child's development through both nature and nurture (genes and environment). Higher maternal education is correlated with better health service use, quality of childcare system and mother–child interactions³¹. Maternal education has positive effects on general cognitive ability³² and verbal and nonverbal functioning in children^{32,33}, patterns also observed in the current study in relation to visual reception and fine motor function.

ICV was significantly lower in children with mothers with primary or secondary education than in children with mothers with tertiary education across the age range. Educational attainment³⁴ and ICV³⁵ are heritable traits with substantial genetic overlap reported from large-scale genome-wide association studies³⁴. These overlapping genetic factors may partially explain the relationship observed here.

At the asymptote, lower maternal education was a predictor for smaller amygdala volume. The amygdala has a critical role in processing threats and learning environmental cues²². Early-life stress and risk exposure can impact amygdala volume and function^{36–38}. Previous studies, all with sample sizes of less than 1,000, have shown inconsistencies in the direction of associations between parental education and amygdala volume in children³⁹. We observe a positive relationship

between maternal education and amygdala volume. Our observation is in line with results from a large cohort study of 9,380 children between 9 and 10 years of age⁴⁰. The discrepancies between studies with different sample sizes underscores the need for large datasets in neuroimaging studies to validate associations between imaging phenotypes and different biological and environmental factors. In the sensitivity analysis with only maternal education as the socioeconomic variable, we found lower maternal education to be significantly associated with smaller hippocampal volume. This is consistent with prior reports suggesting a positive relationship between hippocampal volume and parental education^{41,42}. Maternal education positively predicted thalamic volume at the asymptote. This finding is consistent with a previous study showing that higher childhood SES, as measured by parental education/occupation, was associated with larger thalamus volume⁴³. Anomalous development of the thalamus can significantly affect development of other cortical and subcortical brain structures and can impact cognitive outcomes⁴⁴. At the asymptote, caudate, putamen and pallidum volumes were also significantly lower with lower maternal education, similar to findings observed in previous studies³⁹.

Family income can impact many aspects of a child's environment. Children from lower-income families are more likely to experience stressful environments and possess fewer material and non-material resources, whereas children from higher-income families are reported to have better environmental stimulation that promotes brain development¹⁵.

Children from lower-income families had significantly smaller hippocampal volumes across the age range. This is consistent with previous reports of associations between SES and hippocampal volume where income was used as a measure of SES^{15,39}. The hippocampus mediates long-term memory functions and responses to environmental stress⁴⁵, and larger hippocampal volume is related to better memory performance. Children from high-SES backgrounds have more exposure to stimulating environments that promote learning, including certain educational activities, than children from low-SES backgrounds, and this could improve memory functioning in childhood and boost hippocampal volume⁴⁶.

At the asymptote, thalamic volume was significantly smaller in children from lower-income families, consistent with previous reports³⁹. Basal ganglia structures were also smaller, which is similar to that observed in previous studies^{47,48}. Early-life stress is known to affect development of the caudate and putamen⁴⁹. Thalamus and basal ganglia networks modulate behavioral responses and regulate cortical neurons. Smaller volume in thalamic and basal ganglia regions in children from lower-SES backgrounds might lead to difficulty in coordinating behavioral responses to stress and reward stimuli⁴⁷. In our study, family income was not a significant predictor of ICV. Previous studies on the effects of SES on ICV yielded conflicting observations, with one study reporting a positive correlation with income–poverty ratio⁴⁷ and one study reporting no association between SES and ICV⁵⁰. Similarly, family income was not significantly associated with amygdala volume, and this observation is consistent with previous studies in individuals from late childhood to adulthood⁴⁵. Lower family income predicted lower scores on visual reception and receptive and expressive language scales of MSEL. This is consistent with prior research linking disparities in SES to several cognitive domains, including overall cognitive development, memory, language acquisition, executive control and school achievement¹⁵. Effects became more pronounced at later ages, although the only subscale to show a significant interaction effect with age was visual reception.

Prior studies have reported an impact of socioeconomic differences on cognitive and brain development as early as the first year¹⁵. Our observations provide evidence of differential effects of SES factors on brain development as early as birth. However, it must be noted that split-sample replication analyses suggest that brain volume associations with SES are less robust than associations with sex and adverse

birth outcomes. The subtle effect of these factors on brain structure may be due to genetic, environmental and gene \times environment interaction effects on neurogenesis, synaptogenesis and neuronal morphology¹¹.

Regarding question 3 ('what are the neural correlates of cognitive development in this period?'), in children born full term, globus pallidus and caudate volumes were positively correlated with visual reception scores. These associations make sense as the primary abilities assessed by visual reception scale are visual discrimination, memory, organization, sequencing and spatial awareness⁵¹. Basal ganglia structures are known to interact with the prefrontal cortex to support working memory. Furthermore, the caudate and globus pallidus facilitate visual motor integration via corticostriatal loops⁵². In children born preterm, hippocampal volume was positively correlated with visual reception scores, although sensitivity and split-sample replication analyses suggest that this effect may not be robust. The hippocampus plays a significant role in cognition and memory, particularly sequencing^{23,53}, and lower hippocampal volume has been linked to worse cognitive outcomes in children born preterm¹⁴. ICV was negatively correlated with gross and fine motor scores in children born full term. This relationship may be due to changes in gray matter and white matter composition, myelination or cerebral spinal fluid in the subarachnoid space^{54–56}.

Mediation analyses suggest that the effects of sociodemographic factors on cognitive scores are partially mediated by brain volumes with significant but small effect sizes. The effects of low maternal education and low family income on visual reception scores are partially mediated by lower caudate and pallidum volumes, respectively. Similar to what we observe, the pallidum has been found to partially mediate the relationship between diminished growth and intelligence quotient in children from low-SES conditions⁵⁷. In addition, the association between male sex and lower visual reception is partially ameliorated by caudate and globus pallidus volumes. In children born preterm, the effect of male sex on visual reception score is ameliorated by larger hippocampal volume, and the effect of lower maternal education is partially mediated by lower hippocampal volume.

A particular strength of our study is the large ethnically and socially diverse sample of children who have undergone cognitive assessments and structural MRI measurements in the critical period of infancy and early childhood. Because our study is well powered and we apply stringent corrections for multiple comparisons, results are expected to be rigorous and robust; however, there are some limitations. First, we used maternal education as a binary variable (tertiary versus primary/secondary education). However, in the South African cohort, most of the mothers had some secondary education, and very few had tertiary education. This is an accurate reflection of the country, where 79% of women have an upper secondary education as the highest level achieved, but only 6% of women have a tertiary education⁵⁸. The binary classification might underrepresent some characteristics of that population. Second, as with other multisite observational studies, uncontrolled confounding may influence results. For putamen and pallidum volumes at birth, the group with low maternal education is predominantly from the South African cohort, which is mainly composed of people of African descent and mixed ancestry. They have larger volumes in the pallidum and putamen than people of European descent in this study in the same age range. Hence, the observed effect of maternal education on the putamen and pallidum at birth could reflect differences in segmentation protocols, health history of parents and ancestry within that cohort. Third, although we have an adequate number of observations across the age range to draw representative inferences, we also note that the number of observations is higher from birth to age 3 years. The current results along with others strongly point to the need of inclusion of different ethnic populations from diverse socioeconomic strata in neuroimaging studies. The ORIGINS consortium is a step toward achieving this goal.

In conclusion, our study contributes to the better understanding of the effects of sex, adverse birth outcomes and SES on neurodevelopment and cognitive outcomes in socially and ethnically diverse cohorts. Our approach could be expanded to address other environmental and contextual factors that shape early brain development and, in the long term, inform public health policy and interventions. Furthermore, by defining trajectories of ICV and subcortical volume development in infancy and early childhood, we lay the foundation for large-scale imaging–genetics studies of this critical developmental period.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41593-023-01501-6>.

