LETTER

doi:10.1038/nature21369

Early brain development in infants at high risk for autism spectrum disorder

Heather Cody Hazlett[1,2], Hongbin Gu[1], Brent C. Munsell[3], Sun Hyung Kim[1], Martin Styner[1], Jason J. Wolff[4], Jed T. Elison[5], Meghan R. Swanson[2], Hongtu Zhu[6], Kelly N. Botteron[7], D. Louis Collins[11], John N. Constantino[7], Stephen R. Dager[8,9], Annette M. Estes[9,10], Alan C. Evans[11], Vladimir S. Fonov[11], Guido Gerig[12], Penelope Kostopoulos[11], Robert C. McKinstry[13], Juhi Pandey[14], Sarah Paterson[15], John R. Pruett Jr[7], Robert T. Schultz[14], Dennis W. Shaw[8,9], Lonnie Zwaigenbaum[16], Joseph Piven[1,2] & the IBIS Network*

Brain enlargement has been observed in children with autism spectrum disorder (ASD), but the timing of this phenomenon, and the relationship between ASD and the appearance of behavioural symptoms, are unknown. Retrospective head circumference and longitudinal brain volume studies of two-year olds followed up at four years of age have provided evidence that increased brain volume may emerge early in development[1,2]. Studies of infants at high familial risk of autism can provide insight into the early development of autism and have shown that characteristic social deficits in ASD emerge during the latter part of the first and in the second year of life[3,4]. These observations suggest that prospective brain-imaging studies of infants at high familial risk of ASD might identify early postnatal changes in brain volume that occur before an ASD diagnosis. In this prospective neuroimaging study of 106 infants at high familial risk of ASD and 42 low-risk infants, we show that hyperexpansion of the cortical surface area between 6 and 12 months of age precedes brain volume overgrowth observed between 12 and 24 months in 15 high-risk infants who were diagnosed with autism at 24 months. Brain volume overgrowth was linked to the emergence and severity of autistic social deficits. A deep-learning algorithm that primarily uses surface area information from magnetic resonance imaging of the brain of 6–12-month-old individuals predicted the diagnosis of autism in individual high-risk children at 24 months (with a positive predictive value of 81% and a sensitivity of 88%). These findings demonstrate that early brain changes occur during the period in which autistic behaviours are first emerging.

We first reported increased brain volume in adolescents and adults with ASD over twenty years ago[5]. Subsequent reports suggested that brain overgrowth in ASD may be most apparent during early childhood[6–8]. A study of infants at risk for ASD (33 high-risk and 22 low-risk infants), scanned from 6 to 24 months of age, found enlarged brain volume present at 12 and 24 months in the 10 infants that were later diagnosed with autism at 24 months of age or later[9] (mean age, 32.5 months).

In the present study, we examined data from a subset of individuals from a longitudinal study comprising 318 infants at high familial risk for ASD (HR), of which 70 met clinical best-estimate criteria for ASD (HR-ASD) and 248 that did not meet the criteria for ASD (HR-neg) at 24 months of age, and 117 infants at low familial risk (LR) for ASD, who also did not meet the criteria for ASD at 24 months (see Methods for diagnostic and exclusion criteria). The three groups were comparable in (mean) race/ethnicity (85% white), family income, maternal age at birth (33 years old), infant birth weight (8 lb), and gestational age at birth (39 weeks). The HR-ASD group had more males than the other two groups (83% of the HR-ASD group was male compared to 59% and 57% of the LR and HR-neg groups, respectively) and mothers in the LR group had a higher education level (Extended Data Table 1).

Infants were evaluated at 6, 12 and 24 months of age, which included detailed behavioural assessments and high-resolution magnetic resonance imaging (MRI) of the brain, to prospectively investigate brain and behavioural trajectories during infancy. The analyses described below were conducted on a subset of 106 high-risk ($n = 15$ HR-ASD; $n = 91$ HR-neg) and 42 low-risk infants for whom all three MRI scans were successfully obtained. On the basis of our previous findings at 2–4 years of age[2], we hypothesized that brain overgrowth in ASD begins before 24 months of age; that overgrowth is associated with hyperexpansion of the cortical surface area; and that these early brain changes are temporally linked to the emergence of the defining behaviours of ASD. We also investigated whether differences in the development of brain characteristics might suggest early biomarkers (that is, occurring before the onset of the defining behaviours of ASD) for the detection of ASD.

We first examined group differences in the trajectories of brain growth rate (Fig. 1). The growth rate of the total brain volume (TBV) did not differ between groups from 6 to 12 months of age. However, pairwise comparisons at 24 months showed large effect sizes for HR-ASD compared to LR and HR-ASD compared to HR-neg. The HR-ASD group showed a significantly increased TBV growth rate in the second year compared to both the LR and HR-neg groups (Extended Data Table 2). In addition, the HR-ASD group showed a significantly increased surface area growth rate from 6 to 12 months of age compared to both the HR-neg and LR groups, with the most robust increases observed in the left/right middle occipital gyrus, right cuneus and right lingual gyrus area (see Fig. 2). No group differences were observed in cortical thickness. We observed a significant correlation between surface area growth rate of 6–12 months and enlargement in TBV at 24 months of age in all subjects ($r_{192} = 0.59$, $P < 0.001$), as well as in the combined HR subgroup ($r_{139} = 0.63$, $P < 0.001$). Raw means, standard deviations and effect sizes for the group comparisons of TBV and surface area are provided in Extended Data Table 3. Regional differences in surface area change rate (6–12 months) were observed in the HR-ASD group (Fig. 2).

[1]Department of Psychiatry, University of North Carolina, Chapel Hill, North Carolina 27599, USA. [2]Carolina Institute for Developmental Disabilities, Chapel Hill, North Carolina 27599, USA. [3]College of Charleston, Charleston, South Carolina 29424, USA. [4]Department of Educational Psychology, University of Minnesota, Minneapolis, Minnesota 55455, USA. [5]Institute of Child Development, University of Minnesota, Minneapolis, Minnesota 55455, USA. [6]Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina 27599, USA. [7]Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri 63110, USA. [8]Department of Radiology, University of Washington, Seattle, Washington 98105, USA. [9]Center on Human Development and Disability, University of Washington, Seattle, Washington 98105, USA. [10]Department of Speech and Hearing Sciences, University of Washington, Seattle, Washington 98105, USA. [11]Montreal Neurological Institute, McGill University, Montreal, Quebec H3A 0G4, Canada. [12]Tandon School of Engineering, New York University, New York, New York 10003, USA. [13]Mallinckrodt Institute of Radiology, Washington University, St. Louis, Missouri 63110, USA. [14]Center for Autism Research, The Children's Hospital of Philadelphia and University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [15]Department of Psychology, Temple University, Philadelphia, Pennsylvania 19122, USA. [16]Department of Pediatrics, University of Alberta, Edmonton, Alberta T6G 2R3, Canada.
*A list of participants and their affiliations appears at the end of the paper.

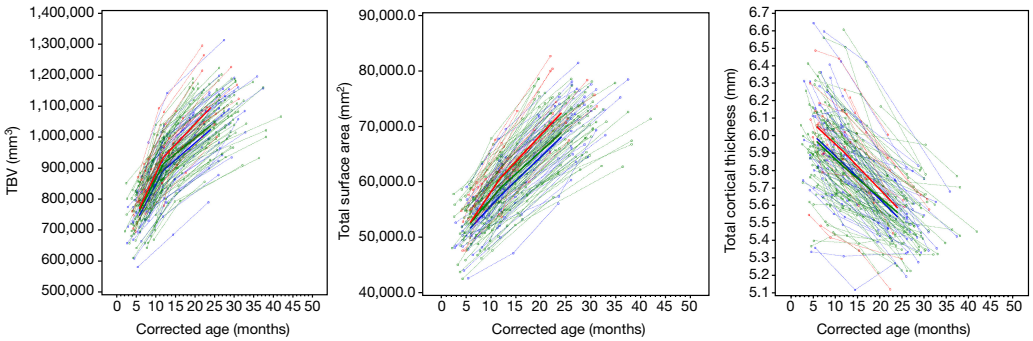


Figure 1 | Longitudinal trajectories of TBV, surface area and cortical thickness from 6 to 24 months. Longitudinal trajectories of TBV, cortical thickness and surface areas from 6 to 24 months for individuals from the HR-ASD (red), HR-neg (green) or LR (blue) groups. Only individuals with complete longitudinal imaging (for 6, 12, and 24 months) were included in the analysis (HR-ASD, $n = 15$; HR-neg, $n = 91$; LR, $n = 42$). Group trajectories were estimated from the random coefficient piecewise linear model (see Methods). The HR-ASD group showed a significantly increased surface area growth rate in the first year of life (from 6 to 12 months) compared to both the HR-neg ($t_{289} = 2.01$, $P = 0.04$) and LR groups ($t_{289} = 2.50$, $P = 0.01$). There were no significant group differences in surface area growth rates in the second year (Extended Data Table 2). Pairwise comparisons of surface area measured at 12 months of age showed medium to large effect sizes for HR-ASD vs LR (Cohen's $d = 0.74$) and HR-ASD vs HR-neg (Cohen's $d = 0.41$), which became more robust by 24 months for HR-ASD vs LR (Cohen's $d = 0.88$) and HR-ASD vs HR-neg (Cohen's $d = 0.70$). There were no significant group differences in trajectories for cortical thickness, with all groups showing a pattern of a decrease in cortical thickness over time. No group differences were observed in the trajectory of cortical thickness growth in either the first ($F_{2,289} = 0.00$; $P = 0.99$) or second year ($F_{2,289} = 1.44$; $P = 0.24$). The age (in months) is corrected by length (body size, in cm).

