Feature similarity gradients detect alterations in the neonatal cortex associated with preterm birth

Citation for published version:

Digital Object Identifier (DOI):
10.1002/hbm.26660

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Human Brain Mapping

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Feature similarity gradients detect alterations in the neonatal cortex associated with preterm birth

Paola Galdi\textsuperscript{a,b,*,}, Manuel Blesa Cabez\textsuperscript{a}, Christine Farrugia\textsuperscript{c,d}, Kadi Vaher\textsuperscript{a}, Logan ZJ Williams\textsuperscript{e,f}, Gemma Sullivan\textsuperscript{a,g}, David Q Stoye\textsuperscript{a}, Alan J Quigley\textsuperscript{h}, Antonios Makropoulos\textsuperscript{e}, Michael J Thrippleton\textsuperscript{g}, Mark E Bastin\textsuperscript{g}, Hilary Richardson\textsuperscript{i}, Heather Whalley\textsuperscript{g,j}, A David Edwards\textsuperscript{e,k}, Claude J Bajada\textsuperscript{d,l},
Emma C Robinson\textsuperscript{e,f,1}, James P Boardman\textsuperscript{a,g,1}

\textsuperscript{a}MRC Centre for Reproductive Health, University of Edinburgh, UK
\textsuperscript{b}School of Informatics, University of Edinburgh, UK
\textsuperscript{c}Faculty of Engineering, University of Malta, Malta
\textsuperscript{d}University of Malta Magnetic Resonance Imaging Platform (UMRI), Malta
\textsuperscript{e}Centre for the Developing Brain, King’s College London, UK
\textsuperscript{f}School of Biomedical Engineering and Imaging Science, King’s College London, UK
\textsuperscript{g}Centre for Clinical Brain Sciences, University of Edinburgh, UK
\textsuperscript{h}Royal Hospital for Children & Young People, Edinburgh, UK
\textsuperscript{i}School of Philosophy, Psychology and Language Sciences, University of Edinburgh, UK
\textsuperscript{j}Centre for Genomic and Experimental Medicine, University of Edinburgh, UK
\textsuperscript{k}MRC Centre for Neurodevelopmental Disorders, King’s College London, UK
\textsuperscript{l}Department of Physiology and Biochemistry, Faculty of Medicine and Surgery, University of Malta, Malta

Abstract

The early life environment programmes cortical architecture and cognition across the life course. A measure of cortical organisation that integrates information from multi-modal MRI and is unbound by arbitrary parcellations has proven elusive, which hampers efforts to uncover the perinatal origins of cortical health. Here, we use the Vogt-Bailey index to provide a fine-grained description of regional homogeneities and sharp variations in cortical microstructure based on feature gradients, and we investigate the impact of being born preterm on cortical development at term-equivalent age. Compared to term-born controls, preterm infants have a homogeneous microstructure in temporal and occipital lobes, and the medial parietal, cingulate, and frontal cortices, compared with term infants. These observations replicated across two independent datasets and were robust to differences that remain in the data after matching samples and alignment of processing and quality control strategies. We conclude that cortical microstructural architecture is altered in preterm infants in a spatially distributed rather than localised fashion.

Keywords: feature similarity gradients, neonatal brain, preterm birth, MRI, neonatal cortex

*Corresponding author: paola.galdi@ed.ac.uk
1Equal contribution
1. Introduction

The cortex undergoes dramatic changes in the last trimester of pregnancy and early postnatal life as a result of concurrent cellular and molecular processes including synaptogenesis, axonal growth, dendritic arborization and myelination [1, 2, 3]. Disruption of normal gestation during this period of rapid growth can have long-lasting consequences on neurodevelopment [4, 5, 6, 7]. Imaging of the preterm brain at term-equivalent age has characterised dysmaturation following preterm birth as a global phenomenon (i.e. encompassing the whole brain), with generalised dysconnectivity and atypical topology of developing neural networks, increased water diffusivity, altered white matter microstructure and reduced brain volume and cortical surface area [8, 9, 10, 11, 12]. Motivated by the search of neural antecedents of cognitive impairment observed in some preterm born individuals, studies focusing on the neonatal cortex have demonstrated that prematurity impacts cortical growth and microstructural development in a dose-dependent fashion [13, 14] and that preterm birth alters cortical folding [15, 16, 17], thalamo-cortical connectivity [18, 19], and different regional MRI metrics with varying spatial distribution, including water diffusion measures, markers of myelination and cortical morphology metrics [20, 21, 22, 23].