References

1. Stiles, J. & Jernigan, T. L. The basics of brain development. *Neuropsychol. Rev.* **20**, 327–348 (2010).
2. Feldman, R. & Eidelman, A. I. Biological and environmental initial conditions shape the trajectories of cognitive and social–emotional development across the first years of life. *Dev. Sci.* **12**, 194–200 (2009).
3. Gao, W. et al. A review on neuroimaging studies of genetic and environmental influences on early brain development. *NeuroImage* **185**, 802–812 (2019).
4. Gilmore, J. H., Knickmeyer, R. C. & Gao, W. Imaging structural and functional brain development in early childhood. *Nat. Rev. Neurosci.* **19**, 123–137 (2018).
5. Vijayakumar, N., Mills, K. L., Alexander-Bloch, A., Tamnes, C. K. & Whittle, S. Structural brain development: a review of methodological approaches and best practices. *Dev. Cogn. Neurosci.* **33**, 129–148 (2018).
6. Kraemer, H. C., Yesavage, J. A., Taylor, J. L. & Kupfer, D. How can we learn about developmental processes from cross-sectional studies, or can we? *Am. J. Psychiatry* **157**, 163–171 (2000).
7. Marek, S. et al. Reproducible brain-wide association studies require thousands of individuals. *Nature* **603**, 654–660 (2022).
8. Klapwijk, E. T., van den Bos, W., Tamnes, C. K., Raschle, N. M. & Mills, K. L. Opportunities for increased reproducibility and replicability of developmental neuroimaging. *Dev. Cogn. Neurosci.* **47**, 100902 (2021).
9. Schmaal, L., Ching, C. R. K., McMahon, A. B., Jahanshad, N. & Thompson, P. M. in *Personalized Psychiatry* (ed Baune, B. T.) 483–497 (Elsevier, 2019).
10. Bethlehem, R. A. I. et al. Brain charts for the human lifespan. *Nature* **604**, 525–533 (2022).
11. Olson, L., Chen, B. & Fishman, I. Neural correlates of socioeconomic status in early childhood: a systematic review of the literature. *Child Neuropsychol.* **27**, 390–423 (2021).
12. Kaczurkin, A. N., Raznahan, A. & Satterthwaite, T. D. Sex differences in the developing brain: insights from multimodal neuroimaging. *Neuropsychopharmacology* **44**, 71–85 (2019).
13. Gurvich, C., Thomas, N. & Kulkarni, J. Sex differences in cognition and aging and the influence of sex hormones. *Handb. Clin. Neurol.* **175**, 103–115 (2020).
14. Lean, R. E., Neil, J. J. & Smyser, C. D. in *Handbook of Pediatric Brain Imaging* Vol. 2 (eds Huang, H. & Roberts, T. P. L.) 429–465 (Academic Press, 2021).
15. Farah, M. J. The neuroscience of socioeconomic status: correlates, causes, and consequences. *Neuron* **96**, 56–71 (2017).

16. Shonkoff, J. P., Boyce, W. T. & McEwen, B. S. Neuroscience, molecular biology, and the childhood roots of health disparities: building a new framework for health promotion and disease prevention. *JAMA* **301**, 2252–2259 (2009).
17. Koziol, L. F., Barker, L. A., Joyce, A. W. & Hrin, S. The small-world organization of large-scale brain systems and relationships with subcortical structures. *3*, 245–252 (2014).
18. Pietschnig, J., Penke, L., Wicherts, J. M., Zeiler, M. & Voracek, M. Meta-analysis of associations between human brain volume and intelligence differences: how strong are they and what do they mean? *Neurosci. Biobehav. Rev.* **57**, 411–432 (2015).
19. Girault, J. B. et al. The predictive value of developmental assessments at 1 and 2 for intelligence quotients at 6. *Intelligence* **68**, 58–65 (2018).
20. Thompson, P. M. et al. ENIGMA and global neuroscience: a decade of large-scale studies of the brain in health and disease across more than 40 countries. *Transl. Psychiatry* **10**, 100 (2020).
21. Whitwell, J. L., Crum, W. R., Watt, H. C. & Fox, N. C. Normalization of cerebral volumes by use of intracranial volume: implications for longitudinal quantitative MR imaging. *AJNR Am. J. Neuroradiol.* **22**, 1483–1489 (2001).
22. Tottenham, N. & Gabard-Durnam, L. J. The developing amygdala: a student of the world and a teacher of the cortex. *Curr. Opin. Psychol.* **17**, 55–60 (2017).
23. Lee, J. K., Johnson, E. G. & Ghetti, S. in *The Hippocampus from Cells to Systems: Structure, Connectivity, and Functional Contributions to Memory and Flexible Cognition* (eds Hannula, D. E. & Duff, M. C.) 141–166 (Springer International Publishing, 2017).
24. Sutton, J. E., Joannisse, M. F. & Newcombe, N. S. Spinning in the scanner: neural correlates of virtual reorientation. *J. Exp. Psychol. Learn Mem. Cogn.* **36**, 1097–1107 (2010).
25. Wierenga, L. et al. Typical development of basal ganglia, hippocampus, amygdala and cerebellum from age 7 to 24. *NeuroImage* **96**, 67–72 (2014).
26. Lenroot, R. K. et al. Sexual dimorphism of brain developmental trajectories during childhood and adolescence. *NeuroImage* **36**, 1065–1073 (2007).
27. Tutunji, R. et al. Thalamic volume and dimensions on MRI in the pediatric population: normative values and correlations: (a cross sectional study). *Eur. J. Radiol.* **109**, 27–32 (2018).
28. Panzica, G. C. & Melcangi, R. C. Structural and molecular brain sexual differences: a tool to understand sex differences in health and disease. *Neurosci. Biobehav. Rev.* **67**, 2–8 (2016).
29. Chawanpaiboon, S. et al. Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis. *Lancet Glob. Health* **7**, e37–e46 (2019).
30. Duncan, G. J. & Magnuson, K. Socioeconomic status and cognitive functioning: moving from correlation to causation. *Wiley Interdiscip. Rev. Cogn. Sci.* **3**, 377–386 (2012).
31. Augustine, J. M., Cavanagh, S. E. & Crosnoe, R. Maternal education, early child care and the reproduction of advantage. *Soc. Forces* **88**, 1–29 (2009).
32. Girault, J. B. et al. White matter microstructural development and cognitive ability in the first 2 years of life. *Hum. Brain Mapp.* **40**, 1195–1210 (2019).
33. Dai, X., Hadjipantelis, P., Wang, J., Deoni, S. C. L. & Müller, H. Longitudinal associations between white matter maturation and cognitive development across early childhood. *Hum. Brain Mapp.* **40**, 4130–4145 (2019).
34. Okbay, A. et al. Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* **533**, 539–542 (2016).
35. Adams, H. H. H. et al. Novel genetic loci underlying human intracranial volume identified through genome-wide association. *Nat. Neurosci.* **19**, 1569–1582 (2016).
36. Evans, G. W. et al. Childhood cumulative risk exposure and adult amygdala volume and function. *J. Neurosci. Res.* **94**, 535–543 (2016).
37. Graham, A. M., Pfeifer, J. H., Fisher, P. A., Carpenter, S. & Fair, D. A. Early life stress is associated with default system integrity and emotionality during infancy. *J. Child Psychol. Psychiatry* **56**, 1212–1222 (2015).
38. Turesky, T. K. et al. The relationship between biological and psychosocial risk factors and resting-state functional connectivity in 2-month-old Bangladeshi infants: a feasibility and pilot study. *Dev. Sci.* **22**, e12841 (2019).
39. Noble, K. G. & Giebler, M. A. The neuroscience of socioeconomic inequality. *Curr. Opin. Behav. Sci.* **36**, 23–28 (2020).
40. Assari, S., Boyce, S. & Bazargan, M. Subjective socioeconomic status and children’s amygdala volume: minorities’ diminish returns. *NeuroSci* **1**, 59–74 (2020).
41. Noble, K. G. et al. Family income, parental education and brain structure in children and adolescents. *Nat. Neurosci.* **18**, 773–778 (2015).
42. Ellwood-Lowe, M. E. et al. Time-varying effects of income on hippocampal volume trajectories in adolescent girls. *Dev. Cogn. Neurosci.* **30**, 41–50 (2018).
43. McDermott, C. L. et al. Longitudinally mapping childhood socioeconomic status associations with cortical and subcortical morphology. *J. Neurosci.* **39**, 1365–1373 (2019).
44. Herrero, M. T., Barcia, C. & Navarro, J. M. Functional anatomy of thalamus and basal ganglia. *Childs Nerv. Syst.* **18**, 386–404 (2002).
45. Hanson, J. L., Chandra, A., Wolfe, B. L. & Pollak, S. D. Association between income and the hippocampus. *PLoS ONE* **6**, e18712 (2011).
46. Zarif, H., Nicolas, S., Petit-Paitel, A., Chabry, J. & Guyon, Alice. in *The Hippocampus—Plasticity and Functions* (ed Stuchlik A.) Ch. 1 (IntechOpen, 2017).
47. Jenkins, L. M. et al. Subcortical structural variations associated with low socioeconomic status in adolescents. *Hum. Brain Mapp.* **41**, 162–171 (2020).
48. Leonard, J. A., Mackey, A. P., Finn, A. S. & Gabrieli, J. D. E. Differential effects of socioeconomic status on working and procedural memory systems. *Front. Hum. Neurosci.* **9**, 554 (2015).
49. Fareri, D. S. & Tottenham, N. Effects of early life stress on amygdala and striatal development. *Dev. Cogn. Neurosci.* **19**, 233–247 (2016).
50. Knickmeyer, R. C. et al. Impact of demographic and obstetric factors on infant brain volumes: a population neuroscience study. *Cereb. Cortex* **27**, 5616–5625 (2016).
51. Mullen, E. *Mullen Scales of Early Learning* (AGS Publishing, 1995).
52. Seger, C. A. How do the basal ganglia contribute to categorization? Their role in generalization, response selection, and learning via feedback. *Neurosci. Biobehav. Rev.* **32**, 265–278 (2008).
53. Lisman, J. E. & Otmakhova, N. A. Storage, recall, and novelty detection of sequences by the hippocampus: elaborating on the SOCRATIC model to account for normal and aberrant effects of dopamine. *Hippocampus* **11**, 551–568 (2001).
54. Groeschel, S., Vollmer, B., King, M. D. & Connelly, A. Developmental changes in cerebral grey and white matter volume from infancy to adulthood. *Int. J. Dev. Neurosci.* **28**, 481–489 (2010).
55. Shen, M. D. Cerebrospinal fluid and the early brain development of autism. *J. Neurodev. Disord.* **10**, 39 (2018).
56. Buyanova, I. S. & Arsalidou, M. Cerebral white matter myelination and relations to age, gender, and cognition: a selective review. *Front. Hum. Neurosci.* **15**, 662031 (2021).
57. Turesky, T. K. et al. Brain morphometry and diminished physical growth in Bangladeshi children growing up in extreme poverty: a longitudinal study. *Dev. Cogn. Neurosci.* **52**, 101029 (2021).
58. OECD. Education at a glance 2019: OECD indicators. <https://doi.org/10.1787/f8d7880d-en> (OECD Publishing, 2019).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this

