Phenotyping the Preterm Brain: Characterizing Individual Deviations From Normative Volumetric Development in Two Large Infant Cohorts

Ralica Dimitrova1,2, Sophie Arulkumaran1, Olivia Carney1, Andrew Chew1, Shona Falconer1, Judit Ciarrusta1,2, Thomas Wolfers3,4, Dafnis Batalle1,2, Lucilio Cordero-Grande1,5, Anthony N. Price1, Rui P.A.G. Teixeira1, Emer Hughes1, Alexia Egloff1, Jana Hutter1, Antonios Makropoulos6, Emma C. Robinson1, Andreas Schuh6, Katy Vecchiato1,2, Johannes K. Steinweg1, Russell Macleod1, Andre F. Marquand3,4, Grainne McAlonan2,7,8, Mary A. Rutherford1, Serena J. Counsell1, Stephen M. Smith9, Daniel Rueckert6, Joseph V. Hajnal1, Jonathan O’Muircheartaigh1,2,7,† and A. David Edwards1,7,†

1Centre for the Developing Brain, School of Biomedical Engineering and Imaging Sciences, King’s College London, London SE1 7EH, UK, 2Department of Forensic and Neurodevelopmental Sciences, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, London SE5 8AF, UK, 3Donders Centre for Cognitive Neuroimaging, Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen 6525EN, the Netherlands, 4Department of Cognitive Neuroscience, Radboud University Medical Centre, Nijmegen 6525EN, the Netherlands, 5Biomedical Image Technologies, ETSI Telecomunicacion, Universidad Politecnica de Madrid and CIBER-BBN, Madrid 28040, Spain, 6Biomedical Image Analysis Group, Department of Computing, Imperial College London, London SW7 2AZ, UK, 7MRC Centre for Neurodevelopmental Disorders, King’s College London, London SE1 1UL, UK, 8South London and Maudsley NHS Foundation Trust, London SE5 8AZ, UK and 9Centre for Functional MRI of the Brain (fMRIB), Wellcome Centre for Integrative Neuroimaging, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford OX3 9DU, UK

Address correspondence to Prof. A. David Edwards, Professor of Paediatrics and Neonatal Medicine, Centre for the Developing Brain, 1st floor South Wing, St. Thomas’ Hospital, London SE1 7EH, UK. Email: ad.edwards@kcl.ac.uk.

†Jonathan O’Muircheartaigh and A. David Edwards contributed equally

Abstract

The diverse cerebral consequences of preterm birth create significant challenges for understanding pathogenesis or predicting later outcome. Instead of focusing on describing effects common to the group, comparing individual infants against robust normative data offers a powerful alternative to study brain maturation. Here we used Gaussian process regression to create normative curves characterizing brain volumetric development in 274 term-born infants, modeling for...
age at scan and sex. We then compared 89 preterm infants scanned at term-equivalent age with these normative charts, relating individual deviations from typical volumetric development to perinatal risk factors and later neurocognitive scores. To test generalizability, we used a second independent dataset comprising of 253 preterm infants scanned using different acquisition parameters and scanner. We describe rapid, nonuniform brain growth during the neonatal period. In both preterm cohorts, cerebral atypicalities were widespread, often multiple, and varied highly between individuals. Deviations from normative development were associated with respiratory support, nutrition, birth weight, and later neurocognition, demonstrating their clinical relevance. Group-level understanding of the preterm brain disguises a large degree of individual differences. We provide a method and normative dataset that offer a more precise characterization of the cerebral consequences of preterm birth by profiling the individual neonatal brain.

Key words: early brain development, heterogeneity, normative modeling, preterm birth, volumetric MRI

Introduction

Preterm birth (or birth before 37 weeks gestational age, GA) affects approximately 10% of pregnancies worldwide (Chawanpaiboon et al. 2019) and is a significant risk predisposing to atypical brain development and lifelong cognitive difficulties including a higher incidence of neurodevelopmental and psychiatric disorders (Nosarti et al. 2012; Agrawal et al. 2018; Thompson et al. 2020). Although early brain correlates of preterm birth have been identified at a group level (Volpe 2019), this vulnerable population is highly heterogeneous, with individuals following diverse clinical and neurocognitive trajectories (Sled and Nossin-Manor 2013; Dimitrova et al. 2020). Indeed, the assumption that premature has a homogenous effect on brain development might help account for the relatively poor predictive power of neonatal magnetic resonance imaging (MRI) for later outcome (de Bruine et al. 2011; Edwards et al. 2018). To better understand brain development, provide accurate prognosis of later functionality, and study the effect of clinical risks and interventions, it is important to provide an individualized assessment of cerebral maturation (O’Muircheartaigh et al. 2020). Comparing individuals against robust normative data avoids the requirement to define quasi-homogenous groups in a search for effects common to the group and offers a powerful alternative to investigate brain development with high sensitivity to pathology at an individual infant level (Towgood et al. 2009; Holland et al. 2014; O’Muircheartaigh et al. 2020).