Parcellation of the brain into spatially contiguous and non-overlapping regions with a locally homogeneous profile is a useful abstraction to study cerebral function and organisation. However, there is no consensus on the optimal delineation of brain parcels, whether they are defined cytoarchitectonically, on the basis of function, or on a combination of features [24, 25, 26, 27, 28]. Individual variability in brain shape and organisation further complicates the definition of a universal brain map, as reference templates might not adapt well to the signal of individual acquisitions [27]. One of the main challenges in defining parcels is the arbitrary assignment of sharp boundaries between regions that may only show graded differences in local structure or function. Feature similarity gradients provide a unified approach to studying graded differences in brain organisation [29, 30, 31]. A gradient is defined as an axis of variance in cortical features, along which different areas are positioned to reflect their similarity [32]. This technique has been used to investigate multiple aspects of cortical organisation of the adult brain, including morphology, microstructure, and structural and functional connectivity [31]. Studies based on functional connectivity have found a whole-brain gradient with primary sensory and motor regions at one end of the spectrum and multimodal association regions at the opposite end [32, 33]. This functional gradient is already detectable during adolescence [34], while closely after birth and through childhood the principal axis of variance spans between sensorimotor
and visual areas [35, 34]. Whole-brain gradients derived from microstructure describe a similar hierarchy to functional gradients, but they become increasingly dissociated from their functional counterparts in multimodal regions [36]. A neonatal study using principal component analysis on measures of cortical morphology and microstructure showed that cortical regions were ordered along a principal axis, with the primary cortex clearly separated from the association cortex [22]. Using local rather than global gradients, Glasser et al. [28] proposed a method to detect boundaries in the cortex in correspondence of spatial transitions in multiple brain maps describing different brain features: myelin-sensitive features (derived from T1/T2 contrasts), cortical thickness, cortical folding, and functional activity patterns and connectivity. They did however rely on expert knowledge and priors from the neuroanatomical literature in the absence of sharp transitions.

Here we introduce a framework for a local gradient analysis of the microstructural organisation of the neonatal cortex based on six microstructural metrics: fractional anisotropy, mean, axial and radial diffusivities, neural density index and orientation dispersion index. These metrics are known to share a substantial amount of variance, but at the same time they carry specific information which differs between preterm and term infants, and their covariance structure varies across the brain [12]. We investigate the impact of being born preterm on cortical microstructure using the Vogt-Bailey index (VB) [30]. By integrating information from multiple features computed at each point of the cortical mantle, this index measures the extent of discontinuity in cortical intra-areal relationships, summarising in a single metric the information contained in a feature similarity gradient. Compared to previous approaches based on parcels or whole-brain gradients, this method works on a more local scale, providing a fine-grained description of regional homogeneities and sharp variations in cortical properties. Another advantage of this method is that local homogeneity is computed in subject native space, allowing for greater anatomical specificity. As an additional contribution, we test the replicability of our results in two independent datasets: the Theirworld Edinburgh Birth Cohort (TEBC, [37]) and the developing Human Connectome Project (dHCP, [38]).

2. Methods and Materials

In the following, we detail data acquisition and processing strategies for the two datasets. Table 1 provides a side-by-side comparison of the pipelines for structural and diffusion processing and cortical registration.
2.1. Data acquisition and processing

**Theirworld Edinburgh Birth Cohort (TEBC):** Participants were recruited as part of a longitudinal study on the long-term effects of preterm birth on neurodevelopment [37] between 2016 and 2021. Ethical approval was obtained from the National Research Ethics Service, South East Scotland Research Ethics Committee (11/55/0061, 13/SS/0143 and 16/SS/0154). Cohort exclusion criteria included major congenital malformations, chromosomal abnormalities, congenital infection, overt parenchymal lesions (cystic periventricular leukomalacia, haemorrhagic parenchymal infarction) or post-haemorrhagic ventricular dilatation.

A Siemens MAGNETOM Prisma 3T MRI clinical scanner (Siemens Healthcare, Erlangen, Germany) and 16-channel phased-array paediatric head receive coil were used to acquire a three-dimensional T2-weighted (T2w) SPACE (Sampling Perfection with Application-optimized Contrasts by using flip angle Evolution) structural images with 1 mm isotropic resolution, repetition time (TR) of 3.2 s and echo time (TE) of 409 ms; and a multishell axial diffusion MRI (dMRI) scan (16×b=0 s/mm², 3×b=200 s/mm², 6×b=500 s/mm², 64×b=750 s/mm², 64×b=2500 s/mm²) with optimal angular coverage [39]. Diffusion MRI images were acquired in two separate acquisitions to reduce the time needed to re-acquire data lost to motion artifacts. In addition, an acquisition of 3 b0 volumes with an inverse phase encoding direction was performed. All dMRI images were acquired using single-shot spin-echo echo planar imaging (EPI) with 2-fold simultaneous multislice and 2-fold in-plane parallel imaging acceleration and 2 mm isotropic voxels; all three diffusion acquisitions had the same parameters (TR/TE 3.4 s/78 ms). Images affected by motion artifacts were re-acquired multiple times as required; dMRI acquisitions were repeated if signal loss was seen in 3 or more volumes. Infants were fed and swaddled and allowed to sleep naturally in the scanner. Pulse oximetry, electrocardiography and temperature were monitored. Flexible earplugs and neonatal earmuffs (MiniMuffs, Natus) were used for acoustic protection. All scans were supervised by a doctor or nurse trained in neonatal resuscitation. Full details on the protocol are provided in [37].