article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

¹Institute for Quantitative Health Sciences and Engineering, Michigan State University, East Lansing, MI, USA. ²Departments of Epidemiology & Biostatistics, Michigan State University, East Lansing, MI, USA. ³Mallinckrodt Institute of Radiology, Washington University School of Medicine, St. Louis, MO, USA. ⁴Department of Medical Psychology, Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany. ⁵Department of Pediatrics, University of California Irvine, Irvine, CA, USA. ⁶Development, Health and Disease Research Program, University of California Irvine, Irvine, CA, USA. ⁷Department of Obstetrics and Gynaecology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore. ⁸Singapore Institute for Clinical Sciences, Agency for Science, Technology and Research, Singapore, Singapore. ⁹Department of Radiology, University of Washington Medical Center, Seattle, WA, USA. ¹⁰Division of Developmental Paediatrics, Department of Paediatrics and Child Health, Red Cross War Memorial Children's Hospital, University of Cape Town, Cape Town, South Africa. ¹¹Neuroscience Institute, University of Cape Town, Cape Town, South Africa. ¹²Masonic Institute for the Developing Brain, University of Minnesota Medical School, Minneapolis, MN, USA. ¹³Department of Diagnostic & Interventional Imaging, KK Women's and Children's Hospital, Singapore, Singapore. ¹⁴Harvard Graduate School of Education, Harvard University, Cambridge, MA, USA. ¹⁵Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. ¹⁶Carolina Institute for Developmental Disabilities, University of North Carolina at Chapel Hill, Carboro, NC, USA. ¹⁷Department of Psychiatry, Oregon Health & Science University, Portland, OR, USA. ¹⁸South African Medical Research Council (SA-MRC) Unit on Child & Adolescent Health, University of Cape Town, Cape Town, South Africa. ¹⁹Department of Psychiatry, University of Cape Town, Cape Town, South Africa. ²⁰Department of Paediatrics and Child Health, University of Cape Town, Faculty of Health Sciences, Cape Town, South Africa. ²¹Department of Radiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. ²²Department of Biomedical Engineering, National University of Singapore, Singapore, Singapore. ²³NUS (Suzhou) Research Institute, National University of Singapore, Suzhou, China. ²⁴The N.1 Institute for Health, National University of Singapore, Singapore, Singapore. ²⁵Institute of Data Science, National University of Singapore, Singapore, Singapore. ²⁶Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA. ²⁷Department of Health Technology and Informatics, Hong Kong Polytechnic University, Hung Hom, China. ²⁸SAMRC Unit on Risk and Resilience in Mental Disorders, Department of Psychiatry, University of Cape Town, Cape Town, South Africa. ²⁹Center for Autism Research, Children's Hospital of Philadelphia and the University of Pennsylvania, Philadelphia, PA, USA. ³⁰Research Group Learning in Early Childhood, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany. ³¹Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. ³²Imaging Genetics Center, Stevens Neuroimaging & Informatics Institute, Keck School of Medicine of University of Southern California, Marina del Rey, CA, USA. ³³Departments of Psychiatry and Human Behavior, Obstetrics & Gynecology, Epidemiology, University of California, Irvine, Irvine, CA, USA. ³⁴A.A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA, USA. ³⁵Department of Statistics & Probability, Michigan State University, East Lansing, MI, USA. ³⁶Department of Pediatrics and Human Development, Michigan State University, East Lansing, MI, USA. *A list of authors and their affiliations appears at the end of the paper.

✉ e-mail: knickmey@msu.edu

the ENIGMA ORIGINS group

Ann M. Alex¹, Claudia Buss^{4,5,6}, Yap-Seng Chong^{7,8}, Kirsten A. Donald^{10,11}, Damien A. Fair¹², Marielle V. Fortier^{8,13}, Nadine Gaab¹⁴, John H. Gilmore¹⁵, Jessica B. Girault^{15,16}, Nynke A. Groenewold^{10,18,19,20}, Weili Lin²¹, Michael J. Meaney⁹, Joseph Piven^{16,17}, Anqi Qiu^{22,23,24,25,26,27}, Jerod M. Rasmussen^{5,6}, Annerine Roos^{10,11,28}, Michael A. Skeide³⁰, Dan J. Stein^{19,28}, Martin Styner^{16,31}, Paul M. Thompson³², Ted K. Turesky¹⁴, Heather J. Zar^{18,20}, Lilla Zöllei³⁴, Gustavo de los Campos^{1,2,35} & Rebecca C. Knickmeyer^{1,36}

Methods

Participants

The data for this project were provided by the following members of ENIGMA-ORIGINS: Max Planck Institute for Human Cognitive and Brain Sciences (Germany), GUSTO (Singapore), the Drakenstein Child Health Study (South Africa), the BCP, Boston Children's Hospital/Harvard Medical School, the IBIS network, University of North Carolina EBDS and UCI. The final dataset includes children representing socially and ethnically diverse backgrounds (Table 1 and Supplementary Table 17). Each project was approved by their respective local review board, and informed consent was obtained from parents/legal guardians and children before data collection. The reviewing organizations include Michigan State University; Max Planck Institute for Human Cognitive and Brain Sciences, Germany; National University of Singapore, Singapore; University of Cape Town, South Africa; the University of North Carolina, Chapel Hill; the University of California, Irvine; and Boston's Children Hospital. For three cohorts (Germany, South Africa and UCI) MRI data were cross-sectional. The other five cohorts had longitudinal data. Overall, the imaging cohort included 2,108 children with a total of 3,607 observations and an age range of 5–2,250 postnatal days (Supplementary Fig. 1). No statistical methods were used to predetermine sample sizes, but our sample sizes are among the highest for pediatric imaging studies. As this was an observational study, blinding does not apply.

Image acquisition and analysis

Structural T1-weighted and T2-weighted scans of each participant were acquired and processed at each study site (Supplementary Tables 18–20). Images were acquired at different field strengths (1.5 T and 3 T). The reported sample size from each cohort is after quality control was performed locally at the respective sites.

Cohort characteristics

Max Planck Institute for Human Cognitive and Brain Sciences. The Max Planck Institute for Human Cognitive and Brain Sciences cohort includes children with and without a family risk of dyslexia who underwent MRI between 3 and 6 years of age. Exclusion criteria included impaired hearing and/or vision, an intelligence quotient below 80, psychiatric disorders, attention-deficit/hyperactivity disorder, previous neurosurgery, contraindication for MRI, medication that modulates brain function and inability and/or unwillingness to follow experimental instructions and/or perform experimental tasks. Participating families received travel cost reimbursement and a small gift.

Segmentation protocol. Scans used for segmentation were preselected for image quality by visual inspection for artifacts, signal dropouts, spatial distortion and anatomical anomalies. In the sample of 3- to 6-year-old children, segmentation was performed using the recon-all procedure implemented into FreeSurfer, which allowed for the extraction of gray matter images from the T1-weighted scans. First, 130 images were skull stripped. Second, white matter and gray matter boundaries were reconstructed. Third, boundary reconstructions were used to calculate the pial surface. These automatic processing steps could not be completed in three datasets, which were then discarded so that cortical surface reconstructions were available for 127 individuals. ICV was calculated using an atlas-based estimation approach implemented in FreeSurfer (<https://surfer.nmr.mgh.harvard.edu/fswiki/eTIV>). Specifically, they computed a volume-scaling factor derived by spatially transforming each individual image to an atlas image for which the ICV is already known. This volume-scaling factor renders a reliable estimation possible because it is highly correlated with the individual ICV. Quality of the automatic surface reconstruction results was assessed by thorough visual inspection. To ensure the neuroanatomical accuracy of each individual dataset, remaining parts of the skull were removed, and removed parts of the cortex were added again by adding control points and rerunning the surface reconstruction if necessary.

GUSTO. This study is comprised of a parent–offspring cohort. Exclusion criteria included mothers receiving chemotherapy or psychotropic drugs or type I diabetes mellitus. Participant compensation for the families was SGD \$100 per trip.