Given that the timing of TBV overgrowth in our study coincided with findings from other studies that have shown emergence of social deficits in the second year of life, we explored whether the rate of volume overgrowth was linked to autism severity. Pearson correlations between TBV and behavioural measurements of autism symptoms and social communication (on the Autism Diagnostic Observation Schedule[10] (ADOS) and Communication and Symbolic Behaviour Scales[11] (CSBS)) were generated, adjusting for multiple comparisons.

We first investigated the relationship between autistic behaviour (ADOS severity score) at 24 months and the TBV change rate at 6–12 and 12–24 months in the HR groups (HR-ASD and HR-neg). We found no significant correlation between the 24-month ADOS severity score and the 6–12 month TBV change rate ($r_{174} = 0.14$, $P = 0.06$); however, a significant correlation was found between the 24-month ADOS severity score and the 12–24 month TBV change rate ($r_{193} = 0.16$, $P = 0.03$). Subsequent analyses designed to investigate the components of overall autism severity (ADOS) during the latter interval revealed a significant correlation between the 12–24 month TBV change rate and the 24-month ADOS social affect score ($r_{194} = 0.17$, $P = 0.01$), but not the ADOS restricted/repetitive behaviour score ($r_{194} = 0.07$, $P = 0.31$).

To follow up on the relationship between change in brain volume and social deficits in the second year described above, we investigated the relationship between TBV change rates and social behaviour at 24 months with an independent measure of social behaviour, the CSBS. Consistent with the findings from the ADOS analysis, the CSBS social composite score was significantly correlated with a more rapid TBV change rate at 12–24 months ($r_{158} = 0.18$, $P = 0.02$) in HR subjects. No significant correlations were observed between CSBS social composite score at 24 months and TBV change rate at 6–12 months ($r_{143} = 0.11$, $P = 0.17$).

As opposed to the ADOS, which was only carried out at 24 months (the ADOS was primarily designed as a tool for diagnosis), measurements of social behaviour were available from the CSBS at both 12 and 24 months. We investigated the change in social behaviour during this time, taking into consideration our observation of changes in brain volume during that same period in the HR-ASD group, and a previous report that social deficits in ASD appear to unfold during the second year of life[3]. We observed a significant group (HR-ASD versus HR-neg) $\times$ time (12–24 months) interaction for the CSBS social composite score ($F_{2,130} = 10.0$, $P < 0.0001$). This finding was further supported by the observation that the CSBS effect size almost tripled from 12 ($d = 0.39$) to 24 ($d = 1.22$) months.

On the basis of earlier findings from our group with regards to surface area, cortical thickness and brain volume[2], we examined whether selected MRI brain measurements at 6 and 12 months of age could be used to accurately identify those infants who would later meet the criteria for ASD at 24 months of age. Independent of knowledge about the results of the above analyses, a machine learning classification algorithm based on a deep-learning network, was employed to investigate how well regional surface area and cortical thickness at 6 and 12 months, intracranial volume (at 6 and 12 months) and sex predicted HR-ASD diagnosis at 24 months of age. We used only data from those infants for whom cortical thickness and surface area data at both 6 and 12 months were available (HR-ASD = 34, HR-neg = 145). A tenfold cross-validation was used to compute classification performance, whereby the whole classification procedure, including network training was performed separately in each fold (see Supplementary Information for details on method, validation and comparison to other approaches).

The classification scheme distinguished the HR-ASD group from the HR-neg group in the cross-validation with 94% accuracy ($n = 168$ out of 179), 88% sensitivity ($n = 30$ out of 34), 95% specificity ($n = 138$ out of 145), 81% positive predictive value ($n = 30$ out of 37) and 97% negative predictive value ($n = 138$ out of 142) (Extended Data Table 4).

Additional analysis of the trained deep-learning network suggests that contributions to the discrimination are mostly on the basis of surface area and not cortical thickness (or TBV or sex), particularly at 6 months of age, as 11 of the top 12 measurements that contributed to the deep-learning network were regional surface area variables and the top six were variables from 6 months of age (Fig. 3 and Extended Data Fig. 1).

Our data suggest that very early, post-natal hyperexpansion of cortical surface areas may have an important role in the development of autism. The rate of cortical surface area expansion from 6 to 12 months was significantly increased in individuals diagnosed with autism at 24 months, and was linked to subsequent brain overgrowth, which, in turn, was linked to the emergence of social deficits. This suggests a sequence whereby hyperexpansion of the cortical surface area is an early event in a cascade leading to brain overgrowth and emerging autistic deficits. In infants diagnosed with autism at 24 months, surface area hyperexpansion in the first year was observed in cortical areas linked to the processing of sensory information (for example, the left middle occipital cortex), consistent with regions previously reported to show the earliest increase in surface area growth rate in typically developing infants[12], and with reports showing early sensory differences in infants who will later develop ASD[13,14].

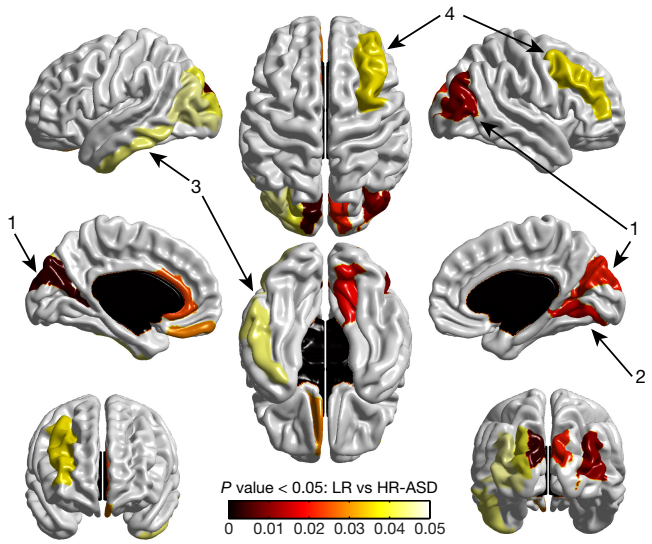


Figure 2 | Cortical regions that show significant expansion in surface area from 6 to 12 months in HR-ASD. A map of significant group differences in surface area from 6 to 12 months. Exploratory analyses were conducted with a surface map containing 78 regions of interest (see Supplementary Information), using an adaptive Hochberg method of $P < 0.05$. The coloured areas show the group effect for the HR-ASD versus LR subjects. Compared to the LR group, the HR-ASD group had significant expansion in the cortical surface area in the left/right middle occipital gyrus and right cuneus (1), right lingual gyrus (2), and to a lesser extent in the left inferior temporal gyrus (3), and middle frontal gyrus (4) (HR-ASD, $n = 34$; LR, $n = 84$).

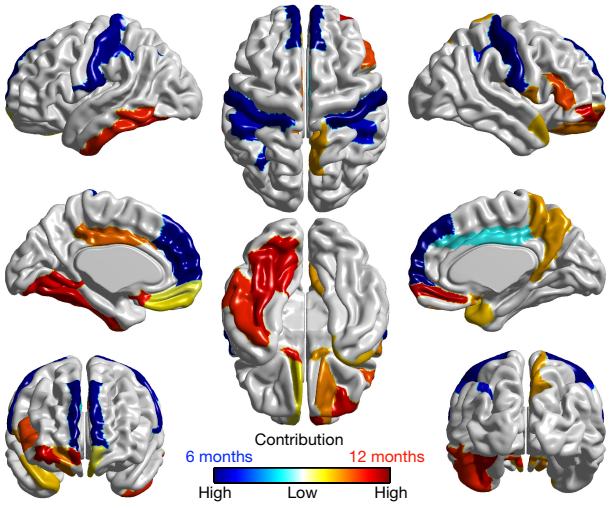


Figure 3 | Visualization of cortical regions with surface area measurements among the top 40 features contributing to the reduction in deep learning dimensionality. The cortical regions with surface area measurements that were among the top 40 features obtained from the nonlinear deep learning approach are visualized. The top 10 deep learning features observed include: surface area at 6 months in the right and left superior frontal gyrus, post-central gyrus, and inferior parietal gyri, and intracranial volume at 6 months. These features produced by the deep learning approach are highly consistent with those observed using an alternative approach (linear sparse learning) (Extended Data Fig. 1). Two tables listing the top 40 features from the deep learning approach and sparse learning are provided in Supplementary Tables 2 and 3.