In this study, we used Gaussian process regression (GPR) to create normative charts of typical volumetric development using a large sample of healthy term-born infants scanned across-sectionally within the first month of life. Analogue to the widely employed pediatric height and weight growth charts, this technique allows the local imaging features of individuals to be referred to typical variation while simultaneously accounting for variables such as age and sex (Marquand et al. 2016, 2019). Having established normative values for brain growth of 14 brain regions, we aimed to 1) quantify deviations from typical development in individual preterm infants, 2) investigate the heterogeneity of these deviations, and 3) examine the association between individual deviations, perinatal clinical factors, and later neurocognitive abilities. To test generalizability, we used a second large independent preterm dataset acquired on a different magnetic resonance (MR) scanner using different imaging parameters.

Materials and Methods

Participants

This study utilized data from two cohorts. A total of 363 (89 preterm) infants recruited for the developing Human Connectome Project (dHCP; http://developingconnectome.org/) were scanned at term-equivalent age (TEA, 37–45 weeks postmenstrual age, PMA) during natural unsedated sleep at the Evelina London Children’s Hospital between 2015 and 2019. The second cohort comprised of further 253 preterm infants born before 33 weeks GA that underwent MRI between 37- and 45-week PMA at the neonatal intensive care unit (NICU) in Hammersmith Hospital between 2010 and 2013 for the Evaluation of Preterm Imaging (EPrime) study. Detailed description of these studies and the scanning procedure used has been previously reported (dHCP, Hughes et al. 2017; EPrime, Edwards et al. 2018). All MRI images were examined by a neonatal neuroradiologist. Exclusion criteria for term-born infants are described in Supplementary Methods. There were no exclusion criteria for the preterm infants, except for major congenital malformations and data included infants from nonsingleton pregnancies. Both studies were approved by the National Research Ethics Committee (dHCP, REC: 14/Lo/1169; EPrime, REC: 09/H0707/98). Informed written consent was given by parents prior to scanning.

MRI Acquisition and Preprocessing

MRI data for the dHCP were collected on a Philips Achieva 3T (Philips Medical Systems) using a dedicated 32-channel neonatal head coil (Hughes et al. 2017). T2-weighted scans were acquired with TR/TE of 12 s/156 ms, SENSE=2.11/2.58 (axial/sagittal), 0.8 × 0.8 mm in-plane resolution, 1.6-mm slice thickness (0.8-mm overlap). Images were motion corrected and super-resolution reconstructed resulting in 0.5-mm isotropic resolution (Cordero-Grande et al. 2018). MRI data collected for EPrime were acquired on a Philips 3T system using an 8-channel phased array head coil. T2-weighted turbo spin echo was acquired with TR/TE of 8670/160 ms, in-plane resolution 0.86 × 0.86 mm, 2-mm slice thickness (1-mm overlap).

Both datasets were preprocessed using the dHCP structural pipeline (Makropoulos et al. 2018). In brief, motion-corrected, reconstructed T2-weighted images were corrected for bias-field inhomogeneities, brain extracted, and segmented. Tissue labels included cerebrospinal fluid (CSF), white matter (WM), cortical...
gray matter (cGM), deep gray matter (dGM), ventricles (including the cisterna magna), a transient fluid-filled cavity located in the midline of the brain between the left and right anterior horns of the lateral ventricles, which if present, usually disappears during the neonatal period, cerebellum, brainstem, hippocampus, and amygdala. dGM was further parcellated into left/right caudate, lentiform, and thalamus. Total tissue volume (TTV) incorporated all brain GM and WM volumes; total brain volume (TBV) included TTV and ventricles; intracranial volume (ICV) included TBV and CSF (Table 1). Given the high correlation between TTV and TBV ($\rho = 0.98$), we reported only TTV. Due to their size and lower tissue contrast in the neonatal brain, the amygdala and the hippocampus are prone to segmentation errors and higher partial volume estimates, especially in the EPrime dataset, where the image resolution was lower (1 mm) compared with dHCP (0.5 mm). Therefore, these structures were excluded from the present analyses. The quality of the preprocessing was visually evaluated using a scoring system detailed elsewhere (Makropoulos et al. 2018) to ensure no images severely affected by motion or with poor segmentation were included (Supplementary Methods and Supplementary Fig. 1).

We estimated regional volumes in absolute (cm$^3$) and relative (%) values. Relative volumes were calculated as the proportion of each tissue volume from TTV, the ventricles from TBV and CSF from ICV (Table 1). To capture the effect of preterm birth, we used relative volumes to 1) ensure results are not driven by extreme individual differences in nonbrain ICV, often seen in preterm infants, 2) partially alleviate differences in data acquisition.

Modeling Volumetric Development Using GPR

To characterize neonatal volumetric development, we used GPR, a Bayesian nonparametric regression, implemented in GPy (https://sheffieldml.github.io/GPy/). GPR simultaneously provides point estimates and measures of predictive confidence for every observation representing the distance of each individual observation from the normative mean at that point on the “curve” accounting for modeled covariates (Marquand et al. 2016).

We first trained a GPR model to describe typical development in the term-born dataset (274 infants) using PMA at scan and sex to predict 14 brain structures separately. Regions included ICV, TTV, cGM, WM, cerebellum, brainstem, CSF, ventricles, and left/right caudate, lentiform, and thalamus. Model accuracy was tested under 5-fold cross-validation, with each fold stratified to cover the whole PMA range (37–45 weeks). The relationship between the volume outputs and model predictors was estimated with a sum of radial basis function, linear, and white noise covariance kernels. Model hyperparameters were optimized using log marginal likelihood. Prediction performance was evaluated using the mean absolute error (MAE) between the predicted and the observed value derived from the 5-fold cross-validation.