Segmentation protocol. For neonates, a Markov random field model was used to automatically segment the subcortical structures. In the Markov random field model, the prior probability of each structure was computed based on the manual segmentation of 20 participants randomly chosen from the participants with the manual labels. The prior probability atlas was obtained in the GUSTO neonatal atlas⁵⁹ where all T2-weighted images were nonlinearly transformed to using large deformation diffeomorphic metric image mapping⁶⁰. Accuracy of this automated segmentation was validated using leave-one-out validation in the manual segmented dataset. ICV was calculated as the number of voxels inside the brain after brain skull removal and scaled by the image resolution, including gray matter, white matter and cerebrospinal fluid of ventricles^{61–63}.

For older children, to eliminate potential profound effects of head motion on our statistical results, we manually checked image quality based on the stringent criteria in Ducharme et al.⁶⁴. Disqualified images were excluded from this study. FreeSurfer software (<https://ddec1-0-en-ctp.trendmicro.com:443/wis/clicktime/v1/query?url=http%3a%2f%2fsurfer.nmr.mgh.harvard.edu%2f&umid=eb095c3f8b31464688adef826b8ef738&auth=8d3ccd473d52f326e51c0f75cb32c9541898e5d5-e74d695f31fbfaabebae9b32a93056f6e20c8e6>) was then used to label each voxel in the usable T1-weighted image as gray matter, white matter, cerebrospinal fluid or subcortical structures. FreeSurfer used a Markov random field model that requests a prior probability obtained from a training dataset with T1-weighted images and their manual structural labels. We reconstructed the prior probability in the Markov random field model based on the manual segmentation of 30 children and embedded it in FreeSurfer. A post-processing quality check was conducted following the instructions in <https://ddec1-0-en-ctp.trendmicro.com:443/wis/clicktime/v1/query?url=https%3a%2f%2fsurfer.nmr.mgh.harvard.edu%2ffswiki%2fFsTutorial%2fTroubleshootingData&umid=eb095c3f8b31464688adef826b8ef738&auth=8d3ccd473d52f326e51c0f75cb32c9541898e5d55e1ff822a43b798d1a48a3fb6506b120f651b211>. Segmentation accuracy was assessed using a volume overlap ratio between the automated and manual segmentations.

Drakenstein Child Health Study, University of Cape Town. This longitudinal cohort aims to investigate the determinants of child growth, health and development in a stable, semiurban, low-socioeconomic-status community in South Africa. For the current study, children with scans acquired in the first month after birth are included. Exclusion criteria were minimal to maximize generalizability and were focused primarily on individuals who did not live in the region and thus could not be readily followed-up or who intended to move out of the district within the following 2 years. Participants were compensated for their time and travel expenses at each study visit with a voucher/gift card to the value of 350 ZAR (South African Rand). Refreshments were made available during the visit. Travel arrangements were offered to those participants who resided outside the study area.

Segmentation protocol. Sagittal three-dimensional T2-weighted images from 2- to 6-week-old infants were brain extracted with FSL v5.0. The output images were preprocessed further in Statistical Parametric Mapping software (SPM8) run in MATLAB R2017B. Images were registered and normalized with modulation to the University of North Carolina neonate T2 template⁶⁵. Hereafter, images were segmented into gray matter, white matter and cerebrospinal fluid based on the corresponding neonate probabilistic maps. Gray matter segmentations from

140 infants passed quality control through visual inspection (exclusion low image quality: 18 images; exclusion poor segmentation: 17 images). Gray matter volumes were extracted according to the automated anatomical labeling atlas⁶⁶, adapted for neonates⁶⁵, for the left and right amygdala, hippocampus, thalamus, caudate, putamen and pallidum.

University of North Carolina EBDS. This prospective longitudinal cohort includes children at high familial risk for schizophrenia and bipolar illness, a 'structural' high-risk group (children with prenatal isolated mild ventriculomegaly), a large sample of twins and an exceptionally large sample of typically developing infants. Exclusion criteria at enrollment included major medical illness in the mother, abnormality on ultrasound and current substance abuse. For participation in the study, parents received US \$50 per child for each MRI visit and US \$50 per child for each developmental assessment visit.

Segmentation protocol. For neonates, hippocampus and amygdala segmentation was performed using a multimodality, multitemplate-based automatic method combining T1- and T2-weighted high-resolution images in AutoSeg v3.3.2 (ref. 67) using the same multitemplate library as in the UCI cohort. Other subcortical structures were determined via a multimodality, single-template-based automatic method combining T1- and T2-weighted high-resolution images in AutoSeg v3.3.2 using the same single template as in the UCI cohort. For participants older than neonate age, all subcortical structure segmentation was performed using a multimodality, multitemplate-based automatic method combining T1- and T2-weighted high-resolution images in MultiSeg Pipeline v 2.2.1 using the same templates as in the IBIS cohort.

IBIS network. This longitudinal study aims to examine the early brain and behavioral development in infants at familial risk for autism and low-risk control infants (LR). For the current study, participants enrolled in the LR group were included. Exclusion criteria included the following: (1) diagnosis or physical signs strongly suggestive of a genetic condition or syndrome (for example, fragile X syndrome) reported to be associated with autism spectrum disorders, (2) a significant medical or neurological condition affecting growth, development or cognition (for example, CNS infection, seizure disorder and congenital heart disease), (3) sensory impairment, such as vision or hearing loss, (4) low birthweight (<2,000 g) or prematurity (<34 weeks gestation), (5) possible perinatal brain injury from exposure to in utero exogenous compounds reported to likely affect the brain adversely in at least some individuals (for example, alcohol and selected prescription medications), (6) non-English-speaking families, (7) contraindication for MRI (for example, metal implants), (8) individuals who were adopted and (9) a family history of intellectual disability, psychosis, schizophrenia or bipolar disorder in a first-degree relative. In addition, LR infants were excluded for autism spectrum disorder based on clinical evaluation at 24 and/or 36–60 months of age. All IBIS families were reimbursed for expenses incurred during study participation (for example, travel, lodging and meals). Families also received compensation for each of the longitudinal study visits, and children were offered small toys for participating.

Segmentation protocol. All subcortical structure segmentation was performed using a multimodality, multitemplate-based automatic method combining T1- and T2-weighted high-resolution images in AutoSeg v3.3.2 (ref. 67), followed by manual correction of selected datasets in ITK-Snap⁶⁸, if necessary. The multitemplate datasets consisted of 16 6-month-old datasets for the 6-month-old participant processing as well as 16 1-year-old and 16 2-year-old datasets for the 1- to 2-year-old participant processing.

UCI. This is a prospective, longitudinal, follow-up study in a population-based cohort. For the current study, infants with MRI

data in the first 2 months after birth were included. Exclusion criteria included (1) preterm birth <34 completed weeks gestation, (2) maternal use of psychotropic medication during pregnancy, (3) maternal use of corticosteroids during pregnancy, (4) maternal smoking and drug use during pregnancy (self-reports verified by urinary cotinine and drug toxicology), (5) congenital or genetic disorder (for example, fetal alcohol syndrome, Down syndrome and fragile X) and (6) major neurologic disorder at birth (for example, bacterial meningitis and epilepsy). Participant compensation was US \$100 per scan.

Segmentation protocol. Hippocampus and amygdala segmentation was performed using a multimodality, multitemplate-based automatic method combining T1- and T2-weighted high-resolution images in AutoSeg v3.3.2 (ref. 67), followed by manual correction of all datasets in ITK-Snap⁶⁸. Images were manually corrected in both original and left–right mirrored presentation to account for asymmetric presentation biases⁶⁹, and volumes were averaged for the two presentations. The multitemplate datasets consisted of eight neonate participants. Other subcortical structures were determined via a multimodality, single-template-based automatic method combining T1- and T2-weighted high-resolution images in AutoSeg v3.3.2. The single template was a single, unbiased average atlas computed from the ALBERT⁷⁰ datasets.

BCP. This sequential cohort with an accelerated longitudinal study design included typically developing children between birth and 5 years of age recruited across two data collection sites (University of North Carolina and The University of Minnesota). Exclusion criteria included gestational age of <37 weeks, birthweight of <2,500 g and any major pregnancy and/or delivery complications. Participation compensation was US \$150 Target gift card or Visa card (US \$135 + US \$15 for travel equivalent reimbursement) for each completed visit (scan and assessments). Participants who completed an MRI retry scan (without an assessment) were given US \$75 (Target or Visa card).

Segmentation protocol. All subcortical structure segmentations were performed using a multimodality, multitemplate-based automatic method combining T1- and T2-weighted high-resolution images in the MultiSeg Pipeline v 2.2.1 without manual correction. The multitemplate datasets consisted of 16 6-month-old datasets for participants younger than 9 months of age as well as 16 1-year- and 16 2-year-old datasets for participants older than 9 months of age using the same templates as in the IBIS cohort.

Boston Children's Hospital/Harvard Medical School. This is a prospective, longitudinal cohort aimed to study neural development in children with and without a familial history of developmental dyslexia. Exclusion criteria include psychiatric or neurological illness, sensory impairment, contraindications for MRI studies (for example, magnetic resonance-incompatible metal implants, such as surgical clips, and probability of metal fragments embedded in the body), treatment with psychotropic medication, prematurity and an atypical hearing screening. Each family received a US \$50 gift certificate for a local bookstore for their participation for each MRI session per participant (infant/child and parent). Families received an additional US \$25 per session for parking (US \$10), transportation costs (US \$10) and small toys/prizes (US \$5).