The finding of brain overgrowth in this sample of young children with 'idiopathic' ASD is consistent with emerging literature demonstrating brain overgrowth in genetically defined ASD subgroups (for example, 16p11 deletions (ref. 15), CHD8 (ref. 16)). Cellular mechanisms and heritability that underlie surface area expansion are thought to differ from mechanisms that underlie cortical thickness[17,18], and surface area hyperexpansion has been reported in genetically-engineered mouse models of autism[19]. Our findings are not inconsistent with the mini-column hypothesis of autism[20], which postulates that symmetrical proliferation of periventricular progenitor cells leads to an increased number of mini-columns. These mini-columns may have a role in the pathogenesis of surface area hyperexpansion and the later emergence of the disorder[18,21]. Overproliferation of cortical progenitor cells may affect other mechanisms of post-natal development (for example, dendritic arborization and decreased pruning[22]). Overproduction of upper-layer neurons in the neocortex was previously shown to be associated with autism-like features in mice[23] and the 16p11.2-deletion mouse has been shown to exhibit altered cortical progenitor proliferation[24]. Furthermore, an imaging study described increased brain volume in individuals with a 16p11 deletion, a genetically defined subgroup of individuals often presenting with 'syndromic autism'[15]. Expansion of basal progenitor cells in rodent models[25] has been shown to regulate cerebral volume size and folding, while the dysregulation of neural-progenitor-cell proliferation has been observed in genetically engineered mouse models of ASD-associated genes (for example, *CHD8*)[26]. The importance of CHD8 in mediating regulatory networks during neurodevelopment was previously demonstrated[27] and suggests a potential role of CHD8 in disrupting the proliferation and differentiation of neurons during early human brain development. In addition, increased rates of proliferation of neural progenitor cells and neuron number compared to controls have been observed in induced pluripotent stem cells derived from individuals with ASD who also had increased brain volume on MRI[28]. Increased proliferation resulted from dysregulation of a β-catenin/BRN2 transcriptional cascade and was associated with reduced synaptogenesis that led to functional defects in neuronal networks, and these proliferation deficits could be rescued by stimulating the insulin growth factor 1 pathway[28]. The findings in the present study together with these recent reports suggest that understanding the mechanisms that underlie surface area hyperexpansion in the first year in human infants can probably provide important insights into the downstream pathogenesis of autism.

Prediction models developed from behaviourally based algorithms during infancy have not provided sufficient predictive power to be clinically useful[29]. We found that a deep-learning algorithm primarily using surface area information from brain MRI at 6 and 12 months of age predicted the 24 month diagnosis of autism in children at high familial risk for autism. This finding may have implications for early detection and intervention, given that this period is before consolidation of the defining features of ASD and the typical age for diagnosis[30]. The latter part of the first and early second years of life are characterized by greater neural plasticity relative to later ages and is a time when the social deficits associated with autism are not yet well established. Intervention at this age may prove more efficacious than later in development. The fact that we demonstrate group differences in surface area growth rate from 6 to 12 months, that very early surface area changes are linked to later brain overgrowth in the second year, and that overgrowth is, in turn, linked to the emergence of core social deficits in autism during this period, provides additional context to support the validity of the prediction model we report. The positive predictive value findings from this study on high-risk infants are probably conservative in nature owing to the likelihood that our HR-ASD group may be higher functioning than those who are clinically referred and diagnosed with ASD at 24 months of age, and that HR-neg groups are known to be more heterogeneous with respect to later development of cognitive, behavioural, social-communication and motor deficits than typical case–control studies[29,31–33]. The algorithm described in this paper will require replication before it could be considered a possible clinical tool for predicting ASD in high familial risk infants, as false diagnostic predictions have the potential to adversely affect individuals

and families. In addition, we do not know whether the brain differences we observed are specific to so-called idiopathic autism or share characteristics with other neurodevelopmental disorders. Although the findings of this study do not have direct application to the larger population of children with ASD who are not known to be at high familial risk for ASD, they provide a proof of principle that early prodromal detection using a brain biomarker may be possible. Future analyses incorporating complementary data from other relevant modalities (for example, behaviour, molecular genetics, electrophysiology and other imaging modalities such as whole brain functional MRI) may improve the accuracy of the prediction we observed.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 15 April 2016; accepted 6 January 2017.

1. Hazlett, H. C. *et al.* Magnetic resonance imaging and head circumference study of brain size in autism: birth through age 2 years. *Arch. Gen. Psychiatry* **62,** 1366–1376 (2005).
2. Hazlett, H. C. *et al.* Early brain overgrowth in autism associated with an increase in cortical surface area before age 2 years. *Arch. Gen. Psychiatry* **68,** 467–476 (2011).
3. Ozonoff, S. *et al.* A prospective study of the emergence of early behavioral signs of autism. *J. Am. Acad. Child Adolesc. Psychiatry* **49,** 256–266.e1, 2 (2010).
4. Zwaigenbaum, L. *et al.* Behavioral manifestations of autism in the first year of life. *Int. J. Dev. Neurosci.* **23,** 143–152 (2005).
5. Piven, J. *et al.* An MRI study of brain size in autism. *Am. J. Psychiatry* **152,** 1145–1149 (1995).
6. Courchesne, E. *et al.* Unusual brain growth patterns in early life in patients with autistic disorder: an MRI study. *Neurology* **57,** 245–254 (2001).
7. Schumann, C. M. *et al.* Longitudinal magnetic resonance imaging study of cortical development through early childhood in autism. *J. Neurosci.* **30,** 4419–4427 (2010).
8. Sparks, B. F. *et al.* Brain structural abnormalities in young children with autism spectrum disorder. *Neurology* **59,** 184–192 (2002).
9. Shen, M. D. *et al.* Early brain enlargement and elevated extra-axial fluid in infants who develop autism spectrum disorder. *Brain* **136,** 2825–2835 (2013).
10. Lord, C., Rutter, M., DiLavore, P. C. & Risi, S. The Autism Diagnostic Observation Schedule. (Western Psychological Services, 2000).
11. Wetherby, A. & Prizant, B. Communication and Symbolic Behavior Scales Developmental Profile, First Normed Edition. (Paul H. Brookes, 2002).
12. Li, G. *et al.* Mapping longitudinal hemispheric structural asymmetries of the human cerebral cortex from birth to 2 years of age. *Cereb. Cortex* **24,** 1289–1300 (2014).
13. Elison, J. T. *et al.* White matter microstructure and atypical visual orienting in 7-month-olds at risk for autism. *Am. J. Psychiatry* **170,** 899–908 (2013).
14. Estes, A. *et al.* Behavioral, cognitive, and adaptive development in infants with autism spectrum disorder in the first 2 years of life. *J. Neurodev. Disord.* **7,** 24 (2015).
15. Qureshi, A. Y. *et al.* Opposing brain differences in 16p11.2 deletion and duplication carriers. *J. Neurosci.* **34,** 11199–11211 (2014).
16. Bernier, R. *et al.* Disruptive CHD8 mutations define a subtype of autism early in development. *Cell* **158,** 263–276 (2014).
17. Panizzon, M. S. *et al.* Distinct genetic influences on cortical surface area and cortical thickness. *Cereb. Cortex* **19,** 2728–2735 (2009).
18. Rakic, P. A small step for the cell, a giant leap for mankind: a hypothesis of neocortical expansion during evolution. *Trends Neurosci.* **18,** 383–388 (1995).
19. Chenn, A. & Walsh, C. A. Regulation of cerebral cortical size by control of cell cycle exit in neural precursors. *Science* **297,** 365–369 (2002).
20. Casanova, M. F. *et al.* Minicolumnar abnormalities in autism. *Acta Neuropathol.* **112,** 287–303 (2006).
21. Hill, J. *et al.* Similar patterns of cortical expansion during human development and evolution. *Proc. Natl Acad. Sci. USA* **107,** 13135–13140 (2010).
22. Packer, A. Neocortical neurogenesis and the etiology of autism spectrum disorder. *Neurosci. Biobehav. Rev.* **64,** 185–195 (2016).
23. Fang, W. Q. *et al.* Overproduction of upper-layer neurons in the neocortex leads to autism-like features in mice. *Cell Reports* **9,** 1635–1643 (2014).
24. Pucilowska, J. *et al.* The 16p11.2 deletion mouse model of autism exhibits altered cortical progenitor proliferation and brain cytoarchitecture linked to the ERK MAPK pathway. *J. Neurosci.* **35,** 3190–3200 (2015).
25. Nonaka-Kinoshita, M. *et al.* Regulation of cerebral cortex size and folding by expansion of basal progenitors. *EMBO J.* **32,** 1817–1828 (2013).
26. Sugathan, A. *et al.* CHD8 regulates neurodevelopmental pathways associated with autism spectrum disorder in neural progenitors. *Proc. Natl Acad. Sci. USA* **111,** E4468–E4477 (2014).
27. Cotney, J. *et al.* The autism-associated chromatin modifier CHD8 regulates other autism risk genes during human neurodevelopment. *Nat. Commun.* **6,** 6404 (2015).
28. Marchetto, M. C. *et al.* Altered proliferation and networks in neural cells derived from idiopathic autistic individuals. *Mol. Psychiatry* http://dx.doi.org/10.1038/mp.2016.95 (2016).
29. Charman, T. Early identification and intervention in autism spectrum disorders: some progress but not as much as we hoped. *Int. J. Speech Lang Pathol.* **16,** 15–18 (2014).
30. Landa, R. J., Gross, A. L., Stuart, E. A. & Faherty, A. Developmental trajectories in children with and without autism spectrum disorders: the first 3 years. *Child Dev.* **84,** 429–442 (2013).
31. Messinger, D. *et al.* Beyond autism: a baby siblings research consortium study of high-risk children at three years of age. *J. Am. Acad. Child Adolesc. Psychiatry* **52,** 300–308.e1 (2013).
32. Georgiades, S. *et al.* A prospective study of autistic-like traits in unaffected siblings of probands with autism spectrum disorder. *JAMA Psychiatry* **70,** 42–48 (2013).
33. Ozonoff, S. *et al.* The broader autism phenotype in infancy: when does it emerge? *J. Am. Acad. Child Adolesc. Psychiatry* **53,** 398–407.e2 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements The IBIS (Infant Brain Imaging Study) Network is an NIH funded Autism Center of Excellence (HDO55741) and consists of a consortium of 8 Universities in the US and Canada. This work was supported by an NIH Autism Center of Excellence grant (NIMH and NICHD HD055741 to J.Pi.), Autism Speaks (6020) and the Simons Foundation (140209). Further support was provided by the National Alliance for Medical Image Computing (NA-MIC), funded by the NIH through grant U54 EB005149, the IDDRC Imaging and Participant Registry cores (NICHD HD003110 to J.Pi.) and R01 MH093510 (to J.R.P.Jr). We thank M. Burchinal and K. Y. Truong for their consultation on the statistical methods and approach. Given the large commitment of time and effort required by this study, we extend our appreciation to the families who have participated in this study and the numerous research assistants and staff who have contributed to this work.