Diffusion MRI processing was performed as follows: for each subject the first two dMRI acquisitions were concatenated and then denoised using a Marchenko-Pastur-PCA-based algorithm [40, 41]; eddy current, head movement and EPI geometric distortions were corrected using outlier replacement and slice-to-volume registration [42, 43, 44, 45]; bias field inhomogeneity correction was performed by calculating the bias field of the mean b0 volume and applying the correction to all volumes [46]. Structural
T2w images (1 mm isotropic) were processed using a modified version of the dHCP minimal processing pipeline [47], detailed below. Finally, the mean b0 EPI volume of each subject was co-registered to their structural T2w volume using boundary-based registration [48].

**The developing Human Connectome Project (dHCP):** Participants were scanned at the Evelina Newborn Imaging Centre, Evelina London Children’s Hospital between 2015 and 2019; the study was approved by the National Research Ethics Committee (REC: 14/Lo/1169) [49]. Subjects were excluded if their scans had incidental findings with possible significance for both clinical and imaging analysis (e.g. destructive white matter lesions). To build a sample that matched as closely as possible the characteristics of the TEBC dataset, we included only preterm infants that had their MRI scans acquired at term-equivalent age and we selected term controls with matching age at scan, age at birth and sex distribution using the nearest neighbour method of the R package Matchit [50].

Preprocessed data from the third dHCP data release were used. In brief, images were acquired using a 3T Philips Achieva system (Philips Medical Systems, Best, The Netherlands). All neonates were scanned without sedation in a scanner environment optimised for neonatal imaging, including a dedicated 32-channel neonatal coil. MR-compatible ear putty and earmuffs were used to provide noise attenuation. Neonates were fed, swaddled and positioned in a vacuum jacket prior to scanning to promote natural sleep (without sedation). All scans were supervised by a neonatal nurse and/or paediatrician who monitored heart rate, oxygen saturation and temperature throughout the scan [49]. T2w images were obtained using a Turbo Spin Echo (TSE) sequence, acquired in two stacks of 2D slices (in sagittal and axial planes), using parameters: TR=12s, TE=156ms, SENSE factor 2.11 (axial) and 2.58 (sagittal) with overlapping slices (resolution 0.8 × 0.8 × 1.6 mm). Motion correction and super-resolution reconstruction techniques were employed resulting in isotropic volumes of resolution 0.5 mm³ [51, 52]. Diffusion MRI data were acquired over a spherically optimised set of directions on three shells (b = 400, 1000 and 2600 s/mm²). A total of 300 volumes were acquired per subject, including 20 with b = 0 s/mm². For each volume, 64 interleaved overlapping slices were acquired (in-plane resolution = 1.5 mm, thickness = 3 mm, overlap = 1.5 mm). The data were then super-resolved along the slice direction to achieve isotropic resolution of 1.5 mm³.

Structural T2w images (0.5 mm isotropic) were processed using the minimal processing pipeline of the dHCP [53, 47]. Diffusion image preprocessing was carried out according to the dHCP diffusion processing pipeline [54]. This includes motion correction and distortion correction [42, 43, 44, 45].
2.2. Cortical surface reconstruction

The surface reconstruction from T2w images was carried out using a modified version of the dHCP surface processing pipeline [47]. The standard pipeline works by fitting an initial surface to the boundary of the white and cortical grey matter tissue segmentation; this is subsequently refined by using T2w image intensity values to detect the closest edges with cortical grey matter or cerebrospinal fluid and ‘pushing’ the surface toward these edges [55]. For the dHCP cohort this has been shown to correct segmentation errors and improve surface reconstruction but the parameters of the optimisation are cohort-specific and challenging to tune, often leading to failure on data with low resolution or intensity contrast. To avoid this issue and to align the processing of the two datasets, all surfaces were extracted by running the pipeline with this feature turned off (available through version 1.1.1\(^2\)), thus using only the tissue segmentation for the surface reconstruction.