To assess the effects of preterm birth, we retrained the model on the entire term-born dataset and applied the model to 89 dHCP preterm infants scanned at TEA. To assess generalizability, we applied the same model to 253 preterm infants from the EPrime study. A Z-score was derived for every infant by estimating the difference between the model prediction and the observed value normalized by the model uncertainty (the square root of the predicted variance). To quantify extreme deviations, prior to analyses, we chose a threshold of $|Z| > 2.6$ (corresponding to $P < 0.005$) following the convention adopted in previous GPR analyses modeling adult brain development (Wolfers et al. 2020). We examined the proportion of infants with volumes lying $>2.6$ standard deviations (SD) above/below the model mean (indicating the top/bottom 0.5% of the typical group values, hereafter described as extreme positive or negative deviations, respectively).

To quantify the effect of image spatial resolution differences between the dHCP and EPrime, we first downsampled the dHCP data to 1-mm isotropic resolution using FSL flirt (-applyisofxm, spline interpolation), reran the tissue segmentation on the new data to 1-mm isotropic resolution, respectively.

We tested the association between deviations from normative development (in relative volumes) and recognized perinatal clinical risks (Boardman and Counsell 2020), including GA at birth, birth weight Z-score, total days receiving mechanical ventilation, continuous positive airway pressure (CPAP), and total parenteral nutrition (TPN, available only for EPrime). Birth weight Z-scores were calculated using the population data from the uk90 growth charts implemented in sitar R package (Cole et al. 2010). Oxygen/respiratory support and nutrition were administered at the NICU. These data were obtained from electronic hospital records and days were counted if the infant spent any part of the day on ventilation, CPAP, or TPN, with higher number of days indicative of poorer health. Bayley III Scales of Infant Development (BSID-III) (Bayley 2006) assessment was carried out by trained developmental pediatricians/psychologists at 18 months for the dHCP and at 20 months for EPrime (corrected age). We used the

### Table 1: Brain regions of interest, the structures they include, and what global brain measures they are taken as a proportion from when calculating relative brain volumes

<table>
<thead>
<tr>
<th>Brain regions</th>
<th>Brain structures included</th>
<th>Relative volume (% from)</th>
</tr>
</thead>
<tbody>
<tr>
<td>cGM</td>
<td>—</td>
<td>TTV</td>
</tr>
<tr>
<td>WM</td>
<td>—</td>
<td>TTV</td>
</tr>
<tr>
<td>Cerebellum</td>
<td>—</td>
<td>TTV</td>
</tr>
<tr>
<td>Brainstem</td>
<td>—</td>
<td>TTV</td>
</tr>
<tr>
<td>CSF</td>
<td>—</td>
<td>ICV</td>
</tr>
<tr>
<td>Ventricles</td>
<td>Lateral ventricles + cisterna magna</td>
<td>TBV</td>
</tr>
<tr>
<td>Caudate (left/right)</td>
<td>—</td>
<td>TTV</td>
</tr>
<tr>
<td>Lentiform (left/right)</td>
<td>Pallidum + putamen</td>
<td>TTV</td>
</tr>
<tr>
<td>Thalamus (left/right)</td>
<td>—</td>
<td>TTV</td>
</tr>
<tr>
<td>TTV</td>
<td>All brain GM + WM tissue</td>
<td>—</td>
</tr>
<tr>
<td>TBV</td>
<td>All brain GM + WM tissue + ventricles</td>
<td>—</td>
</tr>
<tr>
<td>ICV</td>
<td>All brain GM + WM tissue + ventricles + CSF</td>
<td>—</td>
</tr>
</tbody>
</table>
Table 2: Perinatal, demographic, and neurocognitive characteristics of the study sample