Author Contributions All co-authors discussed the results, made critical contributions to the work and contributed to the writing of the manuscript. H.C.H., K.N.B., S.R.D., A.M.E., R.C.M., S.P., J.Pi., R.T.S., J. Pa. and D.W.S. contributed to the data collection. A.C.E., P.K. provided support for data management. B.C.M., S.H.K., M.S., D.L.C., A.C.E., V.S.F. and G.G. conducted image processing. H.G., B.C.M., S.H.K., M.S. and H.Z. analysed the data. H.C.H. wrote the manuscript with J.Pi., H.G., B.C.M., M.S. and with J.J.W., J.T.E., M.R.S., J.N.C., J.R.P.Jr, A.M.E., R.T.S. and L.Z. providing additional feedback.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.C.H. (hcody@med.unc.edu).

Reviewer Information *Nature* thanks M. Johnson, G. Ramsay, T. Yarkoni and the other anonymous reviewer(s) for their contribution to the peer review of this work.

IBIS Network

Clinical Sites J. Piven[1], H. C. Hazlett[1], C. Chappell[1], S. R. Dager[2], A. M. Estes[2], D. W. Shaw[2], K. N. Botteron[3], R. C. McKinstry[3], J. N. Constantino[3], J. R. Pruett Jr[3], R. T. Schultz[4], S. Paterson[4], L. Zwaigenbaum[5], J. T. Elison[6], J. J. Wolff[6]; **Data Coordinating Center** A. C. Evans[7], D. L. Collins[7], G. B. Pike[7], V. S. Fonov[7], P. Kostopoulos[7], S. Das[7]; **Image Processing Core** G. Gerig[8]; M. Styner[1]; **Statistical Analysis Core** H. Gu[1]

[1]University of North Carolina, Chapel Hill, North Carolina 27599, USA. [2]University of Washington, Seattle, Washington 98105, USA. [3]Washington University, St. Louis, Missouri 63130, USA. [4]Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA. [5]University of Alberta, Edmonton, Alberta T6G 2R3, Canada. [6]University of Minnesota, Minneapolis, Minnesota 55455, USA. [7]Montreal Neurological Institute, Montreal, Quebec H3A 0G4, Canada. [8]New York University, New York, New York 10003, USA.

METHODS

Data reporting. No statistical methods were used to predetermine sample size. The experiments were not randomized.

Sample. This study includes data acquired from an NIH-funded Autism Centers of Excellence (ACE) network study, referred to as the 'Infant Brain Imaging Study' (IBIS). The network includes four clinical data collection sites (University of North Carolina at Chapel Hill, University of Washington, Children's Hospital of Philadelphia, Washington University in St. Louis), a Data Coordinating Center at the Montreal Neurological Institute (McGill University), and two image processing sites (University of Utah and UNC). Data collection sites had study protocols approval from their Institutional Review Boards (IRB), and all enrolled subjects had informed consent provided by parent/guardian. Infants at high (HR) and low familial risk (LR) entered the study at 6 months of age (a subset of HR infants entered at 12 months) and were followed-up at 12 and 24 months. Results from the 6 month brain volume findings have previously been reported for a subset of this sample[34].

Subjects were enrolled as HR if they had an older sibling with a clinical diagnosis of ASD confirmed with the Autism Diagnostic Interview-Revised[35] (ADI-R). Subjects were enrolled in the LR group if they had an older sibling without evidence of ASD and no family history of a first or second-degree relative with ASD. Exclusion criteria for both groups 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 ASDs, (2) a significant medical or neurological condition affecting growth, development or cognition (for example, CNS infection, seizure disorder, congenital heart disease), (3) sensory impairment such as vision or hearing loss, (4) low birth weight (<2,000 g) or prematurity (<36 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, selected prescription medications), (6) non-English speaking families, (7) contraindication for MRI (for example, metal implants), (8) adopted subjects, and (9) a family history of intellectual disability, psychosis, schizophrenia or bipolar disorder in a first-degree relative. The sample for this analysis included all children with longitudinal imaging data processed until 31 August 2015. The final sample included 318 HR and 117 LR children, each with 2–3 MRI scans (Extended Data Table 1).

Assessment protocols. *Behavioural assessment.* Infants were assessed at ages 6, 12 and 24 months and received a brain MRI scan in addition to a battery of behavioural and developmental tests. The tests included measurements of cognitive development, adaptive functioning and behaviours associated with autism. Developmental level and adaptive functioning were assessed at each time point using the Mullen Scales of Early Learning[36] and Vineland Scales of Adaptive Behaviour[37]. Autism-oriented assessments included the Autism Diagnostic Interview-Revised[35], Autism Diagnostic Observation Scale[10] (ADOS-WPS) at 24 months and Communication and Symbolic Behaviour Scales of Development Profile[11] (CSBS-DP) at 12 and 24 months. From the CSBS, the total raw score and the social composite raw score were used in the brain–behavioural analyses. Raw scores were used to allow better representation of the distribution of the data.

Diagnostic (outcome) classification. Diagnostic classification was made by an expert clinician at each site using all clinical, behavioural and questionnaire data available at 24 months. A diagnosis of ASD was made using the DSM-IV-TR (diagnostic and statistical manual of mental disorders, edition IV, text revision)[38] criteria for autism and pervasive developmental disorder, not otherwise specified (PDD-NOS)[38] by an expert clinician blind to the outcome of the imaging results. Across the IBIS Network, the expert clinicians met quarterly for diagnostic reliability meetings (via video/telephone) using the DSM-IV-TR criteria independently. The rationale for this conservative approach was to maximize validity of diagnosis at 24 months of age[39]. Reliability between diagnostic raters was maintained throughout the project period.

HR subjects were classified as HR-neg (that is, negative for ASD) if they did not meet either ASD or PDD-NOS criteria on the DSM-IV-TR. In order to have a LR comparison group representing typically developing infants without autism, we also assessed each LR subject at 24 months. The LR subjects included did not meet ASD or PDD-NOS criteria on the DSM-IV-TR clinical best-estimate assessment at 24 months. Three LR subjects met DSM-IV criteria for ASD at their 24-month assessment (one for autism, two for PDD-NOS) and were excluded from the study (Extended Data Table 5). There is strong evidence of differences in the underlying genetic architecture of multiple versus single incidence (or sporadic) cases, with the latter more often being attributed to *de novo* events, that support our exclusion of these LR-ASD subjects from a combined analysis with the HR-ASD subject group, who are siblings of HR infants. The final HR groups included 70 HR-ASD and 248 HR-neg children and the LR group consisted of 117 children.

MRI acquisition. The brain MRI scans were completed on 3T Siemens Tim Trio scanners with 12-channel head coils and obtained while infants were naturally sleeping. The imaging protocol included (1) a localizer scan; (2) 3D T1 MPRAGE: TR = 2,400 ms, TE = 3.16 ms, 160 sagittal slices, FOV = 256, voxel size = 1 mm^3; (3) 3D T2 FSE TR = 3,200 ms, TE = 499 ms, 160 sagittal slices, FOV = 256, voxel size = 1 mm^3; and (4) a 25 direction DTI: TR = 12,800 ms, TE = 102 ms, slice thickness = 2 mm isotropic, variable *b* value = maximum of 1,000 s mm^{-2}, FOV = 190.