2.3. Derivation of microstructural tissue maps

DTI maps were calculated from the dMRI processed images to obtain fractional anisotropy (FA) and mean, axial and radial diffusivities (MD, AD and RD, respectively). To calculate the DTI maps, only the shells of \( b = 750 \) s/mm\(^2\) (TEBC) and \( b = 1000 \) s/mm\(^2\) (dHCP) were used. NODDI maps were calculated for both datasets using the Bingham distribution to obtain the neurite density index (NDI), orientation dispersion index (ODI) and the isotropic water fraction (ISO) [56, 57]. The ISO map was used to obtain the tissue fraction (1-ISO), which was used to modulate the NDI and ODI maps [58]. All diffusion metric maps were propagated to the T2w image in each subject’s native space, z-scored and concatenated to obtain for each voxel a vector comprising all six microstructural metrics.

2.4. Registration to a common template

Cortical surfaces were registered to the 40-week symmetric dHCP cortical template [59] using multimodal surface mapping [60, 61]. For this, the script \texttt{align to template no volumetric initialisation.sh}\(^3\) was used. This script works by registering each subject to its age-specific template and then concatenating the subject-to-template registration with a registration which maps the age-specific template to the 40-week template [59]. The configuration parameters were optimised for this study by including an additional regularisation level. The configuration file is available in the project repository (https://git.ecdf.ed.ac.uk/jbrl/neonatal-vbi).

\(^2\)https://github.com/DevelopingHCP/structural-pipeline/tree/dhcp-v1.1.1
\(^3\)https://github.com/ecr05/dHCP_template_alignment
2.5. **Quality control**

We used the *eddy qc* tool [67] to measure in-scanner motion for dMRI acquisitions. For each volume, motion is quantified by averaging voxel displacement across all voxels (computed as 3 translations and 3 rotations around the x, y and z axes). We report the average relative motion (w.r.t. the previous volume) across all volumes. Note that the motion estimates are not directly comparable across datasets due to differences in resolution and parametrization of input sequences and eddy correction. To check the quality of the registration of cortical surfaces to the template space, individual sulcal depth maps were overlaid onto template surfaces and visually inspected by three independent raters (PG, MBC and KV). Subjects were included in the analyses if five anatomical landmarks were clearly identifiable and correctly registered: central sulcus, superior temporal sulcus, lateral fissure, parieto-occipital fissure and calcarine sulcus. An example is shown in figure 1. Inter-rater agreement was 94%. The quality check was performed separately for the left and right hemispheres, resulting in two different samples, one per hemisphere.

2.6. **Vogt-Bailey index**

Given a set of features computed at each vertex of a cortical surface, the Vogt-Bailey index measures the extent of homogeneity in cortical intra-areal relationships. Using a searchlight approach, at each vertex an affinity graph is built measuring affinity between each vertex and its neighbours; the affinity is calculated using a modified Pearson correlation ([30]) computed over the features used to describe the vertices (in this case the six microstructural metrics); only positive correlations are retained. The VB index is then computed as the (scaled) algebraic connectivity of the affinity graph, that is the second smallest eigenvalue of the unnormalised Laplacian matrix of the graph. The index assumes values close to 0 when there are sharp changes in the cortical features in an area, and values close to 1 if the changes are graded. Compared to the original implementation [30], where the neighbourhood was selected as the set of adjoining vertices on the cortical surface, in this work we used a hybrid approach where the neighbourhood is reconstructed in volumetric native space, in order to avoid spatial correlations induced by artefactual geometric patterns that follow cortical gyrification patterns [68, 69]. The coordinates of a given vertex and its neighbours are used to select the closest corresponding voxels in volumetric space, and the affinity graph is computed over the feature vectors of the matched brain tissue voxels that fall within a 27-voxel cube centred on the voxel corresponding to the input vertex, to guarantee that we only select voxel neighbours within a short radius from the central vertex, as vertices on the surface are not equally
Table 1: Summary of acquisition parameters and processing steps for TEBC and dHCP datasets.