<table>
<thead>
<tr>
<th></th>
<th>Term (dHCP) n = 274</th>
<th>Preterm (dHCP) n = 89</th>
<th>Preterm (EPrime) n = 253</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA at birth (weeks), median (IQR)</td>
<td>40.3 (39.1–41)</td>
<td>31.6 (28.7–34)</td>
<td>30.3 (28–31.6)</td>
</tr>
<tr>
<td>PMA at scan (weeks), median (IQR)</td>
<td>40.7 (39.4–41.4)</td>
<td>41.3 (40.1–42.7)</td>
<td>42.4 (41–43.4)</td>
</tr>
<tr>
<td>Female, No. (%)</td>
<td>129 (47%)</td>
<td>43 (46%)</td>
<td>115 (45%)</td>
</tr>
<tr>
<td>HC at scan (cm), median (IQR)</td>
<td>35 (33.5–36)</td>
<td>35 (33.7–36.1)</td>
<td>36.3 (35–37.2)</td>
</tr>
<tr>
<td>Weight at birth (kg), median (IQR)</td>
<td>3.4 (3–3.6)</td>
<td>3.1 (2.7–3.7)</td>
<td>3.4 (2.6–3.8)</td>
</tr>
<tr>
<td>Weight at birth (kg), median (IQR)</td>
<td>3.4 (3–3.7)</td>
<td>1.6 (1–2)</td>
<td>1.3 (1–1.99)</td>
</tr>
<tr>
<td>Nonsingleton, No. infants, (%)</td>
<td>25 (28%)</td>
<td>6 (25%)</td>
<td>62 (25%)</td>
</tr>
<tr>
<td>Days of ventilation, No. infants, (%)</td>
<td>45 (51%)</td>
<td>116 (46%)</td>
<td></td>
</tr>
<tr>
<td>median (IQR)</td>
<td>6 (2.5–20)</td>
<td>2 (1–5)</td>
<td></td>
</tr>
<tr>
<td>Days of TPN, No. infants, (%)</td>
<td>63 (71%)</td>
<td>204 (81%)</td>
<td></td>
</tr>
<tr>
<td>median (IQR)</td>
<td>7 (1–30)</td>
<td>11 (4–30)</td>
<td></td>
</tr>
<tr>
<td>Days of CPAP, No. infants, (%)</td>
<td>34 (12%)</td>
<td>22 (25%)</td>
<td>43 (17%)</td>
</tr>
<tr>
<td>median (IQR)</td>
<td>27 (30%)</td>
<td>52 (21%)</td>
<td></td>
</tr>
<tr>
<td>IUGR, No. infants, (%)</td>
<td>34 (12%)</td>
<td>5 (6%)</td>
<td>16 (6%)</td>
</tr>
<tr>
<td>Cerebellar hemorrhages, No infants, (%)</td>
<td>6 (2%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PVL, No. infants, (%)</td>
<td>222 (81%)</td>
<td>68 (76%)</td>
<td>237 (94%)</td>
</tr>
<tr>
<td>Age (months), median (IQR)</td>
<td>18.4 (18–18.7)</td>
<td>18.5 (18–19.1)</td>
<td>20.1 (20–20.6)</td>
</tr>
<tr>
<td>BSID-III Motor, median (IQR)</td>
<td>103 (97–107)</td>
<td>100 (94–107)</td>
<td>95 (85–100)</td>
</tr>
<tr>
<td>BSID-III Language, median (IQR)</td>
<td>97 (88–106)</td>
<td>97 (86–106)</td>
<td>91 (79–103)</td>
</tr>
<tr>
<td>BSID-III Cognition, median (IQR)</td>
<td>100 (95–105)</td>
<td>100 (90–105)</td>
<td>95 (88–103)</td>
</tr>
</tbody>
</table>

Notes: GA – Gestational age; IQR – interquartile range; PMA – Postmenstrual age; HC – head circumference; TPN – total parenteral nutrition; CPAP – continuous positive airway pressure; IURG – intrauterine growth restriction; PWMLs – punctate white matter lesions; HPI – Haemorrhagic Parenchymal Infarction; PVL – Periventricular leukomalacia; BSID-III – Bayley Scales of Infant Development III; Missing data: Birth HC – 11 preterm infants; birth weight–1 preterm. TPN data were not available for dHCP preterm infants. * median (IQR) calculated only for infants on ventilation/CPAP/TPN.

Results

The perinatal, demographic and neurocognitive characteristics are presented in Table 2. EPrime infants were born earlier (p < 0.05, d = 0.28) and had lower birth weight (p < 0.05, d = 0.19) compared with dHCP preterm infants. On average, they had poorer motor (p < 0.05, d = 0.22) skills at follow-up (language p = 0.1; cognition p = 0.12). There were no differences in days on CPAP (p = 0.57) but dHCP preterm infants required mechanical ventilation for longer (p < 0.05, d = 0.16). The 2 preterm cohorts did not differ in punctate WM lesions (PWMLs, p = 0.09) or cerebellar hemorrhages (p = 0.82) incidence, nor in proportion of infants with intraventricular growth restriction (IUGR, p = 0.15); yet hemorrhagic parenchymal infarction (HPI) and periventricular leukomalacia (PVL) were observed only in the EPrime cohort (Table 2).

Typical Volumetric Development in Term-Born Infants During the Neonatal Period

We found an increase in all absolute volumes except the ventricles, where no change was detected (Fig. 1A, Supplementary Fig. 2). The increase was greatest in cGM (10.4% per week [pw]) and cerebellum (9.9% pw) compared with ICV (6.1% pw), TTV (6% pw), and CSF (7% pw). Subcortical structures increased between 4% and 6% pw (caudate L:4.1%, R:4.3%; lentiform L:6.6%, R:5.3%; thalamus L:4.1%, R:4.7% pw) with smaller increases in brainstem (3.9% pw) and WM (2% pw).

The greatest changes in relative volumes were observed in cGM and WM (Fig. 1B, Supplementary Fig. 2). cGM represented 36% of TTV at 37-week PMA and increased to 44% at 44-week PMA, whereas the relative WM volume decreased from 48% to 38% of TTV. The relative cerebellar volume increased from 6% to 7%. There was an increase in relative lentiform volume, a subtle decrease in caudate, and no change in thalamus. We observed a slight increase in CSF proportion of ICV and a steady decrease in the proportion that ventricles contributed to TBV. MAE for all models is shown in Supplementary Table 1.