A number of quality control procedures were employed to assess scanner stability and reliability across sites, time and procedures. Geometry phantoms were scanned monthly and human phantoms (two adult subjects) were scanned annually to monitor scanner stability at each site across the study period. Details on the stability procedures for IBIS and scanner quality control checks are described elsewhere[34].

We further examined the three subject groups (HR-ASD, HR-neg, LR) for any differences in scan success rates (that is, proportion of completed scans). We found a significant difference between groups ($\chi^2_2 = 16.9$, $n = 1{,}305$, $P = 0.02$). Overall, the HR-ASD subjects had proportionately fewer successful scans (69%) compared to the HR-neg (78%) and LR (76%) groups, particularly at the 12-month visit. We hypothesize that this may be owing to more behavioural difficulties in the ASD group (for example, problems with sleep disturbance).

Radiologic review. All scans were reviewed locally by a paediatric neuroradiologist for radiologic findings that, if present, were communicated to the participant. In addition, a board certified paediatric neuroradiologist (R.C.M., Washington University) blindly reviewed all MRI scans across the IBIS network and rated the incidental findings. A third neuroradiologist (D.W.S., University of Washington) provided a second blind review for the Washington University site, and contributed to a final consensus rating if there were discrepancies between the local site reviews and the network review. The final consensus review was used to evaluate whether there were group differences in the number and/or type of incidental findings. Scans were rated as either normal, abnormal, or with incidental findings. No scans rated as abnormal were included in the analysis, and previous examinations of our data did not find group differences in incidental findings[34]. Scans rated as clinically abnormal by a site paediatric neuroradiologist, and independently confirmed by two study paediatric neuroradiologists, were excluded ($n = 3$).

Image processing. Image processing was performed to obtain global brain tissue volumes, regional brain tissue volumes and cortical measurements (surface area, cortical thickness). All image processing was conducted blind to the subject group and diagnostic information. The brain volumes were obtained using a framework of atlas-moderated expectation maximization including co-registration of multi-modal (T1w/T2w) MRI, bias correction, brain stripping, noise reduction and multivariate classification with the AutoSeg toolkit[40] (http://www.nitrc.org/projects/autoseg/). Population average templates and corresponding probabilistic brain tissue priors, for white matter (WM), grey matter (GM), and cerebrospinal fluid (CSF) were constructed for the 6–24-month-old brain. The following brain volumes were generated at all ages: intracranial volume (ICV), total brain volume = GM + WM, total cerebrospinal fluid (CSF), cerebrum, cerebellum, and lateral ventricles. Intracranial volume was defined as the sum of WM, GM and CSF. Total brain tissue volume (TBV) was defined as the sum of all WM and GM contained in the brain cavity (that is, cerebrum, cerebellum and a portion of midbrain/brainstem). Subjects were included in the volumetric analyses if they had successfully segmented scans at 6, 12 and 24 months and corresponding body length measurements.

Cortical thickness and surface area measurements for 12 and 24 month data were obtained via a CIVET workflow[41,42] adapted for this age using an age-corrected automated anatomical labelling (AAL) atlas[43,44]. CIVET includes shrink-wrap deformable surface evolution of WM, local Laplacian distance and local surface area, mapping to spherical domain, co-registration using cortical sulcal features and extraction of regional measurements through a deformably co-registered fine-scale lobar parcellation. Surface area was measured at the mid-cortical surface. Cortical thickness and surface area measurements for data of 6-month olds were extracted from surfaces propagated via deformable multi-modal, within-subject, co-registration[44] of MRI data at 12 months.

Statistical analysis. We used χ^2 tests to examine group differences in the categorical demographic variables, including race, gender, social economic income categories and mother's education (see Extended Data Table 1). For the continuous variables, including birth weight, mother's age at child birth, children's age at visits and behavioural measurements, we used an ANOVA to test group differences. A random-coefficient, piecewise longitudinal mixed model was used as a coherent framework to model brain growth trajectories in the first and second year and to test for group differences in growth trajectories. The three outcome variables investigated were total brain volume (TBV), surface area and cortical thickness. Each model included random coefficients for the first year growth rate (6–12 months)

and change of growth rate in the second year (12–24 months), and random intercepts for each child to account for individual differences and correlated repeated measures collected at 6, 12 and 24 months. For subject *i* from group *k* at month *j*, the brain measure is:

$$Y_{ikj} = (\beta_{0k} + b_{0ik}) + (\beta_{1k} + b_{1ik})t_{ikj} + (\beta_{2k} + b_{2ik})(t_{ikj} - 12) + e_{ikj}$$

The mean group growth rate in the first year will be β_{1k}, and the growth rate for the 12 months beyond will be $\beta_{1k} + \beta_{2k}$. The inclusion of the change of slope after 12 months is to capture the change in growth rate from the first to the second year. The first two years of life is a period of rapid brain development, with growth rate being faster in the first than the second year[45]. The two-piece linear mixed model was chosen to capture the change in growth rate from the first to the second year. We required all subjects in this analysis to have 3 completed scans at 6, 12, and 24 months. This reduced the HR-ASD sample from 70 to 15 subjects. This requirement is to ensure that we captured the individual growth rate change from the first to the second year without the potential bias caused by partial visits and changes in study cohorts at different visits. We examined possible bias in the HR-ASD subjects with three completed visits versus those with only one or two visits and found no significant differences in demographic (for example, sex, age) or outcome measure (for example, TBV, surface area). Results are shown in Extended Data Tables 6 and 7.

To model the unique brain overgrowth separate from the general body growth, we modelled the brain growth relative to normative body growth in the first two years[45,46]. Normative age based on body length was used instead of chronological age in order to capture brain overgrowth in the context of body growth. The normative age for each infant's body size (t_{ij}, corrected age) was used in the model as the continuous growth variable. The corrected age correlates highly with chronological age while taking into account the infant's sex and body size, which is necessary to determine the relative brain overgrowth. Sex-specific WHO height norms[47] were used to determine the corrected age based on an infant's sex and height (length).

We addressed potential sex-related brain differences in two ways. First, in order to account for sex-related body size differences and their effects on brain volume[48], we normalized differences in body size by using the sex-specific WHO height norms. Second, we included sex as a covariate in the analysis model to account for remaining sex-related differences[49]. The approach to include sex as a model covariate will account for a linear, fixed effect of sex differences in brain volume. However, for developmental studies, the sex differences in body size and brain volume may be nonlinear, with an unknown function form. Using a body size standardization based on normative sex-specific height data are more likely to account for nonlinear sex-related differences.

The final model covariates include site and sex. Despite regular cross-site calibration in both behavioural and imaging protocols, a site covariate was included to account for the possibility of cohort differences or potential administrative differences in a multi-site study. Sex was included as a covariate in the analysis model to account for remaining sex-related differences not accounted for by sex-specific body growth. However, when we analysed only males for group differences, our results remain unchanged (Extended Data Fig. 2 and Extended Data Table 8).

As a sensitivity analysis, we also tested the model with other demographic, familial and child birth-related covariates (race, social economic status, mother's education, mother's age at birth, birth weight and gestational age), and only the site and sex remained in the model with $P < 0.01$.

The association of 24-month clinical outcome (ADOS, CSBS) with brain growth rates (TBV) from 6–12 and 12–24 month intervals was assessed among HR subjects using a Pearson correlation. Family income, mother's education, subject sex and birth weight were examined as potential covariates, but none contributed significantly and were excluded from the final analysis.