Segmentation protocol. All images were processed using (1) infant FreeSurfer⁷¹ for scans that were taken between 0 and 3 years of life and (2) a modified FreeSurfer pipeline adjusted for processing MRI data from children acquired at age 4.5 years or older. Infant FreeSurfer is an automated segmentation and surface extraction pipeline designed to accommodate clinical MRI studies of infant brains in a population

of 0- to 2-year-old children. The algorithm relies on a single channel of T1-weighted MRI images to achieve automated segmentation of cortical and subcortical brain areas, producing volumes of subcortical structures and surface models of the cerebral cortex. Infant FreeSurfer is equipped with niftyreg (<https://sourceforge.net/p/niftyreg>) for automated nonlinear registration between template and individual brains. The standard FreeSurfer pipeline⁷² has been optimized to perform well on adult acquisitions; however, it has been shown that with expert guidance and good-quality data, the tools can be used on images of participants as young as 4.5 years of age⁷³. ICV was calculated by the approach mentioned in infant FreeSurfer (see Max Planck Institute for Human Cognitive and Brain Sciences) but recomputed for infants using templates/images from the Developmental Human Connectome Project. After inspection of the segmentation and surface reconstruction outcomes, 85 of the youngest scans were processed with a variation of infant FreeSurfer. Instead of relying on the default multiatlas label-fusion segmentation framework, they used the newly released sequence-adaptive whole-brain segmentation⁷⁴ framework with an infant atlas for volumetric segmentation, improving their accuracy.

Cognitive assessment

Cognitive functioning was measured using the MSEL⁵¹; this assessment has a high test–retest reliability. The battery consists of 144 items that are distributed across five main subtests: expressive and receptive language, visual reception and fine and gross motor function. Raw scores can be used to generate standardized norm-referenced *T* scores, percentile ranks and age-equivalent scores. We chose to focus on raw scores, as we were interested in actual changes in children’s abilities over time rather than their degree of difference from a normative sample⁷⁵. Raw scores for gross motor scales were available within the range of 75 to 1,275 d. Fine motor scale and visual reception scale data were available within the range of 75 to 1,776 d. Expressive and receptive language scores were available within the range of 75 to 2,963 d. The demographic distribution of children ($N = 1,238$; observations = 2,530) in the cognitive development analysis is provided in Supplementary Table 21.

Predictive measures

Birth measures were obtained from hospital records. A gestational age of <259 d was considered preterm, and a birthweight of <2,500 g was considered low birthweight. Parent-reported measures of their educational attainment and income were used to assess socioeconomic status. Maternal education was categorized as primary, secondary and tertiary based on The International Standard Classification of Education. Primary and secondary education were classified as low maternal education. The low-income variable was defined per country-specific norms. For Singapore, low income was <SGD \$2,000 per month^{76,77}; for South Africa, low income was <1,000 Rand per month⁷⁸; for the United States, low income was <US \$50,000 per year⁷⁹. Income was not collected in the German sample (Max Planck). Consequently, this sample is not included in the main analysis but is included in the first sensitivity analysis.

Statistical analysis

To analyze the age-related growth of ICV and subcortical structures ($l = 1, \dots, L$), we used nonlinear mixed models⁸⁰ with the following asymptotic function:

$$f_l(x_{ij}, \theta_{jk}) = \theta_{jk1} + (\theta_{jk2} - \theta_{jk1})e^{-\theta_3 x_{ij}} \quad (1)$$

where x_{ij} is the age of the i th observation for the j th participant, θ_{jk} is a vector of participant- and covariate-specific parameters defining the function, where θ_{jk1} is the asymptote, θ_{jk2} is the intercept, and θ_3 is the rate constant that is proportional to the relative rate of increase. This last parameter is not indexed by participant or covariate because it

is fixed in the model, whereas the asymptote and intercept had both fixed and random effects as

$$\theta_{jk1} = \theta_1 + \sum_{k=1}^K \beta_{k1} + u_{j1} \quad (2)$$

where θ_1 is a fixed population parameter, the fixed effect β_{k1} is covariate specific ($k = 1, \dots, K$), and the random effect u_{j1} is participant specific ($j = 1, \dots, J$). Likewise, θ_{jk2} follows Eq. (2). The fixed effects were preterm birth, sex, low birthweight, low maternal education, low family income and cohort. All fixed effects were coded as binary effects using dummy variables. The simplified form of the nonlinear mixed model is

$$y_{ijl} = f_l(x_{ij}, \theta_{jk}) + \varepsilon_{ijl} \quad (3)$$

where y_{ijl} is the ij th observation for the l th subcortical structure, and ε_{ijl} is an error term. The model error was assumed to follow a normal distribution with heterogenous variances between cohorts, but this was not formally tested.

Multiple-comparisons correction for 35 tests (7 volumes and 5 covariates) was applied by using Bonferroni correction with an original α level set at 0.05, resulting in a *P* value significance threshold of $P = 0.001$. For the main analysis and first sensitivity analysis, we used raw volumes, which are advantageous for comparisons across studies and for creating normative volumetric values for the age range⁸¹. Prior studies have shown that ICV correction can influence the interpretability of brain–behavior associations across ages. For example, Dhamala et al. reported that ICV correction reduces predictive accuracies for cognitive ability from gray matter volumes⁸². Moreover, ICV correction methods reduce both univariate sex differences and the accuracy of multivariate sex prediction based on gray matter volume^{82–84}. Finally, ICV and different brain regional volumes have unique growth trajectories across development such that correction for ICV will have different effects across the different periods^{81,85}.

For the second sensitivity analysis, the volumes were ICV scaled and modeled with a linear mixed model with covariate- and cohort-specific fixed effects, participant-specific random effects and cohort-specific error variances. The linear mixed model used for this sensitivity analysis is

$$y_{ijl}/ICV_{ijl} * 1000 = \mu_l + \sum_{k=1}^K \beta_{k1} + u_{j1} + \varepsilon_{ijl} \quad (4)$$

where μ_l is the overall mean for the l th subcortical structure, β_{k1} is covariate-specific and cohort-specific fixed effects, and u_{j1} is the j th participant random effect. ε_{ijl} has the same specifications as before.

To investigate if genetic relatedness affects the results of the analysis, we performed a third set of sensitivity analyses, removing a single twin/sibling from the EBDS cohort and rerunning the main model.

Cognitive changes across age were modeled with a linear mixed model for each cognitive scale. The linear mixed model used on the evaluation of cognitive changes across age is

$$y_{ijm} = \mu_m + x_{ijm} + \sum_{k=1}^K \beta_{km} + u_{jm} + \varepsilon_{ijm} \quad (5)$$

where y_{ijm} is the i th observation on the j th participant for the m th cognitive score, μ_m is the overall mean, x_{ijm} is the slope given by age, $\sum_{k=1}^K \beta_{km}$ is the sum of the covariate and cohort fixed effects, u_{jm} is the j th participant’s random effect, and ε_{ijm} is the error term for the m th model. Both random effects and the error term follow a normal distribution, where the random effects are assumed independent, and the error term has cohort-specific variances (heterogenous variances). The analysis was performed using the lme4 package⁸⁶ v1.1.31 in R. Multiple-comparisons correction for 25 tests (5 cognitive scores and 5

covariates) was applied by using a Bonferroni correction with an original α level set at 0.05, resulting in a P value significance threshold of $P = 0.002$.

To analyze the relationship between ICV and subcortical volumes and cognitive scores, Pearson's correlation was used. The predicted volumes at age 2 were tested for correlation with predicted cognitive scores at age 2 separately in children born preterm and full term. The brain volumes and cognitive measures were not acquired contemporaneously. We focused on age 2 for this analysis as it had dense data points, and gross motor skills is used only until 33 months. Furthermore, cognitive ability at age 2 is a strong predictor of cognitive outcomes at school age^{19,87}. A multiple-comparisons correction for 35 tests (5 cognitive scores and 7 brain volume measures) was applied by using a Bonferroni correction with an original α level set at 0.05, resulting in a P value significance threshold of $P = 0.001$. We also performed a bootstrap method (bootcorci package v 0.0.0.9 in R⁸⁸) to compute confidence intervals for the differences in correlation coefficients between the full-term and preterm groups (Supplementary Table 19).

We additionally performed a replication analysis to test the robustness of the results. The whole sample was randomly split into two folds, and we replicated the analysis (volume trajectory, development of cognitive and motor scores and correlation analysis) 100 times. We have reported the proportion of times both the folds showed the same direction of effect and proportion of times where the results from both folds were of the same sign and significant, which is a more stringent approach.

To test if brain volumes mediate relationships between predictors and cognition, we examined the indirect effects wherever significant brain volume–cognitive score correlations were observed and for those predictors (sex, birthweight, maternal education and family income) that were associated with volumetric measures and cognitive scores. The effects were reported as 95% confidence intervals (significant when they did not include 0) based on 10,000 bootstrapped samples using the mediation package V.4.5.0 in R⁸⁹.