<table>
<thead>
<tr>
<th></th>
<th>TEBC</th>
<th>dHCP</th>
<th>Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Image Resolution</strong></td>
<td>1 mm³</td>
<td>0.5 mm³</td>
<td>No</td>
</tr>
<tr>
<td><strong>Preprocessing</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brain Extraction</td>
<td>BET</td>
<td>BET</td>
<td>Yes</td>
</tr>
<tr>
<td>T1 - T2 registration</td>
<td>Rigid Alignment, BBR</td>
<td>Rigid Alignment, BBR</td>
<td>Yes</td>
</tr>
<tr>
<td>Bias Correction</td>
<td>N4</td>
<td>N4</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Segmentation/Surface extraction</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2-weighted images</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Draw-EM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Surface Inflation</strong></td>
<td>FreeSurfer re-implementation</td>
<td>FreeSurfer re-implementation</td>
<td>Yes</td>
</tr>
<tr>
<td>Spherical Projection</td>
<td>sMDS</td>
<td>sMDS</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Diffusion MRI</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brain Extraction</td>
<td>BET</td>
<td>BET</td>
<td>No</td>
</tr>
<tr>
<td>EPI correction</td>
<td>Topup</td>
<td>Topup</td>
<td>No</td>
</tr>
<tr>
<td>Bias correction</td>
<td>N4</td>
<td>None</td>
<td>No</td>
</tr>
<tr>
<td>DTI fitting</td>
<td>distfit</td>
<td>distfit</td>
<td>Yes</td>
</tr>
<tr>
<td>NODDI fitting</td>
<td>Noddi-Bingham cuDIMOT</td>
<td>Noddi-Bingham cuDIMOT</td>
<td>Yes</td>
</tr>
<tr>
<td>Inter modality registration</td>
<td>T2 to b=0 Rigid Alignment, BBR</td>
<td>T2 to b=1000 Rigid Alignment, BBR</td>
<td>No</td>
</tr>
<tr>
<td>Registration to 40-week cortical template 40</td>
<td>MSM</td>
<td>MSM</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Figure 1: Sulcal map of an exemplar 38-week subject (left) compared to the 40-week sulcal template from dHCP (right). White dots demarcate five anatomical landmarks used for visual inspection: central sulcus, superior temporal sulcus, lateral fissure, parieto-occipital fissure and calcarine sulcus.

2.7. Statistical analysis

The PALM tool [70] from the FSL suite [66] was used to perform a surface-based, vertex-wise term vs. preterm comparison of the VB maps,
controlling for PMA at scan and sex. An analysis comparing VB maps between datasets was also run, controlling for prematurity, PMA at scan, sex, and vertex-wise cortical thickness and surface area, while sulcal depth was regressed out directly from the VB maps before running PALM as we hypothesised it to be the main contributor to between-dataset discrepancies as a consequence of the difference in image resolution (see also Appendix A and Appendix B). Separate analyses were run for each hemisphere, and independently for the two datasets. Prior to the analyses, Gaussian-weighted smoothing was applied to individual VB maps, with a smoothing kernel with full-width at half-maximum set at 4 mm. Permutation p-values were computed over 10000 random shuffles with threshold-free cluster enhancement [71] and family-wise error rate corrections. Statistical significance was set at \( p < 0.0253 \) (equivalent to \( \alpha = 0.05 \), after Šidák correction over the two hemispheres, [72]).

3. Results

In the TEBC sample, 254 infants had multimodal MRI data suitable for the proposed analysis. Quality control was performed separately for the left and right hemispheres. After visual inspection of the cortical surfaces, 221 subjects (147 preterm and 72 term-born infants) were included in either the left or right hemisphere analysis, or both (Table 2). Of the preterm infants, 36 had bronchopulmonary dysplasia (BPD), 8 had necrotising enterocolitis (NEC) and 6 required treatment for retinopathy of prematurity (ROP). After quality control, the dHCP matching sample (matched on age at scan, age at birth and sex distribution) contained 43 preterm and 84 term-born infants (Table 3). None of the preterm infants in the dHCP sample had NEC or ROP; no information was recorded regarding BPD. In both datasets, there were no differences in in-scanner relative motion estimates of volume-to-volume displacement between the preterm and term groups (Table 2). The distributions of gestational age at birth and age at scan for the two datasets are shown in Figure 2. Table 4 reports the sample size for the preterm and term groups for each experiment. There were no statistically significant differences in post-menstrual age (PMA) distributions between the left and right hemisphere samples, or between any individual hemisphere and the whole sample, in either dataset (Wilcoxon rank-sum test, \( p\)-values > 0.43).

For each subject, VB maps were generated from six microstructural metrics (fractional anisotropy, mean, axial and radial diffusivities, neural density index and orientation dispersion index). Figure 3 shows two exemplar VB maps for one preterm and one term subject. Figure 4 shows two exemplar VB maps for one preterm and one term subject. Figure 4
shows the distribution of VB index (across both hemispheres) by group (preterm vs. term) for both datasets.

3.1. The impact of preterm birth on cortical organisation

In both datasets, when comparing preterm and term infants controlling for age at scan and sex, we found more homogeneous cortical microstructure (higher VB index) in the preterm group compared to the term group in widespread areas of the cortex including the right temporal lobe, occipital lobe, medial parietal cortex, right cingulate cortex and left frontal cortex. Figure 5 reports the standardised t-maps of the preterm vs. term comparisons for the two datasets, while the p-values maps in Figures 6 and 7 highlight regions where the VB index is significantly higher in the preterm group or in the term group, respectively. In the TEBC dataset only, the VB index is lower (decreased homogeneity) in the preterm group compared to term controls in the medial orbitofrontal cortex. There were no areas of lower VB index in the preterm group in the dHCP dataset.