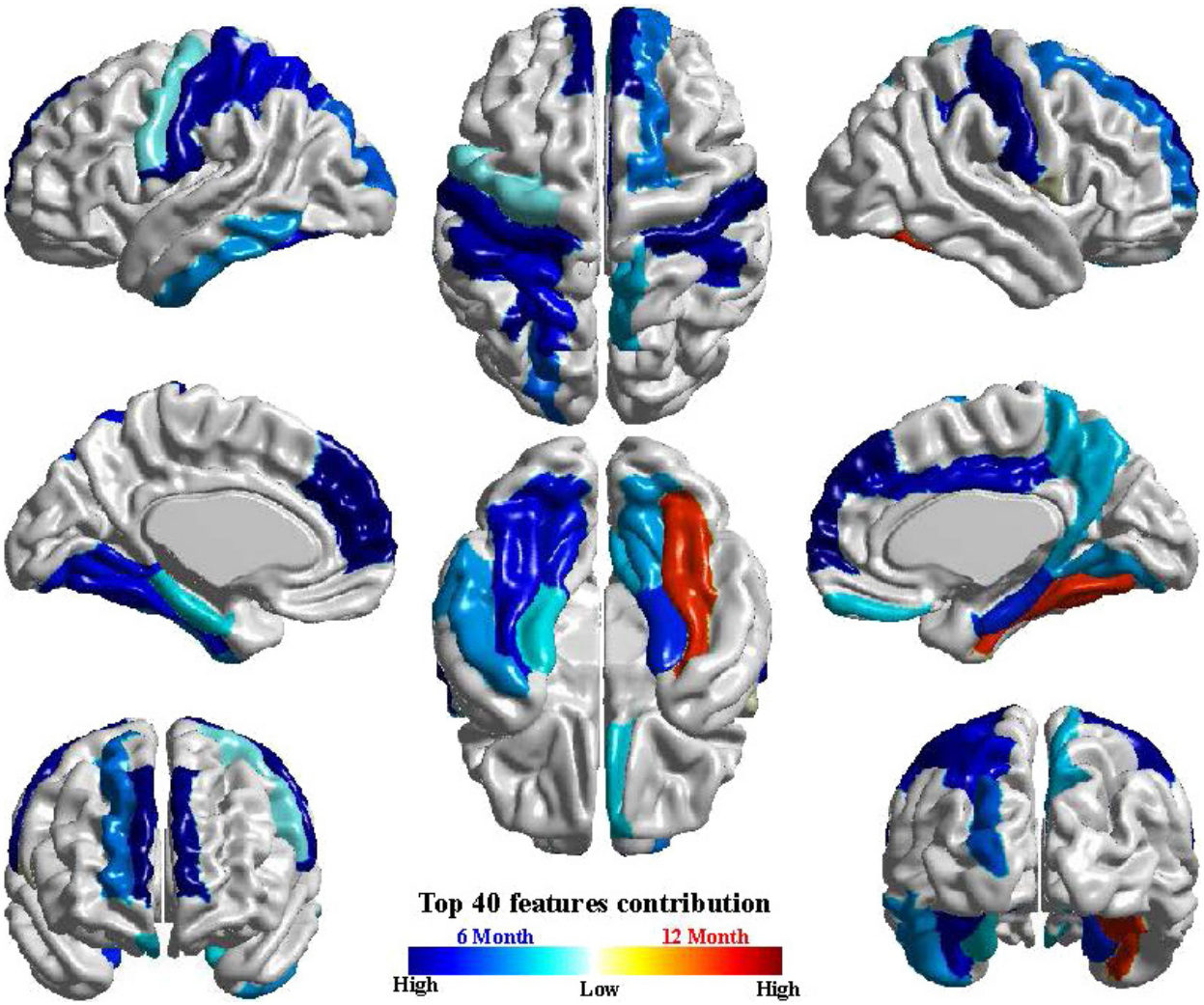
Multiple-comparison adjustments were performed for all pairwise comparisons and the correlation analyses, which followed the tests for overall group differences (*F* test, reported in Extended Data Table 2). All pairwise comparisons and correlation analyses used adaptive Hochberg multiple-comparison adjustments[50]. Only those comparisons that remained significant after the multiple comparison adjustment are reported (Fig. 1 and Extended Data Table 2).

The machine learning analysis used a nonlinear prediction model based on a standard three-stage deep-learning network and included the following unbiased/unweighted information: sex, age-corrected intracranial volume, and age-corrected surface area and cortical thickness measurements from 39 left and 39 right cortical hemisphere regions at 6 and 12 months (approximately 312 measurements). This analysis included 34 HR-ASD and 145 HR-neg subjects. The model was evaluated via a standard tenfold cross-validation. The core of the prediction model is a weighted three-stage neural/deep-learning network[51], where the first stage reduces 315 measurements to 100, the second stage reduces 100 to 10, and the third stage reduces 10 to only 2 measurements. At each stage, the measurements (in the progressively smaller sets) are the weighted combination of input measurements from the previous stage. In general, the training process determines (1) those network weights that retain information that are capable of distinguishing the affected condition (for example, HR-ASD) from the unaffected condition (HR-neg), as well as (2) the linear support vector machine based classification decision that separates the group label (HR-ASD and HR-neg) in the two-dimensional final network space. Thus to apply the prediction model, the data are first inserted into the two-dimension final network space using the trained deep-learning network, and then classified in the final network space using the trained support vector machine. All training was performed purely on the training data in each fold. Once training was achieved, this prediction model was applied to the testing data in each fold. Classification measurements of accuracy, sensitivity, specificity, positive predictive value and negative predictive value are combined and reported across the 10 folds. Details of our machine learning procedures and validity tests are provided in the Supplementary Information.

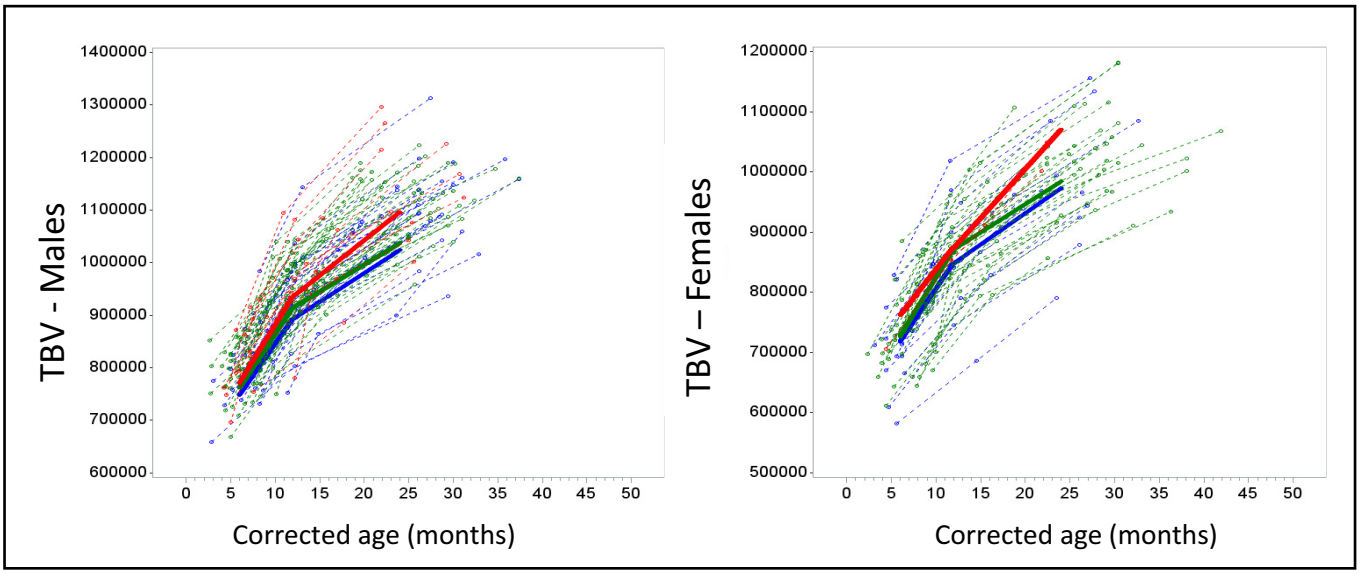
Data availability. The raw data that support the findings from this study are publically available from the NIH National Database for Autism Research (NDAR). Any additional data may be available from the corresponding author upon reasonable request. Source data for Figs 1–3 are provided with the paper.

Information on the following tools used in our analyses (AutoSeg, HeadCirc and ITK-SNAP) is freely available for download: http://www.med.unc.edu/psych/research/niral/download/download-software and http://www.nitrc.org. All of the Matlab source code used to construct, train and test the prediction pipeline is also available at https://github.com/munsellb/dl_autism.git.

34. Hazlett, H. C. *et al.* Brain volume findings in 6-month-old infants at high familial risk for autism. *Am. J. Psychiatry* **169,** 601–608 (2012).
35. Lord, C., Rutter, M. & Le Couteur, A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J. Autism Dev. Disord.* **24,** 659–685 (1994).
36. Mullen, E. M. Mullen Scales of Early Learning: AGS edn (American Guidance Service, 1995).
37. Sparrow, S., Balla, D. & Cicchetti, D. Vineland Scales of Adaptive Behavior: A Survey Form Manual. (American Guidance Service, 1984).
38. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (4th edn) (Washington, 2000).
39. Guthrie, W., Swineford, L. B., Nottke, C. & Wetherby, A. M. Early diagnosis of autism spectrum disorder: stability and change in clinical diagnosis and symptom presentation. *J. Child Psychol. Psychiatry* **54,** 582–590 (2013).
40. Wang, J. *et al.* Multi-atlas segmentation of subcortical brain structures via the AutoSeg software pipeline. *Front. Neuroinform.* **8,** 7 (2014).
41. Shaw, P. *et al.* Development of cortical surface area and gyrification in attention-deficit/hyperactivity disorder. *Biol. Psychiatry* **72,** 191–197 (2012).
42. Shaw, P. *et al.* Neurodevelopmental trajectories of the human cerebral cortex. *J. Neurosci.* **28,** 3586–3594 (2008).
43. 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).
44. Kim, S. H. *et al.* Adaptive prior probability and spatial temporal intensity change estimation for segmentation of the one-year-old human brain. *J. Neurosci. Methods* **212,** 43–55 (2013).
45. Gilmore, J. H. *et al.* Longitudinal development of cortical and subcortical gray matter from birth to 2 years. *Cereb. Cortex* **22,** 2478–2485 (2012).
46. Chawarska, K. *et al.* Early generalized overgrowth in boys with autism. *Arch. Gen. Psychiatry* **68,** 1021–1031 (2011).
47. WHO Multicentre Growth Reference Study Group. WHO Child Growth Standards: Growth Velocity based on Weight, Length and Head Circumference: Methods and Development. (WHO, 2009).
48. Im, K. *et al.* Brain size and cortical structure in the adult human brain. *Cereb. Cortex* **18,** 2181–2191 (2008).
49. Gilmore, J. H. *et al.* Regional gray matter growth, sexual dimorphism, and cerebral asymmetry in the neonatal brain. *J. Neurosci.* **27,** 1255–1260 (2007).
50. Benjamini, Y., Krieger, A. M. & Yekutieli, D. Adaptive linear step-up procedures that control the false discovery rate. *Biometricka* **93,** 491–507 (2006).
51. Hinton, G. E. & Salakhutdinov, R. R. Reducing the dimensionality of data with neural networks. *Science* **313,** 504–507 (2006).