Image Resolution and Volumetric Development

Overall, the majority of observations in both dHCP and EPrime preterm samples fit within 2.6 SD of the term-born model, indicating good agreement between the two studies (Fig. 2A). Differences were most profound in fluid-filled structures, likely attributable to partial voluming of high T2-signal CSF. In agreement, when compared with the models built using the original dHCP resolution of 0.5 mm, the matched 1-mm resolution models showed a mean shift (increase) for the CSF

Data Availability

Normative term-born data and the GPR code used in this study are freely available on GitHub (https://github.com/ralidimitrova). All imaging data collected for the dHCP will be publicly available in early 2021 at http://developingconnectome.org/.

Notes: GA – Gestational age; IQR – interquartile range; PMA – Postmenstrual age; HC – head circumference; TPN – total parenteral nutrition; CPAP – continuous positive airway pressure; IURG – intrauterine growth restriction; PWMLs – punctate white matter lesions; HPI – Haemorrhagic Parenchymal Infarction; PVL – Periventricular leukomalacia; BSID-III – Bayley Scales of Infant Development III; Missing data: Birth HC – 11 preterm infants; birth weight–1 preterm. TPN data were not available for dHCP preterm infants. * median (IQR) calculated only for infants on ventilation/CPAP/TPN.

comparisons corrected.

**Results**

The perinatal, demographic and neurocognitive characteristics are presented in Table 2. EPrime infants were born earlier (p < 0.05, d = 0.28) and had lower birth weight (p < 0.05, d = 0.19) compared with dHCP preterm infants. On average, they had poorer motor (p < 0.05, d = 0.22) skills at follow-up (language p = 0.1; cognition p = 0.12). There were no differences in days on CPAP (p = 0.57) but dHCP preterm infants required mechanical ventilation for longer (p < 0.05, d = 0.16). The 2 preterm cohorts did not differ in punctate WM lesions (PWMLs, p = 0.09) or cerebellar hemorrhages (p = 0.82) incidence, nor in proportion of infants with intraventricular growth restriction (IUGR, p = 0.15); yet hemorrhagic parenchymal infarction (HPI) and periventricular leukomalacia (PVL) were observed only in the EPrime cohort (Table 2).

**Typical Volumetric Development in Term-Born Infants During the Neonatal Period**

We found an increase in all absolute volumes except the ventricles, where no change was detected (Fig. 1A, Supplementary Fig. 2). The increase was greatest in cGM (10.4% per week [pw]) and cerebellum (9.9% pw) compared with ICV (6.1% pw), TTV (6% pw), and CSF (7% pw). Subcortical structures increased between 4% and 6% pw (caudate L:4.1%, R:4.3%; lentiform L:6.6%, R:5.3%; thalamus L:4.1%, R:4.7% pw) with smaller increases in brainstem (3.9% pw) and WM (2% pw).

The greatest changes in relative volumes were observed in cGM and WM (Fig. 1B, Supplementary Fig. 2). cGM represented 36% of TTV at 37-week PMA and increased to 44% at 44-week PMA, whereas the relative WM volume decreased from 48% to 38% of TTV. The relative cerebellar volume increased from 6% to 7%. There was an increase in relative lentiform volume, a subtle decrease in caudate, and no change in thalamus. We observed a slight increase in CSF proportion of ICV and a steady decrease in the proportion that ventricles contributed to TBV. MAE for all models is shown in Supplementary Table 1.

**Image Resolution and Volumetric Development**

Overall, the majority of observations in both dHCP and EPrime preterm samples fit within 2.6 SD of the term-born model, indicating good agreement between the two studies (Fig. 2A). Differences were most profound in fluid-filled structures, likely attributable to partial voluming of high T2-signal CSF. In agreement, when compared with the models built using the original dHCP resolution of 0.5 mm, the matched 1-mm resolution models showed a mean shift (increase) for the CSF.
Figure 1. Normative modeling of volumetric development during the neonatal period. The model means for both female and male term infants are shown in purple and blue respectively, together with ±1, ±2, and ±3 SDs from the model means for absolute (A) and relative (B) volumes (tissue volumes represented as a proportion from TTV, ventricles from TBV, and CSF from ICV). Normative charts are shown only for right dGM structures (left structures are shown in Supplementary Fig. 2).

and ventricular volumes (Fig. 2B; Supplementary Figs 3 and 4). As a result, when using the lower resolution normative charts, the proportion of extreme positive deviations in EPrime decrease from 53% to 29% for CSF and from 44% to 32% for ventricles (Fig. 2C). All infants who showed extreme deviations in the matched 1-mm resolution showed extreme deviations in the original 0.5-mm resolution. Changes in the proportion of extreme deviations associated with image resolution for the rest
of the structures were more subtle. Unless stated otherwise, data are presented for the 0.5-mm resolution models. The overall proportion of extreme deviations in the term-born sample was very low in all brain structures, with no more than 2% of the sample with $Z$-scores $>2.6$ in the original 0.5-mm resolution.

**Infants With Reduced TTV Suffered More Extreme Prematurity**

Six dHCP preterm infants (7%) and 8 EPrime (3%) infants showed extreme negative deviations in TTV (Fig. 2C), all of which (except 1 EPrime infant) were born at GA $<30$ weeks and weighed less than 1 kg at birth. Significantly reduced TTV was accompanied by enlarged CSF and ventricles (dHCP: 3/6; EPrime: 7/8). Four infants (dHCP: 1/6; EPrime 3/8) also had associated reduction in cerebellar volumes. A further 3 preterm infants (dHCP: 1/6; EPrime 2/8) had associated reduction in the thalami, bilaterally. All infants required oxygen support after birth. Infants who had TTV 2.6 SD above the model mean (dHCP: 4 [5%]; EPrime: 18 [7%]) were born GA $>30$ weeks and had no incidental findings, other than PWMLs, and short need for oxygen support and TPN after birth.

