

Neonatal morphometric similarity mapping for predicting brain age and characterizing neuroanatomic variation associated with preterm birth

Paola Galdi^{a,1,*}, Manuel Blesa^{a,1}, David Q. Stoye^a, Gemma Sullivan^a, Gillian J. Lamb^a, Alan J. Quigley^b, Michael J. Thrippleton^{c,d}, Mark E. Bastin^c, James P. Boardman^{a,c}

^aMRC Centre for Reproductive Health, University of Edinburgh, Edinburgh EH16 4TJ, UK

^bDepartment of Radiology, Royal Hospital for Sick Children, Edinburgh EH9 1LF, UK

^cCentre for Clinical Brain Sciences, University of Edinburgh, Edinburgh EH16 4SB, UK

^dEdinburgh Imaging, University of Edinburgh, Edinburgh EH16 4SB, UK

Abstract

Multi-contrast MRI captures information about brain macro- and micro-structure which can be combined in an integrated model to obtain a detailed “fingerprint” of the anatomical properties of an individual’s brain. Inter-regional similarities between features derived from structural and diffusion MRI, including regional volumes, diffusion tensor metrics, neurite orientation dispersion and density imaging measures, can be modelled as morphometric similarity networks (MSNs). Here, individual MSNs were derived from 105 neonates (59 preterm and 46 term) who were scanned between 38 and 45 weeks postmenstrual age (PMA). Inter-regional similarities were used as predictors in a regression model of age at the time of scanning and in a classification model to discriminate between preterm and term infant brains. When tested on unseen data, the regression model predicted PMA at scan with a mean absolute error of 0.70 ± 0.56 weeks, and the classification model achieved 92% accuracy. We conclude that MSNs predict chronological brain age accurately; and they provide a data-driven approach to identify networks that characterise typical maturation and those that contribute most to neuroanatomic variation associated with preterm

*Correspondence: Paola Galdi, Queen’s Medical Research Institute, 47 Little France Crescent, Edinburgh EH16 4TJ, UK. Email: paola.galdi@ed.ac.uk

¹These authors contributed equally to the work.

birth.

Keywords: morphometric similarity networks, preterm, developing brain, brain age, multi-modal data, MRI

Highlights

1. Multiple MRI features are integrated in a single model to study brain maturation in newborns.
2. Morphometric similarity networks (MSNs) provide a whole-brain description of the structural properties of neonatal brain.
3. The information encoded in MSNs is predictive of chronological brain age in the perinatal period.
4. MSNs provide a novel data-driven method for investigating neuroanatomic variation associated with preterm birth.

1. Introduction

Preterm birth is closely associated with increased risk of neurodevelopmental, cognitive and psychiatric impairment that extends across the life course (Nosarti et al., 2012; Anderson, 2014; Mathewson et al., 2017; Van Lieshout et al., 2018). Structural and diffusion MRI (sMRI and dMRI) support the conceptualisation of atypical brain growth after preterm birth as a process characterised by micro-structural alteration of connective pathways due to impaired myelination and neuronal dysmaturation (Boardman et al., 2006; Anjari et al., 2007; Counsell et al., 2008; Ball et al., 2013b; Back and Miller, 2014; Van Den Heuvel et al., 2015; Eaton-Rosen et al., 2015; Thompson et al., 2016; Batalle et al., 2017; Telford et al., 2017; Batalle et al., 2018); this leads to a “dysconnectivity phenotype” that could form the basis for long term functional impairment (Boardman et al., 2010; Caldinelli et al., 2017; Keunen et al., 2017; Cao et al., 2017; Batalle et al., 2018b). However, there has not been a unified approach that incorporates information from sMRI and dMRI to study brain maturation

16 in the perinatal period so the set of image features that best capture brain
17 maturation, and support image classification, are unknown.

18 The majority of neonatal connectomics studies have used single modes of
19 data such as dMRI tractography (Brown et al., 2014; Batalle et al., 2017; Blesa
20 et al., 2019) or resting-state functional connectivity (Ball et al., 2016; Smyser
21 et al., 2016a). An alternative connectome model is the structural covariance
22 network (SCN) approach (Alexander-Bloch et al., 2013) in which covariance be-
23 tween regional measurements is calculated across subjects, resulting in a single
24 network for the entire population. Other approaches have constructed subject-
25 specific SCNs (Li et al., 2017; Mahjoub et al., 2018) or higher order morpho-
26 logical networks to model the relationship between ROIs across different views
27 (Soussia and Rekik, 2018), but these techniques have been restricted to the use
28 of morphometric variables available through standard structural T1-weighted
29 MRI sequences and by using a single metric (e.g. cortical thickness) to assess
30 the “connectivity” between nodes (Shi et al., 2012).