Extended Data Figure 1 | Visualization of cortical regions with surface area measurements among the top 40 features contributing to the linear sparse learning classification. The cortical features produced by the deep learning approach (Fig. 3) are highly consistent with those observed using an alternative approach (linear sparse learning) shown here. Results from this alternative approach are included for comparison in Supplementary Tables 2 and 3.



Extended Data Figure 2 | Trajectories of TBV for males (left) and females (right). For illustrative purposes, we provide plots for TBV for males and females from the same sample. The longitudinal trajectories of total brain volume (TBV) from 6 to 24 months for the three groups examined are shown with males and females displayed separately. The trajectory of TBV for males among the three groups is similar to the pattern we see in the full sample (Fig. 1). The female HR-ASD group is quite small ($n = 2$), which makes the pattern of trajectory difficult to interpret. These figures support the general similarity of the findings in the combined sample and the male-only sample. Red, HR-ASD; green, HR-neg; blue, LR. TBV is shown in mm^3. The age (in months) is corrected by length (body size, in cm).

Extended Data Table 1 | Subject demographics (including tests for group differences)

	LR (N=117)		HR-neg (N=248)		HR-ASD (N=70)		P
	N	(%)	N	(%)	N	(%)	
Sex							<0.01
Female	48	(41)	106	(43)	12	(17)	
Male	69	(59)	142	(57)	58	(83)	
Race							0.14
Non-white	15	(13)	29	(12)	9	(13)	
White	102	(87)	213	(86)	61	(87)	
Not reported	0	(0)	6	(2)	0	(0)	
Family Income							0.13
Not Answered	5	(4)	20	(8)	1	(2)	
<50K	17	(15)	52	(21)	19	(27)	
50-75K	27	(23)	42	(17)	17	(24)	
75-100K	20	(17)	40	(16)	12	(17)	
100-150K	29	(25)	50	(20)	12	(17)	
>150K	19	(16)	44	(18)	9	(13)	
Mother's Highest Education							<0.01
Not Answered	0	(0)	8	(3)	0	(0)	
No College	16	(14)	75	(30)	29	(41)	
College Degree	47	(40)	108	(44)	25	(36)	
Graduate Degree	54	(46)	57	(23)	16	(23)	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	
Mother's Age at Child Birth							0.99
Years	33.2	(4.2)	33.2	(4.7)	33.3	(4.1)	
Birth Weight							0.52
Pounds (LB)	8.0	(1.0)	7.9	(1.0)	7.9	(1.3)	
Gestational Age at Birth							0.30
Weeks	39.3	(1.3)	39.1	(1.2)	38.9	(1.1)	
Age at Study Visit							
6 month visit	6.7	(0.7)	6.6	(0.7)	6.6	(0.7)	0.54
12 month visit	12.7	(0.7)	12.7	(0.6)	12.7	(0.7)	0.48
24 month visit	24.6	(0.8)	24.7	(0.9)	24.6	(0.6)	0.59
24 month Behavioral Scores							
Mullen ELC	109.7	(13.8)	101.8	(15.9)	79.3	(17.6)	0.01
Vineland ABC	105.0	(7.4)	101.0	(9.0)	88.1	(10.0)	<0.01

No significant group differences (between HR-ASD, HR-neg and LR) were observed in race/ethnicity, family income, maternal age at birth, infant birth weight, gestational age at birth, or age at visit. As expected, on the basis of the disproportionately higher rates of ASD in males, the HR-ASD group contained significantly more males than the LR group ($\chi^2_2 = 15.7, P < 0.01$). We also observed that the LR group had higher maternal education compared to the other two groups ($\chi^2_2 = 36.4, P < 0.01$). As expected, on the basis of the association between intellectual disability and ASD, the HR-ASD group had significantly lower Mullen and Vineland scores at 24 months than the other two groups. Mullen ELC, Early Learning Composite standard score; Vineland ABC, Vineland Adaptive Behaviour Composite standard score; SD, standard deviation.

Extended Data Table 2 | Group differences in developmental trajectories and cross-sectional volumes by age

		LR[a] (N=42)		HR-neg[b] (N=91)		HR-ASD[c] (N=15)		Overall Group Comparison			Pairwise Group Comparison
Measure	**Trajectory**	**Slope**	**SE**	**Slope**	**SE**	**Slope**	**SE**	**df1, df2**	**F**	**P**	
Total brain volume	1st Year	24046	1416	25329	980	27298	2399	2,289	0.72	0.49	
	2nd Year	10928	523	10277	360	13318	1060	2,289	3.83	0.02	c>(a, b)
Surface Area	1st Year	972	71	1061	49	1331	125	2,289	3.13	0.04	c>(a, b)
	2nd Year	864	38	814	26	960	75	2,289	1.93	0.15	
Cortical Thickness	1st Year	-0.024	0.003	-0.024	0.002	-0.024	0.006	2,289	0.00	0.99	
	2nd Year	-0.025	0.002	-0.021	0.001	-0.026	0.004	2,289	1.44	0.24	
	Cross-sectional Visit age	**Adjusted Group Mean**	**SE**	**Adjusted Group Mean**	**SE**	**Adjusted Group Mean**	**SE**	**df1,df2**	**F**	**P**	
Total brain volume	6 months	748890	9269	762420	6297	771403	15604	2,289	1.08	0.34	
	12 months	893167	11550	914391	7843	935189	19590	2,289	2.10	0.12	
	24 months	1024297	12475	1037714	84442	1095002	21563	2,289	4.14	0.02	c>(a, b)
Surface area	6 months	51736	541	52611	367	52779	914	2,289	1.03	0.36	
	12 months	57570	671	58979	455	60764	1144	2,289	3.32	0.04	c>a
	24 months	67933	765	68746	518	72281	1323	2,289	4.16	0.02	c>(a, b)
Cortical thickness	6 months	5.98	0.040	5.96	0.027	6.05	0.067	2,289	0.69	0.50	
	12 months	5.84	0.037	5.82	0.025	5.91	0.064	2,289	0.82	0.44	
	24 months	5.54	0.029	5.56	0.020	5.59	0.051	2,289	0.48	0.62	

Brain volume measurements for trajectory and cross-sectional analyses are in mm^3, surface area measurements are in mm^2 and cortical thickness measurements are in mm. The slope is presented as change/months. Adjusted group mean is the model estimated group mean at the specified time point. Year 1, 6–12 month period; year 2, 12–24 month period; pairwise group comparison ($P < 0.05$); the sample for the piecewise linear model included subjects with complete data at all three visits (6, 12 and 24 months).

Extended Data Table 3 | Raw means and standard deviations for TBV and surface area group comparisons showing effect size and confidence intervals

Visit Age	Group 1 N	Raw Mean (SD)	Group 2 N	Raw Mean (SD)	t	Effect Size d'	Effect size Confidence interval
Total Brain Volume (TBV) Comparisons							
6	HR-ASD N=15	800001 (69515)	HR-neg N=91	771089 (63012)	0.54	0.15	-0.40 - 0.70
6	HR-ASD N=15	800001 (69515)	LR-neg N=42	770886 (77012)	1.25	0.38	-0.23 - 0.98
6	HR-neg N=91	771089 (63012)	LR-neg N=42	770886 (77012)	1.22	0.23	-0.14 - 0.60
12	HR-ASD N=15	959305 (83486)	HR-neg N=91	922692 (69138)	0.99	0.28	-0.28 - 0.83
12	HR-ASD N=15	959305 (83486)	LR-neg N=42	917106 (85631)	1.87	0.57	-0.05 - 1.17
12	HR-neg N=91	922692 (69138)	LR-neg N=42	917106 (85631)	1.53	0.29	-0.09 - 0.66
24	HR-ASD N=15	1111639 (101094)	HR-neg N=91	1069671 (80844)	2.48	0.70	0.13 - 1.25
24	HR-ASD N=15	1111639 (101094)	LR-neg N=42	1066273 (99719)	2.86	0.88	0.24 - 1.48
24	HR-neg N=91	1069671 (80844)	LR-neg N=42	1066273 (99719)	0.9	0.17	-0.20 - 0.54
Surface Area (SA) Comparisons							
6	HR-ASD N=15	54886 (3671)	HR-neg N=91	53017 (3723)	0.17	0.05	-0.51 - 0.60
6	HR-ASD N=15	54886 (3671)	LR-neg N=42	52785 (4102)	1.0	0.31	-0.30 - 0.91
6	HR-neg N=91	53017 (3723)	LR-neg N=42	52785 (4102)	1.35	0.25	-0.12 - 0.62
12	HR-ASD N=15	61745 (4206)	HR-neg N=91	59576 (4046)	1.45	0.41	-0.15 - 0.96
12	HR-ASD N=15	61745 (4206)	LR-neg N=42	59011 (4652)	2.44	0.75	0.12 - 1.35
12	HR-neg N=91	59576 (4046)	LR-neg N=42	59011 (4652)	1.75	0.33	-0.04 - 0.70
24	HR-ASD N=15	73254 (5293)	HR-neg N=91	70757 (4642)	2.49	0.70	0.13 - 1.25
24	HR-ASD N=15	73254 (5293)	LR-neg N=42	70686 (5450)	2.87	0.88	0.24 - 1.49
24	HR-neg N=91	70757 (70757)	LR-neg N=42	70686 (5450)	0.88	0.17	-0.21 - 0.53