Table 2: Participant characteristics for the TEBC by group. Mean, minimum and maximum values are reported for continuous variables and ratios for binary variables. The last column reports the p-values of the group differences computed with the Wilcoxon rank-sum test for continuous variables and with the chi-squared test for binary variables.

<table>
<thead>
<tr>
<th></th>
<th>preterm (N=147)</th>
<th>term (N=72)</th>
<th>preterm vs. term</th>
</tr>
</thead>
<tbody>
<tr>
<td>PMA at birth (weeks)</td>
<td>29.48 (22.14-32.86)</td>
<td>39.47 (36.43-42.00)</td>
<td>p = 2.96e-33</td>
</tr>
<tr>
<td>Birth weight (grams)</td>
<td>1318 (370-2510)</td>
<td>3462 (2410-4560)</td>
<td>p = 3.14e-33</td>
</tr>
<tr>
<td>PMA at scan (weeks)</td>
<td>40.71 (36.57-45.86)</td>
<td>41.79 (38.29-43.86)</td>
<td>p = 4.23e-9</td>
</tr>
<tr>
<td>In-scanner motion (mm)</td>
<td>0.24 (0.10-1.04)</td>
<td>0.22 (0.12-0.71)</td>
<td>p = 0.4449</td>
</tr>
</tbody>
</table>

PMA = Postmenstrual age, M = male, F = female.

Table 3: Participant characteristics for the dHCP by group. Mean, minimum and maximum values are reported for continuous variables and ratios for binary variables. The last column reports the p-values of the group differences computed with the Wilcoxon rank-sum test for continuous variables and with the chi-squared test for binary variables.

<table>
<thead>
<tr>
<th></th>
<th>preterm (N=44)</th>
<th>term (N=89)</th>
<th>preterm vs. term</th>
</tr>
</thead>
<tbody>
<tr>
<td>PMA at birth (weeks)</td>
<td>28.98 (23-32.71)</td>
<td>39.03 (37.14-41.86)</td>
<td>p = 3.51e-20</td>
</tr>
<tr>
<td>Birth weight (grams)</td>
<td>1178 (450-1960)</td>
<td>3184 (2060-4300)</td>
<td>p = 3.63e-20</td>
</tr>
<tr>
<td>PMA at scan (weeks)</td>
<td>40.86 (38.14-44.29)</td>
<td>41.07 (37.57-44.71)</td>
<td>p = 0.5838</td>
</tr>
<tr>
<td>M:F ratio</td>
<td>25:18</td>
<td>47:37</td>
<td>p = 7.53e-06</td>
</tr>
<tr>
<td>In-scanner motion (mm)</td>
<td>2.04 (0.75-6.05)</td>
<td>1.72 (0.65-3.64)</td>
<td>p = 0.1277</td>
</tr>
</tbody>
</table>

PMA = Postmenstrual age, M = male, F = female.
3.2. Replicability across datasets

We tested the replicability of our findings in two datasets, and we aimed to reduce avoidable variations in the samples, structural preprocessing and quality control strategies. However, some discrepancies remain, such as scanner differences, input resolution and variation in acquisition protocols; these had inevitable consequences on the down-stream processing, that are reflected in differences in cortical thickness (Appendix A) and single-metric maps between the two datasets (Appendix B) that were spread throughout the cortex.

With the exception of the temporal and occipital lobes, throughout the cortex the VB index was on average higher in the TEBC dataset, possibly as a consequence of lower image resolution, with the strongest effect size measured in the medial frontal lobe and right insular cortex (Figures 8-10). There was overlap in between-dataset differences in cortical thickness and between-dataset differences in VB maps in the temporal lobe and insular cortex and in medial parts of the occipital lobe and the frontal gyrus, but not in the rest of the frontal cortex. Single-metric difference maps also showed overlap with the cortical thickness differences but with variable patterns.

Figure 2: Distribution of postmenstrual age in weeks at birth (left) and at scan (right) for the TEBC and dHCP datasets.

<table>
<thead>
<tr>
<th></th>
<th>left hemisphere</th>
<th>right hemisphere</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>preterm</td>
<td>term</td>
</tr>
<tr>
<td>TEBC</td>
<td>103</td>
<td>57</td>
</tr>
<tr>
<td>dHCP</td>
<td>37</td>
<td>67</td>
</tr>
</tbody>
</table>
Figure 3: Vogt-Bailey index maps of the right hemisphere of two exemplar subjects, one preterm (left) and one term (right), displayed on midthickness (top row) and inflated (bottom row) surfaces registered to the 40-week dHCP template. Values close to zero indicate local discontinuity in the measured microstructural properties, while values close to 1 indicate a high local homogeneity.