Addressing site-dependent variability. Data generated from different cohorts may be subject to systematic differences due to the technologies used to collect and process imaging data as well as systematic differences due to biological effects not fully accounted for in the model (for example, geographical differences). Therefore, our models contemplated the possibility that cohorts may have systematic differences in the mean and the scale of the traits. Specifically, all of our models included the random effect of the cohort on the outcomes (this adjusts for mean differences between cohorts) as well as cohort-specific error variances, which account for possible scale differences. Furthermore, we checked the distribution of model residuals (Supplementary Figs. 9 and 10). The residual plots show that before modeling the distribution, the volumes are clearly bimodal. This primarily reflects differences in age both within and between cohorts but may also reflect effects due to birth outcomes, sociodemographic factors and cohorts. The histogram of residuals shows that after modeling, the residuals are reasonably normal, suggesting that the model accounted for differences due to the effects included in it, including cohort. For thalamus, amygdala, putamen and pallidum volume, we do observe that the Cape Town and UCI cohorts deviate slightly from the normal distribution. We performed another sensitivity analysis removing these cohorts, and the results reaffirm that our inferences are robust (Supplementary Table 22).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data for the study came from eight different cohorts. The data for four of the cohorts are deposited in the National Institute of Mental

Health Data Archive (NDA) and can be accessed by submitting a data access request to the NDA. Imaging data for twins in the EBDS cohort are available through NDA 1974 and NDA 2384 and for singletons via NDA 4314. Imaging data from IBIS are available via NDA 19 and NDA 2027. Imaging data for UCI are available via NDA 1890. Imaging data for BCP are available via NDA 2848. Imaging data from the Harvard cohort and a subset of EBDS, IBIS and BCP data will also be made available through NDA 3905. Cognitive data from all cohorts and imaging data for the other cohorts may be available upon request to the parent cohort and may require Institutional Review Board approval or data use agreements. Investigators interested in further information on ORIGINS dataset access and sharing can contact the corresponding author.

Code availability

All code used in the analysis in this paper is available in the GitHub repository at https://github.com/knickmeyer-lab/ORIGINS_ICV-and-Subcortical-volume-development-in-early-childhood.

References

- Zhu, J. et al. Integrated structural and functional atlases of Asian children from infancy to childhood. *NeuroImage* **245**, 118716 (2021).
- Du, J., Younes, L. & Qiu, A. Whole brain diffeomorphic metric mapping via integration of sulcal and gyral curves, cortical surfaces, and images. *NeuroImage* **56**, 162–173 (2011).
- Poh, J. S. et al. Developmental synchrony of thalamocortical circuits in the neonatal brain. *NeuroImage* **116**, 168–176 (2015).
- Qiu, A. et al. COMT haplotypes modulate associations of antenatal maternal anxiety and neonatal cortical morphology. *Am. J. Psychiatry* **172**, 163–172 (2015).
- Qiu, A. et al. Maternal anxiety and infants' hippocampal development: timing matters. *Transl. Psychiatry* **3**, e306 (2013).
- Ducharme, S. et al. Trajectories of cortical thickness maturation in normal brain development—the importance of quality control procedures for the Brain Development Cooperative Group HHS Public Access. *NeuroImage* **125**, 267–279 (2016).
- Shi, F. et al. Infant brain atlases from neonates to 1- and 2-year-olds. *PLoS ONE* **6**, e18746 (2011).
- Tzourio-Mazoyer, N. et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* **15**, 273–289 (2002).
- Wang, J. et al. Multi-atlas segmentation of subcortical brain structures via the AutoSeg software pipeline. *Front. Neuroinform.* **8**, 7 (2014).
- Yushkevich, P. A. et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *NeuroImage* **31**, 1116–1128 (2006).
- Maltbie, E. et al. Asymmetric bias in user guided segmentations of brain structures. *NeuroImage* **59**, 1315–1323 (2012).
- Gousias, I. S. et al. Magnetic resonance imaging of the newborn brain: automatic segmentation of brain images into 50 anatomical regions. *PLoS ONE* **8**, e59990 (2013).
- Zöllei, L., Iglesias, J. E., Ou, Y., Grant, P. E. & Fischl, B. Infant FreeSurfer: an automated segmentation and surface extraction pipeline for T1-weighted neuroimaging data of infants 0–2 years. *NeuroImage* **218**, 116946 (2020).
- Fischl, B. FreeSurfer. *NeuroImage* **62**, 774–781 (2012).
- Ghosh, S. S. et al. Evaluating the validity of volume-based and surface-based brain image registration for developmental cognitive neuroscience studies in children 4 to 11 years of age. *NeuroImage* **53**, 85–93 (2010).
- Puonti, O., Iglesias, J. E. & Van Leemput, K. Fast and sequence-adaptive whole-brain segmentation using parametric Bayesian modeling. *NeuroImage* **143**, 235–249 (2016).

75. Naigles, L. R. et al. Neural correlates of language variability in preschool-aged boys with autism spectrum disorder. *Autism Res.* **10**, 1107–1119 (2017).
76. Padmapriya, N. et al. Association of physical activity and sedentary behavior with depression and anxiety symptoms during pregnancy in a multiethnic cohort of Asian women. *Arch. Women's Ment. Health* **19**, 1119–1128 (2016).
77. Quah, P. L. et al. Validation of the Children's Eating Behavior Questionnaire in 3 year old children of a multi-ethnic Asian population: the GUSTO cohort study. *Appetite* **113**, 100–105 (2017).
78. Donald, K. A. et al. Risk and protective factors for child development: an observational South African birth cohort. *PLoS Med.* **16**, e1002920 (2019).
79. U. S. Department of Health and Human Services. 2021 Poverty Guidelines. <https://aspe.hhs.gov/2021-poverty-guidelines> (2021).
80. Davidian, M. & Giltinan, D. M. Nonlinear models for repeated measurement data: an overview and update. *J. Agric. Biol. Environ. Stat.* **8**, 387–419 (2003).
81. Voevodskaya, O. et al. The effects of intracranial volume adjustment approaches on multiple regional MRI volumes in healthy aging and Alzheimer's disease. *Front. Aging Neurosci.* **6**, 264 (2014).
82. Dhamala, E. et al. Proportional intracranial volume correction differentially biases behavioral predictions across neuroanatomical features, sexes, and development. *NeuroImage* **260**, 119485 (2022).
83. Sanchis-Segura, C., Ibañez-Gual, M. V., Aguirre, N., Gómez-Cruz, Á. J. & Forn, C. Effects of different intracranial volume correction methods on univariate sex differences in grey matter volume and multivariate sex prediction. *Sci. Rep.* **10**, 12953 (2020).
84. Sanchis-Segura, C. et al. Sex differences in gray matter volume: how many and how large are they really? *Biol. Sex Differ.* **10**, 32 (2019).
85. Caspi, Y. et al. Changes in the intracranial volume from early adulthood to the sixth decade of life: a longitudinal study. *NeuroImage* **220**, 116842 (2020).
86. Bates, D., Mächler, M., Bolker, B. M. & Walker, S. C. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
87. McCall, R. B., Hogarty, P. S. & Hurlburt, N. Transitions in infant sensorimotor development and the prediction of childhood IQ. *Am. Psychol.* **27**, 728–748 (1972).
88. Rousselet, G., Pernet, D. C. & Wilcox, R. R. An introduction to the bootstrap: a versatile method to make inferences by using data-driven simulations. *Meta-Psychol.* (in the press).
89. Tingley, D., Yamamoto, T., Hirose, K., Keele, L. & Imai, K. mediation: R package for causal mediation analysis. *J. Stat. Softw.* **59**, 1–38 (2014).

Acknowledgements

We thank the families and children who participated in this study. ORIGINS is supported by the National Institute of Mental Health (R01MH123716 to R.K.). The ENIGMA consortium is supported by the National Institutes of Health (NIH) Big Data to Knowledge award for foundational support and consortium development (grant number U54 EB020403 to P.M.T.). The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. M.A.S. is supported by a Heisenberg Program Grant of the German Research Foundation (project number 433758790) and a Research Fellowship of the

Jacobs Foundation (fellowship number 2020136212). The GUSTO study was supported by the Singapore National Research Foundation under its Translational and Clinical Research (TCR) Flagship Programme and administered by the Singapore Ministry of Health's National Medical Research Council (NMRC; NMRC/TCR/004-NUS/2008 and NMRC/TCR/012-NUHS/2014). The Drakenstein Child Health Study was funded by the Bill and Melinda Gates Foundation (OPP 1017641), the South African Medical Research Council and the National Research Foundation of South Africa. K.A.D. received support from the Academy of Medical Sciences Newton Advanced Fellowship (NAFO02/1001), funded by the UK Government's Newton Fund, National Institute on Alcohol Abuse and Alcoholism via R21AA023887, the Collaborative Initiative on Fetal Alcohol Spectrum Disorders developmental grant (U24 AA014811), the US Brain and Behavior Foundation Independent Investigator grant (24467) and the Carnegie Corporation of New York. K.A.D., D.J.S. and H.J.Z. received financial support from the South African Medical Research Council. A.R. was supported by the National Research Foundation of South Africa. J.H.G. was supported by grants from the NIH (MH070890, HD053000 and MH064065). M.S. was supported by R33 MH104330 and P50 HD103573. The IBIS study was supported by grants from the NIH (R01-HD055741, P30-HD003110 and T32-HD040127) and the Simons Foundation (140209). J.B.G. was supported by K01-MH122779. The UCI study was supported by US PHS (NIH) grants R01 HD-060628, R01 MH-105538 and UH3 OD-023349 and European Research Council grant ERC-Stg 639766. P.D.W. was supported by NIH R01 MH-091361, R01 MH-105538 and UH3 OD-023349. S.E. was supported by R01 HD065825. W.L. was supported by 5U01MH110274. The Boston study was funded by NIH NICHD R01 HD065762-10, the William Hearst Fund (Harvard University) and the Harvard Catalyst/NIH (5UL1R025758).