Brain volume measurements are in mm^3, surface area measurements are in mm^2. Visit age is shown in months. Raw mean is the group mean.

Extended Data Table 4 | Prediction model using cortical data to classify groups at 24 months

Prediction	Diagnosis HR+ N=34	Diagnosis HR- N=145		
HR+	30	7	81%	37
HR-	4	138	97%	142
	88%	95%		
	34	145		179

	A known	B known		
A test	TP	FP	PPV	TP+FP
B test	FN	TN	NPV	FN+TN
	Sensitivity	Specificity		
	TP+FN	FP+TN		(TP+FN+FN+TN)

A nonlinear prediction model included the following unbiased/unweighted information: sex, age-corrected intracranial volume and age-corrected surface area and cortical thickness measurements from 39 left and 39 right cortical hemisphere regions at 6 months and 12 months. The prediction model was evaluated using a standard tenfold cross-validation approach. Classification performance of the prediction model is at 94% overall accuracy, 88% sensitivity, 95% specificity, 81% positive predictive value and 97% negative prediction value.
TP, true positive; FP, false positive; PPV, positive predictive value; NPV, negative predictive value; diagnosis, outcome on the basis of DSM-IV-TR criteria[38].

Extended Data Table 5 | Clinical characteristics for LR subjects who met ASD criteria at 24 months

LR	DSM	Sex	Mullen ELC	Vineland ABC	ADOS SA	ADOS RBx	ADOS Sev	TBV	SA
Case 1	PDD	F	113	94	12	1	6	1034400	69839
Case 2	PDD	M	82	78	8	4	4	NA	NA
Case 3	AUT	M	59	89	12	4	7	1110231	72244

All data presented is for the visit at 24 months old. DSM, DSM-IV diagnostic criteria; PDD, pervasive developmental disorder, not otherwise specified; AUT, autism; Mullen ELC, Mullen Early Learning Composite standard score; Vineland ABC, Vineland Adaptive Behaviour Composite standard score, ADOS SA, ADOS social affective total score; ADOS RBx, ADOS repetitive behaviour total; ADOS Sev, ADOS severity score; SA, surface area; NA, no MRI data at 24 months.

Extended Data Table 6 | Subject demographics (including tests for group differences) for subjects with all 3 longitudinal visits and those with 1–2 visits completed

		HR-ASD[c]				
Measure	**Visit**	**Subjects with All 3 Visits** (N=15)		**Subjects with Partial Visits** (N=53)		
		N	**%**	**N**	**%**	**P**
Sex						.73
	Male	13	87	44	83	
	Female	2	13	9	17	
Race						.99
	White	13	87	46	87	
	Non-white	2	13	7	13	
	Not reported	0	0	0	0	
Family Income						.54
	Not answered	0	0	1	2	
	<50K	3	20	14	26	
	50K-75K	3	20	14	26	
	75K-100K	5	33	7	13	
	100K-150K	3	20	9	17	
	>150K	1	7	8	15	
Mother's Highest Education						.58
	No college	5	33	23	43	
	College Degree	5	33	19	36	
	Graduate Degree	5	33	11	21	
		Mean	**SD**	**Mean**	**SD**	**P**
Mother's Age at Childbirth (Years)		34.5	3.6	32.9	4.2	.20
Gestational Age at Birth (Weeks)		38.6	1.0	39.1	1.1	.24
Birth Weight in pounds (lbs)		7.6	1.7	7.9	1.2	.54
Age at study visit	6 months	6.7	0.8	6.6	0.7	.61
	12 months	12.9	0.7	12.6	0.6	.10
	24 months	24.6	0.6	24.6	0.6	.72

Extended Data Table 7 | Group differences in developmental level, TBV and surface area

		HR-ASD[c]				
	Age	**Subjects with All 3 Visits** (N=15)		**Subjects with Partial Visits** (N=30, 26, 26)		
		Mean	SD	Mean	SD	P
Mullen Scales of Early Learning Composite Standard Score	6	98	12	97	15	.73
	12	97	14	91	14	.22
	24	79	18	77	19	.80
Total Brain Volume	6	800001	69514	782607	64296	.41
	12	959305	83485	975594	75371	.53
	24	1111639	101093	1093702	106536	.60
Surface Area	6	54886	3671	54312	3830	.68
	12	61745	4206	61545	4893	.90
	24	73254	5293	71641	6388	.41

We further tested whether there were any group differences in developmental function (Mullen) and TBV and surface area. Groups did not differ in developmental function at any visit age, indicating that the subjects in the 3-visit and 1–2-visit subgroups are similar in their developmental capabilities. No group differences were observed for either TBV or surface area at any visit age, suggesting groups appear to have similar profiles for the brain measurements. Age is the visit age (in months), brain volume measurements are in mm^3 and surface area measurements are mm^2.

Extended Data Table 8 | Group differences in developmental trajectories and cross-sectional volumes by age for males

		LR[a] (N=28)		HR-neg[b] (N=51)		HR-ASD[c] (N=13)		Overall Group Comparison			Pairwise Group Contrast
Measure	Trajectory	Slope	SE	Slope	SE	Slope	SE	df1, df2	F	P	
Total brain volume	1st Year	21559	1683	24038	1124	18633	4522	2, 106	1.25	0.29	
	2nd Year	10398	896	9396	450	16255	3057	2, 106	2.83	0.06	c>b
Surface Area	1st Year	1020	105	1070	75	1421	153	2, 177	2.56	0.08	c>(a,b)
	2nd Year	858	47	891	37	921	80	2, 177	0.29	0.75	
Cortical Thickness	1st Year	-0.020	0.004	-0.021	0.003	-0.025	0.006	2, 177	0.19	0.83	
	2nd Year	-0.026	0.002	-0.023	0.002	-0.026	0.004	2, 177	0.41	0.67	
Cross-Sectional	Visit age	LSM	SE	LSM	SE	LSM	SE	df1,df2	F	P	
Total brain volume	6 months	718589	16333	727226	9585	762604	42125	2, 106	0.49	0.62	
	12 months	847944	20012	871452	11539	874399	52450	2, 106	0.53	0.59	
	24 months	972725	21498	984204	12307	1069460	57570	2, 106	1.24	0.29	
Surface area	6 months	53167	660	54885	473	54463	939	2, 177	2.32	0.10	
	12 months	59287	781	61307	588	62990	1170	2, 177	4.01	0.02	c>a
	24 months	69582	886	72001	667	74036	1354	2, 177	4.42	0.01	c>a
Cortical thickness	6 months	6.04	0.05	5.99	0.033	6.09	0.067	2, 177	1.10	0.33	
	12 months	5.92	0.045	5.86	0.034	5.95	0.07	2, 177	0.81	0.45	
	24 months	5.61	0.035	5.58	0.026	5.63	0.05	2, 177	0.37	0.69	

The primary analysis of brain volume trajectories included only those male and female study participants with three completed visits (6, 12 and 24 months), to best depict longitudinal trajectories over time. Separate analyses on males and females are likely to be inadequately powered owing to small subsample size (males = 13, females = 2) and therefore provide inconclusive results. With that caveat, we provide the results of our male-only analysis for the three groups for total brain volume. We do not see any group differences in the first year (6–12 months). The HR-ASD males show a pattern of TBV brain enlargement by the end of the second year, compared to the LR and HR-neg groups. Brain volume measurements for trajectory and cross-sectional analyses are in mm^3, surface area measurements are in mm^2 and cortical thickness measurements are in mm. Slope is presented as change/months. The sample for the piecewise linear model included subjects with complete data at all three visits (6, 12 and 24 months). 1st year, 6–12 month period; 2nd year, 12–24 month period; LSM, least square means; pairwise group comparison ($P < 0.05$); SE, standard error.