**Infants With Reduced Thalamic Volume Also Had PWMLs**

In the dHCP preterm sample, all 8 infants with extreme negative thalamic deviations had PWMLs, 7 of 8 had multiple lesions. Four out of these 7 infants had lesions involving the corticospinal tract (Fig. 3). Seven out of the 8 infants were on CPAP, but none of them for a long period of time (5 infants $>4$ days; 1 infant 11 days; 1 infant 18 days), and all 7 did not require ventilation. None of these infants had a birthweight of less than 1 kg. In EPrime, 17 infants had bilateral reduced thalamic volume and 10 unilateral extreme deviations (with structure in the other hemisphere close to but not reaching $Z < -2.6$). In total, 78% of these infants had PWMLs compared with 16% incidence in the rest of the sample. Overall, across the whole cohort, infants with PWMLs had significantly reduced left ($d = 0.56$) and right ($d = 0.53$) thalamic volumes, compared with infants without (both $p < 0.05$). In EPrime, infants with reduced thalamic volumes often had CSF or ventricular volumes significantly bigger than the normative values for their age/sex. In 5 infants, this was associated with PVL and in a further 2, with HPI.

**Atypical Ventricular Development in Preterm Infants: Frequent But Highly Heterogeneous**

 Widening of the fluid-filled structures was the most frequently observed deviation from normative development in both cohorts. In the dHCP 29% and 17% of the preterm infants showed extreme deviations in ventricular and CSF volumes, respectively. This number was higher in EPrime where increased ventricles and CSF were seen in 44% and 53% of infants with the original 0.5-mm dHCP resolution and in 29% and 32% with the downsampled 1-mm resolution. Figure 4 shows the most extreme cases where infants’ ventricles were 10 SD above the model mean. These extreme deviations in ventricular volume were associated with overt focal brain injuries including HPI (infants 1, 2, 4, 6) and PVL (infants 5, 7). In all of these infants we also observed significant negative deviations in TTV or thalamus and increased CSF. These infants performed poorly at follow-up (Fig. 4).

**Association Between Perinatal Risks and Deviations From Normative Development**

In the dHCP cohort, decreased GA at birth related to reduced TTV ($\rho = 0.45$) and increased proportion of CSF ($\rho = -0.44$), whereas in EPrime to reduced TTV ($\rho = 0.27$) and increased relative ventricular volumes ($\rho = -0.26$) (all $p_{corr} < 0.05$) (Fig. 5; Supplementary Table 2). In both samples, greater birth weight $Z$-score was related to bigger ICV (dHCP: $\rho = 0.41$, EPrime: $\rho = 0.36$) and TTV ($\rho = 0.40$, $\rho = 0.37$) at TEA and in EPrime alone, to reduced relative brainstem ($\rho = -0.26$) and bilateral thalamic volumes (right: $\rho = -0.28$, left: $\rho = -0.25$) (all $p_{corr} < 0.05$; Supplementary Table 3).

In the dHCP preterm sample, longer requirement for CPAP related to smaller TTV ($\rho = -0.47$) and ICV ($\rho = -0.37$) as well as to increased proportion CSF ($\rho = 0.39$) and ventricles ($\rho = 0.35$). Longer need for mechanical ventilation was associated with reduced TTV ($\rho = -0.44$) and relative left caudate volume ($\rho = -0.34$) as well as with increased relative CSF ($\rho = 0.39$) and ventricular volumes ($\rho = 0.42$) (all $p_{corr} < 0.05$). Consistent with this, in EPrime, longer requirement for both CPAP and mechanical ventilation were related to reduced TTV ($\rho = -0.30$, $\rho = -0.38$, respectively) and increased relative ventricular volume (both at $\rho = 0.22$) (all $p_{corr} < 0.05$). Increased number of days requiring TPN were related to reduced TTV and ICV ($\rho = -0.43$, $\rho = -0.35$, $p_{corr} < 0.05$) and increased relative ventricular volume ($\rho = 0.25$, $p_{corr} < 0.05$; Supplementary Table 4).

**Association Between Deviations From Normative Development and Neurocognitive Outcome**

In the dHCP preterm sample, increased ventricular volume was related to poorer motor ($\rho = -0.40$) and language ($\rho = -0.37$) scores at 18 months (all $p_{corr} < 0.05$). In EPrime, increased CSF ($\rho = -0.22$) and cGM ($\rho = -0.21$) and reduced WM ($\rho = 0.20$) were associated with poorer language abilities, whereas reduced TTV ($\rho = 0.20$), increased CSF ($\rho = -0.25$), and cGM ($\rho = -0.20$) were linked to poorer cognitive performance (all $p_{corr} < 0.05$). While most of these associations were of similar effect size in the dHCP, they did not survive multiple comparison correction (Supplementary Table 5).