Author contributions

R.C.K. is chair of the ENIGMA-ORIGINS group. A.M.A., G.d.I.C., P.M.T. and R.C.K. designed the research. A.M.A., F.A., G.d.I.C. and R.C.K. performed the research. M.A.S., Y.-S.C., M.J.M., M.V.F., A.Q., K.A.D., D.J.S., H.J.Z., N.A.G., J.H.G., J.P., H.H., K.B., S.R.D., R.T.S., M.S., J.B.G., C.B., S.E., D.A.F., A.M.G., P.D.W., J.M.R., A.R., W.L., N.G., T.K.T. and L.Z. contributed data. A.M.A., F.A. and G.d.I.C. analyzed the data. A.M.A. and R.C.K. wrote the paper. All authors edited the paper.

Competing interests

The authors declare no conflicts of interest.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41593-023-01501-6>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41593-023-01501-6>.

Correspondence and requests for materials should be addressed to Rebecca C. Knickmeyer.

Peer review information *Nature Neuroscience* thanks Ashok Panigrahy and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Extended Data Table 1 | Birth outcomes and socio-demographic factors have significant effects on structural brain development

		ICV					Thalamus					Hippocampus				
		Estimate	Std.Error	DF	t-value	p-value	Estimate	Std.Error	DF	t-value	p-value	Estimate	Std.Error	DF	t-value	p-value
Reference	Intercept	402952.30	2676.98	1317	150.52	0.00E+00 **	8924.17	42.15	1395	211.72	0.00E+00 **	2190.29	13.16	1369	166.45	0.00E+00 **
	Asymptote	1285208.00	6309.91	1317	203.68	0.00E+00 **	13292.35	80.24	1395	165.67	0.00E+00 **	5159.63	31.85	1369	162.01	0.00E+00 **
	Growth Rate	-5.66	0.01	1317	-757.85	0.00E+00 **	-6.64	0.02	1395	-285.01	0.00E+00 **	-6.02	0.01	1369	-521.95	0.00E+00 **
Male Sex	Intercept	19177.04	2690.45	1317	7.13	1.68E-12 **	423.13	42.57	1395	9.94	1.54E-22 **	80.27	13.34	1369	6.02	2.30E-09 **
	Asymptote	112624.10	6517.97	1317	17.28	1.88E-60 **	990.70	74.91	1395	13.23	1.05E-37 **	365.18	32.32	1369	11.30	2.28E-28 **
Preterm Birth	Intercept	-48142.85	4056.34	1317	-11.87	6.12E-31 **	-282.32	67.89	1395	-4.16	3.39E-05 **	-179.79	20.14	1369	-8.93	1.38E-18 **
	Asymptote	17063.40	9981.36	1317	1.71	8.76E-02	-184.55	112.09	1395	-1.65	9.99E-02	-93.51	47.82	1369	-1.96	5.07E-02
Low Birthweight	Intercept	-45116.71	4263.44	1317	-10.58	3.57E-25 **	-295.56	69.66	1395	-4.24	2.35E-05 **	-176.08	20.97	1369	-8.40	1.12E-16 **
	Asymptote	-26109.07	10355.73	1317	-2.52	1.18E-02	-110.19	115.98	1395	-0.95	3.42E-01	-87.64	49.65	1369	-1.77	7.78E-02
Primary/Secondary Maternal Education	Intercept	-16269.64	3203.33	1317	-5.08	4.34E-07 **	-186.31	55.66	1395	-3.35	8.38E-04 *	-4.71	16.63	1369	-0.28	7.77E-01
	Asymptote	-45624.15	8944.40	1317	-5.10	3.88E-07 **	-394.91	101.29	1395	-3.90	1.01E-04 **	-106.46	44.70	1369	-2.38	1.74E-02 **
Low Income Family	Intercept	4526.54	3174.30	1317	1.43	1.54E-01	-92.39	53.61	1395	-1.72	8.50E-02	-58.86	16.10	1369	-3.66	2.65E-04 **
	Asymptote	-14994.01	8452.62	1317	-1.77	7.63E-02	-398.43	95.18	1395	-4.19	3.01E-05 **	-162.85	41.57	1369	-3.92	9.40E-05 **
Number of observations		3168					3276					3242				
Number of Individuals		1835					1865					1857				
		Amygdala					Caudate					Putamen				
		Estimate	Std.Error	DF	t-value	p-value	Estimate	Std.Error	DF	t-value	p-value	Estimate	Std.Error	DF	t-value	p-value
Reference	Intercept	398.78	5.80	1380	68.76	0.00E+00 **	3711.22	22.54	1389	164.66	0.00E+00 **	3345.64	51.25	762	65.28	2.656518e-31 **
	Asymptote	2082.93	13.30	1380	156.59	0.00E+00 **	6947.72	53.36	1389	130.20	0.00E+00 **	8535.24	55.41	762	154.04	0.00E+00 **
	Growth Rate	-5.60	0.01	1380	-451.94	0.00E+00 **	-6.26	0.02	1389	-372.63	0.00E+00 **	-6.01	0.01	762	-446.26	0.00E+00 **
Male Sex	Intercept	19.84	5.48	1380	3.62	3.01E-04 **	114.07	21.93	1389	5.20	2.28E-07 **	226.31	50.16	762	4.51	7.45E-06 **
	Asymptote	190.05	13.37	1380	14.21	6.77E-43 **	325.09	53.41	1389	6.09	1.49E-09 **	671.36	59.57	762	11.27	2.34E-27 **
Preterm Birth	Intercept	-109.84	8.86	1380	-12.40	1.57E-33 **	-82.43	34.63	1389	-2.38	1.74E-02 *	-218.48	87.78	762	-2.49	1.30E-02 *
	Asymptote	-41.81	19.84	1380	-2.11	3.52E-02 *	-143.08	79.84	1389	-1.79	7.34E-02	-56.33	89.71	762	-0.63	5.30E-01
Low Birthweight	Intercept	-74.96	9.16	1380	-8.19	6.07E-16 **	-185.32	35.98	1389	-5.15	2.97E-07 **	-385.05	90.67	762	-4.25	2.44E-05 **
	Asymptote	-24.73	20.63	1380	-1.20	2.31E-01	-146.68	82.80	1389	-1.77	7.67E-02	-165.53	93.09	762	-1.78	7.58E-02
Primary/Secondary Maternal Education	Intercept	7.27	6.52	1380	1.12	2.65E-01	-120.10	24.50	1389	-4.90	1.06E-06 **	383.86	70.81	762	5.42	7.96E-08 **
	Asymptote	-88.94	18.62	1380	-4.78	1.97E-06 **	-298.09	71.94	1389	-4.14	3.63E-05 **	-260.66	84.23	762	-3.09	2.04E-03 **
Low Income Family	Intercept	-13.37	6.46	1380	-2.07	3.87E-02 *	10.61	25.67	1389	0.41	6.80E-01	-254.62	60.30	762	-4.22	2.71E-05 **
	Asymptote	-45.89	17.25	1380	-2.66	7.91E-03 *	-233.97	68.13	1389	-3.43	6.12E-04 **	-336.27	77.43	762	-4.34	1.59E-05 **
Number of observations		3257					3270					2245				
Number of Individuals		1861					1865					1469				
		Globus Pallidum														
		Estimate	Std.Error	DF	t-value	p-value										
Reference	Intercept	1141.27	16.79	762	67.98	0.00E+00 **										
	Asymptote	2114.39	15.88	762	133.17	0.00E+00 **										
	Growth Rate	-6.16	0.03	762	-224.00	0.00E+00 **										
Male Sex	Intercept	40.48	13.33	762	3.04	2.47E-03 **										
	Asymptote	182.94	16.34	762	11.19	4.89E-27 **										
Preterm Birth	Intercept	-71.87	22.75	762	-3.16	1.64E-03 **										
	Asymptote	-34.43	23.57	762	-1.46	1.45E-01										
Low Birthweight	Intercept	-96.05	24.03	762	-4.00	7.05E-05 **										
	Asymptote	-69.30	24.47	762	-2.83	4.75E-03 *										
Primary/Secondary Maternal Education	Intercept	205.47	17.43	762	11.79	1.42E-29 **										
	Asymptote	-92.64	22.38	762	-4.14	3.88E-05 **										
Low Income Family	Intercept	-51.61	14.59	762	-3.54	4.30E-04 **										
	Asymptote	-85.68	20.71	762	-4.14	3.91E-05 **										
Number of observations		2254														
Number of Individuals		1469														

Reference: Female, Full Term, Normal birthweight , Tertiary maternal education, medium high-income family . p value <0.05 *** and <0.0007 **** [Bonferroni threshold for p value for significance]. Two-tailed t-test.