Figure 4: Normalised histograms of the VB index distribution by group in the TEBC (left) and the dHCP dataset (right).
Figure 5: Standardised t-statistic maps resulting from the preterm vs. term comparison for the TEBC sample (top panel) and the dHCP sample (bottom panel). Positive values in the maps indicate regions where the Vogt-Bailey index was higher in the preterm group, and negative values regions where the Vogt-Bailey index was higher in the term group.
Figure 6: Results of the preterm vs. term comparison for the TEBC sample (top panel) and the dHCP sample (bottom panel). Highlighted areas indicate regions where the Vogt-Bailey index was higher in the preterm group at an alpha level of 0.05 after Šidák correction. The colour bar reports negative log10 p-values.
Figure 7: Results of the preterm vs. term comparison for the TEB sample (top row) and the dHCP sample (bottom row). Highlighted areas indicate regions where the Vogt-Bailey index was higher in the term group at an alpha level of 0.05 after Šidák correction. The colour bar reports negative log10 p-values.
Figure 8: Standardised t-statistic maps showing the results of the comparison of Vogt-Bailey maps between dataset after controlling for prematurity, age at scan, sex, cortical thickness, surface area and sulcal depth. Positive values in the maps indicate regions where the Vogt-Bailey index was higher in the TEBC sample, and negative values regions where the Vogt-Bailey index was higher in the dHCP sample.

Figure 9: P-value maps showing the regions where the Vogt-Bailey index was higher in the TEBC sample, after controlling for prematurity, age at scan, sex, cortical thickness, surface area and sulcal depth.
Figure 10: P-value maps showing the regions where the Vogt-Bailey index was higher in the dHCP sample, after controlling for prematurity, age at scan, sex, cortical thickness, surface area and sulcal depth.
4. Discussion

In this work, we demonstrate that the VB index can be used to explore graded changes in the microstructural properties of the neonatal cortex. Consistent with previous descriptions of the dysmaturation following preterm birth as a substantially whole-brain phenomenon [8, 73, 12], across two independent datasets we find that a significant proportion of cortical microstructure is consistently altered in association with preterm birth, in a spatially distributed rather than localised fashion.

The increased homogeneity measured by the VB index in preterm infants compared to term controls could reflect increased water content, potentially as a consequence of delayed maturation, as supported by previous findings [14, 23]. In a previous study that used a region-based approach to investigate neuroanatomic variation associated with preterm birth, term infants showed less homogeneous microstructure throughout cortical grey matter regions compared to preterm infants, indicating a more advanced state of maturation [74]. Our results are also consistent with prior evidence that prematurity alters microstructural organisation throughout the cortex: Bouyssi-Kobar et al. [20] reported a higher diffusivity in preterm infants in the prefrontal, parietal, motor, somatosensory, and visual cortices, and Dimitrova et al. [23] reported increased cortical tissue water content and reduced neurite density index in posterior cortical regions, and greater cortical thickness in anterior cortex, in preterm compared to term-born infants. Sensory areas and posterior regions of the cortex have the largest maturational changes in microstructural and morphological profile around birth [75, 76], while the network of cortico-cortical connections involving sensory-motor and transmodal integration regions is characterised by a higher complexity in term compared to preterm infants [11]. Together with our findings, this suggests that cortical maturation in these regions is vulnerable to the impact of preterm birth.