**Discussion**

The diverse cerebral consequences of preterm birth create significant challenges for understanding pathogenesis or predicting later neurocognitive outcomes. Focusing on individuals and their unique cerebral development can offer new insights. In this study, we first characterized normative volumetric development during the neonatal period, to then describe the effect of preterm birth at an individual infant level. We showed deviations from the normative curves consistent with previous studies but with marked variability among individuals. These individual deviations are associated with perinatal risks and later neurocognition.

We previously demonstrated that GPR could be used to detect subtle WM injury with high sensitivity (O’Muircheartaigh et al. 2020) and to characterize the heterogeneous consequences of preterm birth on the developing brain microstructure (Dimitrova et al. 2020). The present application of GPR to volumetric data offers more straightforward clinical translation. GPR provides normative charts describing typical volumetric development and can detect and quantify atypical maturation in individual
Figure 2. Characterizing the effects of preterm birth on the developing brain. (A) Deviations from normative volumetric development in preterm infants. Observations for individual preterm infants from both dHCP and EPrime cohorts are shown with model means for both female and male term-born infants together with ±1, ±2, and ±3 SDs. ICV, TBV, and TTV are in cm³; cGM, WM, cerebellum, brainstem, and subcortical structures shown as a proportion of TTV, CSF as a proportion of ICV and ventricles as a proportion of TTV. Horizontal lines show \( Z > 2.6 \), the threshold used to define extreme deviations. The normative curves for the ventricles show data within 10 SD from the mean, full range is shown in Fig. 4 and discussed below. (B) Mean differences in fluid filled structures between GPR models build using 0.5- and 1-mm dHCP imaging resolution. (C) Proportion (%) of extreme deviations from the normative model in preterm infants. Extreme negative deviations (\( Z < -2.6 \)) are depicted in blue, whereas extreme positive deviations (\( Z > 2.6 \)) are shown in orange.
Figure 3. Extreme negative deviations in thalamic volume were often accompanied by PWML in the preterm brain. Depicted are four infants (A–D) with bilateral thalamic volumes significantly below the model mean. Thalamic segmentation (dark blue) is overlaid onto the T2-weighted images. T1-weighted images are shown with and without the manual outlined PWMLs (light blue). Note T1-weighted images were not used in the preprocessing but are shown here due to better contrast for detecting PWMLs.

infants (Ou et al. 2017). The GPR approach generalized to a cohort of infants with MRI data collected on a different MR scanner with different acquisition parameters. In the future, this method could be integrated into automatic tools that complement radiological decisions regarding infant development (Duerden and Thompson 2020). Our method and normative dataset are freely available for researchers to use for understanding pathogenesis, trialing interventions, and defining neurocognitive prognosis for vulnerable preterm infants.

We quantified rapid postnatal brain growth consistent with previous imaging and postmortem studies describing change in the size, organization, and complexity of the brain during the perinatal period (Huttenlocher and Dabholkar 1997; Hüppi et al. 1998; Knickmeyer et al. 2008; Kersbergen et al. 2016; Makropoulos et al. 2016). This dramatic growth is a sum product of a number of heterochronous developmental processes that take place in the developing brain including synaptogenesis, dendritic arborization, and early stages of myelination (Huttenlocher 1990; Huttenlocher and Dabholkar 1997; Travis et al. 2005; Petanjek et al. 2008; Lebenberg et al. 2019). Abrupt preterm extraterine exposure represents a significant stressor to these events and may lead to widespread deviations from the normative trajectories in any or many of these processes as seen in pathology and preclinical models (Elovitz and Mrinalini 2004; Burd et al. 2012; Volpe 2019) associated with atypical trajectory of brain growth and a wide range of neurodevelopmental consequences (Inder et al. 2005; Bora et al. 2014; Ball et al. 2017; Gui et al. 2019). However, these alterations are not a result of loss of intrauterine environment alone but are a product of the cumulative effects of clinical and genetic factors creating individualized circumstances for every infant. GPR applied to a large normative dataset offers a powerful approach to study how preterm birth shapes the brain at an individual infant level and offers the means to capture important differences in single infants that may be missed by analysis of the means/medians of quasi-homogenous groups, which “averages-out” personal effects.

By quantifying this interindividual variability, our analysis clarified the relationship between reduced global brain growth and preterm birth. Many but not all studies show group-level differences in TTV between preterm and term-born infants (Boardman et al. 2007). We report a subset of infants in both preterm cohorts that deviated significantly from normative brain volumes. These infants were born very early, very small and had...
Figure 4. Capturing heterogeneity and extreme deviations in ventricular development in the preterm brain at TEA. (A) Normative curves are shown for both female and male infants (in upper right corner curves excluding the outliers, also shown in Fig. 2A). The figure also depicts the T1-weighted images for infants with ventricular volume lying 10 SD above the mean, separate for females (top) and males (bottom), together with their neurocognitive scores (M—motor, C—cognitive, L—language). Ventricular development in EPrime preterm infants is highly heterogeneous both in shape and size as illustrated in (B) showing ventricular volumes of various Z-scores.