Extended Data Table 2 | Demographic, birth outcomes and socioeconomic factors significantly influence cognitive and motor development

	Gross Motor Score				Visual Reception Score				Fine Motor Score				Receptive Language Score				Expressive Language Score			
	Estimate	Std.Error	t-value	p-value	Estimate	Std.Error	t-value	p-value	Estimate	Std.Error	t-value	p-value	Estimate	Std.Error	t-value	p-value	Estimate	Std.Error	t-value	p-value
Reference	7.65	0.18	43.35		6.13	0.20	30.69		7.50	0.15	49.88		5.76	0.20	28.42		5.52	0.20	27.53	
Age	0.03	0.0002	113.36	<2e-16 **	0.03	0.0003	113.55	<2e-16 **	0.03	0.0002	125.12	<2e-16 **	0.03	0.0002	114.56	<2e-16 **	0.03	0.0002	112.54	<2e-16 **
Male Sex	-0.19	0.12	-1.60	1.09E-01	-0.51	0.13	-3.93	8.89E-05 **	-0.30	0.10	-3.12	1.86E-03 **	-0.40	0.15	-2.66	8.00E-03 *	-0.40	0.15	-2.63	8.53E-03 *
Preterm Birth	-0.92	0.19	-4.71	2.68E-06 **	-0.88	0.22	-4.03	5.77E-05 **	-0.91	0.16	-5.60	2.58E-08 **	-1.32	0.25	-5.22	2.05E-07 **	-1.28	0.25	-5.04	5.18E-07 **
Low Birthweight	-0.26	0.20	-1.28	2.02E-01	-0.64	0.23	-2.78	5.43E-03 *	-0.52	0.17	-3.08	2.13E-03 *	-0.06	0.26	-0.23	8.14E-01	-0.18	0.26	-0.67	5.03E-01
Primary/Secondary Maternal Education	0.16	0.20	0.80	4.26E-01	-0.74	0.22	-3.35	8.15E-04 **	-0.34	0.16	-2.08	3.79E-02	-0.46	0.26	-1.81	7.02E-02	-0.27	0.26	-1.05	2.94E-01
Low Family Income	0.01	0.16	0.05	9.62E-01	-0.60	0.18	-3.31	9.62E-04 **	-0.11	0.13	-0.85	3.98E-01	-0.80	0.21	-3.87	1.16E-04 **	-0.53	0.21	-2.55	1.08E-02 *
Number of observations	2384				2404				2410				2472				2477			
Number of Individuals	1209				1203				1205				1222				1223			

p value <0.05 – **; <0.002 – *** [Bonferroni threshold for p value for significance]. Two-tailed t-test.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a | Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Data analysis

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The data for the study came from 8 different cohorts. The data for 4 of the cohorts is deposited in the NIMH Data Archive (NDA) and can be accessed by submitting a Data Access Request to the NDA. Imaging data for twins in the EBDS cohort is available through NDA #1974 and NDA #2384 and for singletons via NDA #4314.

Imaging data from IBIS is available via NDA #19 and NDA #2027. Imaging data for UCI is available via NDA #1890. Imaging data for BCP is available via NDA #2848. Imaging data from the cohorts HARVARD, some of IBIS and BCP will also be made available through NDA#3905. The cognitive data from all cohorts and imaging data for the other cohorts (GUSTO, DCHS, Max Planck, Boston Children's Hospital/Harvard Medical school) can be available upon request to the parent cohort and pending IRB approval.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

Sex was included in all models for analyzing the influence on brain volume and cognitive trajectories. Of the overall sample size of 2108 individuals, 1102 were males and 1006 were females.

Reporting on race, ethnicity, or other socially relevant groupings

Maternal ethnicity distribution among the participants is provided in the supplementary table 1.

Population characteristics

Overall, the imaging cohort included 2,108 children with a total of 3,607 observations. The age-range of acquired data spans 5–2,250 postnatal days. Cognitive scores data were available within range 75 – 2,963 days. 52.3% of the participants were male and 47.7% were female.

Recruitment

Recruitment was done by the parent cohorts.

Ethics oversight

Each project was approved by their respective local review board and informed consent was obtained from parents/legal guardian and children prior to data collection. The reviewing organizations include Michigan State University, USA; Max Planck Institute for Human Cognitive and Brain Sciences, Germany; National University of Singapore, Singapore; University of Cape town, South Africa; University of North Carolina, Chapel Hill, USA; University of California, Irvine, USA; Boston's Children Hospital, USA

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

Sample size was determined based on the availability of data that had all the relevant information from each of the cohorts.

Data exclusions

The individual cohorts had data exclusion criteria which are described in the supplementary section. We analyzed all the data that was send in.

Replication

The whole sample was randomly split into two folds, and we replicated the analysis (volume trajectory, development of cognitive and motor scores, correlation analysis) 100 times. We have reported the proportion of times both the folds showed same direction of effect and proportion of times where the results from both folds were of the same sign and significant, which is a more stringent approach. The results of the replication analysis are similar to findings in the main analysis. However, some of the associations are less robust and could be due to lower effect sizes and smaller sample sizes in the replication analysis.

Randomization

As this was an observational study, randomization does not apply.

Blinding

As this was an observational study, blinding does not apply.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input type="checkbox"/>	<input checked="" type="checkbox"/> MRI-based neuroimaging

Magnetic resonance imaging

Experimental design

Design type	This is an observational study using structural MRI data
Design specifications	As this was an observational study, number of blocks, trials or experimental units per session and/or subject does not apply.
Behavioral performance measures	This study used structural MRI measures. No tasks performed in scanner; we do integrate subcortical volume measures with cognitive development assessed with the Mullen Scales of Early Learning

Acquisition

Imaging type(s)	Structural
Field strength	Primarily 3T, one site used 1.5T
Sequence & imaging parameters	Varies by site, full details are provided in supplementary tables 2 and 3 and 4
Area of acquisition	whole brain
Diffusion MRI	<input type="checkbox"/> Used <input checked="" type="checkbox"/> Not used

Preprocessing

Preprocessing software	Varies by site, full details are provided in the cohort characteristics in methods section and in Supplementary Table 4
Normalization	Varies by site, full details are provided in the cohort characteristics in methods section and in Supplementary Table 4
Normalization template	Varies by site, full details are provided in the cohort characteristics in methods section and in Supplementary Table 4
Noise and artifact removal	Varies by site, full details are provided in the cohort characteristics in methods section and in Supplementary Table 4
Volume censoring	No volume censoring

Statistical modeling & inference

Model type and settings	To map longitudinal brain development, we fitted (mixed-effects) subject-specific non-linear longitudinal growth curves to ICV and subcortical structures (thalamus, amygdala, hippocampus, caudate, putamen, and pallidum). Our growth curve models have subject-specific intercepts (i.e., volume at birth), asymptote, and growth rate parameters. In our hierarchical model, we included effects of birth outcomes and socio-demographic factors on intercepts and asymptotes, and random effects of cohort and subject. To assess brain-cognition correlations, we used Pearson's correlation between predicted brain volumes (ICV and subcortical structures), and cognitive scores at 2 years of age.
Effect(s) tested	We tested the effects of sex, birth outcomes (preterm birth and low birth weight) and socio-demographic factors (low maternal education, low maternal income) on intercepts and asymptotes of the growth curves. To assess brain-cognition correlations, we used Pearson's correlation between predicted brain volumes (ICV and subcortical structures), and predicted cognitive scores at 2 years of age. Cognitive scores included raw scores for expressive and receptive language, visual reception, and fine and gross motor function, assessed with the Mullen Scales of Early Learning
Specify type of analysis:	<input type="checkbox"/> Whole brain <input checked="" type="checkbox"/> ROI-based <input type="checkbox"/> Both

Anatomical location(s) Focus was on ICV and subcortical structures (thalamus, amygdala, hippocampus, caudate, putamen, and pallidum). Processing pipelines varied by site and full details are provided in the cohort characteristics in methods section and in Supplementary Table 4

Statistic type for inference

neither voxel or cluster-based approaches were used in the study.

(See [Eklund et al. 2016](#))

Correction

Multiple comparisons correction for 35 tests (7 volumes and 5 covariates) was applied by using Bonferroni correction with original alpha level set at 0.05 resulting in a p-value significance threshold of $p=0.001$.

Models & analysis

- | n/a | Involvement in the study |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Functional and/or effective connectivity |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Graph analysis |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Multivariate modeling or predictive analysis |