Previous studies have characterised the maturation of microstructure in the preterm brain by measuring deviation from developmental trajectories estimated in term subjects [77, 23] and have shown a high inter-individual heterogeneity in microstructure maturation throughout the cortex. Compared to these works, which examined each microstructural metric in isolation, we studied the joint variation of multiple microstructural properties to observe the impact of prematurity on cortical organisation. Our results complement these previous findings by showing that, overall, despite individual variability in deviation from normal development, prematurity results in locally
homogeneous patterns of variation in a substantial portion of the cortex. In the TEBC dataset, a region of the medial frontal lobe showed increased homogeneity of microstructure in the term subjects compared to the preterm group (Figure 7). We note that care must be taken when interpreting results along the midline sagittal plane, as they might be influenced by image processing artifacts (e.g., due to a brain-size bias causing alignment errors in the preterm group) or cerebrospinal fluid contamination due to partial volume effects. The presence of diverging patterns of increased and decreased homogeneity could indicate that early exposure to the extra-uterine environment causes dysmaturation in a region-specific manner. Different cortical regions mature at different rates, with the insular cortex developing first, and the sensorimotor cortex maturing earlier than association cortices. This is reflected by increasing cortical anisotropy which occurs first in primary somatosensory regions and later in the frontal cortex [16, 78]. Prematurity can cause deviations from typical maturational trajectories, for example with primary and non-primary auditory cortices showing a delayed maturation in preterm infants at term-equivalent age [79]. However, we did not observe decreased homogeneity of microstructure in any region in the dHCP dataset. The variation in results across datasets may be explained partly by unmodelled sample variability (e.g., in illness severity, ethnicity, socioeconomic factors and nutrition), and power differences, but any interpretation needs to be made in the context of the different experimental conditions. We investigated whether the observed variability was a consequence of differences in cortical metrics (Appendix A), that in turn might be driven by partial volume effects, and found that they could explain only in part our results. We also compared single-metric cortical maps between datasets (Appendix B, finding differences in most part of the cortex, but with varying spatial patterns. In another experiment (results not shown), we regressed sulcal depth directly from VB maps, but it did not have any effect on the results. As our goal was to assess the replicability of VB index differences in independent datasets we did not attempt any data harmonisation, but these results serve as a caveat for future retrospective multi-site studies. State of the art harmonisation methods have the potential to minimise site-related differences especially in diffusion data [80], for which the processing was not harmonised in the current study. Allowing for these limitations, the existence of corresponding findings across the two datasets provides evidence of generalisability and indicates that the VB index is robust to scanner differences and methodological variation.

To conclude, the VB index is a promising metric for investigating upstream determinants of cortical development. In contrast to whole-brain gradient
methods, this approach provides superior anatomical precision, enabling detailed analysis of specific brain regions or structures. Through the integration of multiple MRI features into a single localised metric, the VB index offers a comprehensive framework for evaluating how clinical and environmental variables impact the cortex in an integrated analysis. For example, future work could investigate whether the pattern described by the VB index underlies differences in cognitive and functional development observed in preterm infants. Elucidating the link between the widespread alterations in cortical microstructure observed in the preterm brain in the neonatal period and long-term outcomes will be important for study designs that would benefit from robust risk prediction in early life.

5. Data availability

Requests for TEBC anonymised data will be considered under the study's Data Access and Collaboration policy and governance process (https://www.ed.ac.uk/centre-reproductivehealth/tebc/about-tebc/for-researchers/data-access-collaboration). Any additional data supporting the findings of the study are available from the corresponding author, upon reasonable request. Information on how to access dHCP data is available at the project homepage (https://biomedia.github.io/dHCP-release-notes/). The cortical template and the registrations for the dHCP dataset are available at https://brain-development.org/brain-atlases/atlases-from-the-dhcp-project/cortical-surface-template/. Code to reproduce the analyses presented in this paper is available at https://git.ecdf.ed.ac.uk/jbrl/neonatal-vbi. The VB toolbox is available at https://github.com/VBIndex/py_vb_toolbox.

6. Declaration of competing interests

The authors have no conflict of interest to disclose.

7. CRediT author statement

PG: Conceptualisation, Methodology, Software, Formal Analysis, Writing - Original Draft. MBC: Methodology, Software, Formal Analysis, Data

8. Acknowledgements

We thank the families who took part in the study and the Edinburgh Imaging research staff for providing the infant scanning. This work was supported by Theirworld and was undertaken in the MRC Centre for Reproductive Health, which is funded by MRC Centre Grant [MRC G1002033]. JPB would like to acknowledge funding from the UKRI MRC programme grant (MR/X003434/1). PG was partly supported by the Wellcome Trust - University of Edinburgh Institutional Strategic Support Fund 3 [IS3-R1.1320/21]. CJB would like to acknowledge funding for the Boundaries of the Brain project (REP 2020 005) by the MCST Research Excellence Programme 2020 Call and the BE-BOP University of Malta Research Excellence Fund Grant Agreement No. 202201. MJT is supported by NHS Lothian Research and Development Office. KV was funded by the Wellcome Translational Neuroscience PhD Programme at the University of Edinburgh (108890/Z/15/Z). ECR was supported by the Academy of Medical Sciences/the British Heart Foundation/the Government Department of Business, Energy and Industrial Strategy/the Wellcome Trust Springboard Award [SBF003/1116] and a Wellcome Collaborative Award [215573/Z/19/Z]. LZJW is supported by funding from the Commonwealth Scholarship Commission, United Kingdom. The dHCP is funded by the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement no. [319456]. For the purpose of open access, the authors have applied a
References


L. Z. Williams, S. P. Fitzgibbon, J. Bozek, A. M. Winkler, R. Dimitrova,