The period encompassing mid gestation and the last trimester of pregnancy is a critical phase for the development and establishment of the thalamocortical network (Kostovic et al. 2014). During this short period, there are dynamic changes in thalamocortical efferent fiber organization and cortical lamination, including rapid axonal growth and the dissolution of the subplate (Kostovic and Rakic 1990; Vasung et al. 2011; Kostovic et al. 2014). This makes the thalamus and connecting WM projections particularly vulnerable to injury as a result of preterm birth (Boardman et al. 2006; Ball et al. 2015) with studies suggesting abnormal development may persist beyond TEA (Lin et al. 2001). We reported a subset of preterm infants with thalamic volumes significantly below the model mean (Z < -2.6). These infants had a high load of PWMls, and 5 of the EPrime infants had PVL, supporting previous findings of a close...
Figure 5. Association between degree of prematurity and deviations from normative brain development. In the dHCP preterm sample, increased degree of prematurity (lower GA at birth) was related to reduced TTV and increased CSF. In the EPrime sample, increased degree of prematurity was associated with reduced TTV and increased ventricular volume. Individual preterm observations are plotted against the normative model mean for female (purple) and male (blue) term infants. The plots also show ±1, ±2, and ±3 SDs from the normative means together with lines indicating $Z > 2.6$, the threshold used to define extreme deviations. Ventricular data are shown only for infants with volume ± 10 SDs from the model mean.

The link between thalamic development and WM abnormalities, including a previous group analysis of the EPrime dataset (Boardman et al. 2006; Pierson et al. 2007; Ligam et al. 2009; Volpe 2009; Ball et al. 2015; Wisnowski et al. 2015; Tusor et al. 2017). The exact mechanisms that underlie reduced thalamic growth, possibly including neuronal loss and/or atypical developmental trajectory triggered by preterm extrauterine exposure, however, remain elusive (Volpe 2009).

Compared with the dHCP preterm cohort, the EPrime study comprised extremely preterm infants, which were sicker during clinical care, had overall poorer motor outcomes, and were scanned using different acquisition parameters. These factors in combination likely underlie some of the differences in associations between extreme deviations and later neurocognitive scores observed between the 2 datasets. The lower spatial resolution in EPrime in particular, contributed to the mean shift (increase) in CSF and ventricular volumes observed in the EPrime. With all this in mind, it was reassuring that deviations in brain development and their association with perinatal risks found in the dHCP broadly replicated in EPrime, indicating good generalization of the model to independent data collected on a different MRI scanner. We chose to use volumetric measures that are easy to calculate in research studies or routine clinical examinations. This could offer a direct clinical application, although given the regional heterochrony of early brain development (Lebenberg et al. 2019), future work should focus on more finely parcellated regions or more sophisticated MRI-derived features, including cortical thickness and surface area. We reported an association between deviations from normative brain development at TEA and behavior at 18–20 months. An important step for future research is to investigate whether these early brain deviations persist in later life and are predictive of childhood and later neurodevelopment (Boardman et al. 2020; George et al. 2020).

The argument that every brain is different is not novel, and the expectation that the effects of preterm birth are homogeneous and exactly alike in every infant is equally untenable. Individualized methodologies have been successfully applied in other fields (e.g., neuropsychology, Towgood et al. 2009; aging, Ziegler et al. 2014) and hold significant promise for the preterm infant. Although a group-mean difference is detectable using the conventional case–control approach, the significant heterogeneity would not be captured and effects of clinical significance to individual infants would be averaged out (Sled and Nossin-Manor 2013). Additionally, visually subtle effects may have prognostic significance when combined with other deviations from normative brain growth, for example reduced thalamic volume, and further analytic power may be gained by including covariates in the GPR model.

In summary, our approach offers a readily interpretable, generalizable, and more precise understanding of the cerebral consequences of preterm birth by focusing on the individual rather than the group average atypicality and in future might improve the predictive power of neuroimaging.

Supplementary Material
Supplementary material can be found at Cerebral Cortex online.

Funding
The dHCP project was funded by the European Research Council (ERC) under the European Union Seventh Framework Programme (FR/2007-2013)/ERC grant agreement no. 319456.
The EPrime study was funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research Programme (grant reference no. RP-PG-0707-10154). The results leading to this publication have received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 777394 for the project AIMS-2-TRIALS. This Joint Undertaking receives support from the European Union’s Horizon 2020 research and innovation programme and EFPIA and AUTISM SPEAKS, Autistica, SFARI. The authors acknowledge infrastructure support from the National Institute for Health Research (NIHR) Biomedical Research Centre at Guy’s and St Thomas’ Hospitals NHS Foundation Trust. The study was supported in part by the Wellcome Engineering and Physical Sciences Research Council Centre for Medical Engineering at King's College London (grant WT 203148/Z/16/2) and the Medical Research Council (UK) (grant MR/K006355/1). G.M. is supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (grant 206675/Z/17/Z), G.M. received support from the Sackler Foundation Trust. The study was supported in part by the National Institute for Health Research (NIHR) Biomedical Research Centre at King’s College London and from National Institute for Health Research (NIHR) Maudsley Biomedical Research Centre (BRC), J.O., A.D.E. and G.M. received support from the Medical Research Council Centre for Neurodevelopmental Disorders, King's College London (grant MR/N026063/1). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, the Department of Health and IMI 2JU.

Notes
The authors would like to thank all the families who dedicated their time to take part in these studies. Conflict of Interest: The authors declare no conflict of interest.

References