31 Based on observations that integrating data from different MRI sequences
32 enhances anatomic characterization (Melbourne et al., 2014; Kulikova et al.,
33 2015; Ball et al., 2017; Thompson et al., 2018a), we investigated whether whole-
34 brain structural connectomes derived from multi-modal data within a prediction
35 framework can capture novel information about perinatal brain development.
36 We used morphometric similarity networks (MSNs) to model inter-regional cor-
37 relations of multiple macro- and micro-structural multi-contrast MRI variables
38 in a single individual. This approach was originally devised to study how hu-
39 man cortical networks underpin individual differences in psychological functions
40 (Seidlitz et al., 2018), and we adapted it to describe both cortical and subcor-
41 tical regions in the developing brain. The method works by computing for
42 each region of interest (ROI) a number of metrics derived from different MRI
43 sequences which are arranged in a vector. The aim is to obtain a multidimen-
44 sional description of the structural properties of the ROIs. The MSN is then
45 built considering the ROIs as nodes and modelling connection strength as the
46 correlation between pairs of ROI vectors, thus integrating in a single connectome

the ensemble of imaging features. The pattern of inter-regional correlations can be conceptualised as a “fingerprint” of an individual’s brain.

We investigated the utility of MSNs for describing brain maturation, and for patient classification. The edges of individual MSNs were used to train two predictive models: a regression model to predict postmenstrual age (PMA) at scan and identify the set of image features that best model chronological brain age; and a classification model to discriminate between preterm infants at term equivalent age and term neonates, and thereby identify the networks that explain neuroanatomic variation associated with preterm birth. We hypothesized that predictive models based on MSNs, which integrate information from multiple data modalities, would outperform models based on single metrics and single data modalities.

2. Material and methods

2.1. Participants and data acquisition

Participants were recruited as part of a longitudinal study designed to investigate the effects of preterm birth on brain structure and long term outcome. The study was conducted according to the principles of the Declaration of Helsinki, and ethical approval was obtained from the UK National Research Ethics Service. Parents provided written informed consent. One hundred and twelve neonates underwent MRI at term equivalent age at the Edinburgh Imaging Facility: Royal Infirmary of Edinburgh, University of Edinburgh, UK, and 105 had multi-modal imaging suitable for MSN analysis (7 acquisitions did not yield usable datasets across all modalities due to motion or wakefulness during one or more sequences). The study group contained 46 term and 59 preterm infants (details are provided in Table 1). The distribution of PMA at scan for all participants, for the term and preterm groups, and the distribution by gender are shown in Fig. 1. Of the preterm infants, 12 had bronchopulmonary dysplasia, 3 had necrotising enterocolitis and 3 required treatment for retinopathy of prematurity.

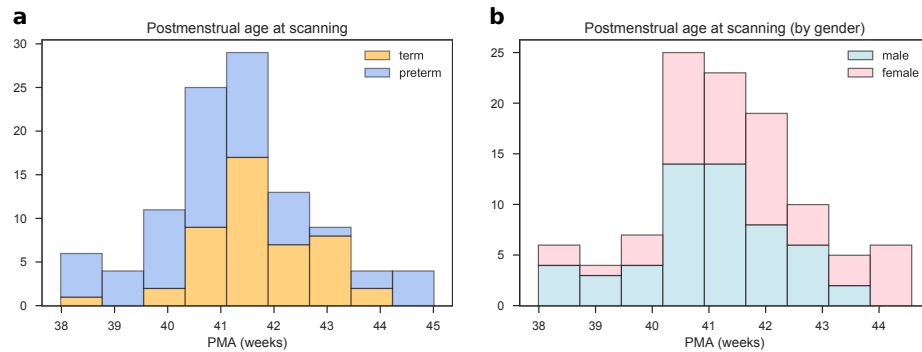


Figure 1: Distribution of postmenstrual age at scan for all subjects. a) Age distribution for the for term (blue) and preterm (orange) groups. b) Age distribution for male (blue) and female (pink) participants.

Table 1: Participant characteristics. 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 categorical variables.

	preterm (N=59)	term (N=46)	all (N=105)	preterm vs. term
PMA at birth (weeks)	23.42-32.00	37.00-42.00	23.42-42.00	$p = 1.88 \times 10^{-18}$
Birth weight (grams)	454-2100	2556-4560	454-4560	$p = 1.93 \times 10^{-18}$
PMA at scan (weeks)	38.00-44.56	38.28-43.84	38.00-44.56	$p = 0.0035$
M:F ratio	29:30	26:20	55:50	$p = 0.4532$

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

76 A Siemens MAGNETOM Prisma 3 T MRI clinical scanner (Siemens Health-
77 care Erlangen, Germany) and 16-channel phased-array paediatric head coil were
78 used to acquire: 3D T1-weighted MPRAGE (T1w) (acquired voxel size = 1mm
79 isotropic) with TI 1100 ms, TE 4.69 ms and TR 1970 ms; 3D T2-weighted
80 SPACE (T2w) (voxel size = 1mm isotropic) with TE 409 ms and TR 3200 ms;
81 and axial dMRI. dMRI was acquired in two separate acquisitions to reduce the
82 time needed to re-acquire any data lost to motion artefact: the first acquisition
83 consisted of 8 baseline volumes ($b = 0$ s/mm² [b0]) and 64 volumes with $b =$
84 750 s/mm², the second consisted of 8 b0, 3 volumes with $b = 200$ s/mm², 6
85 volumes with $b = 500$ s/mm² and 64 volumes with $b = 2500$ s/mm²; an op-
86 timal angular coverage for the sampling scheme was applied (Caruyer et al.,

2013). 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 3400/78.0 ms).

Infants were fed and wrapped and allowed to sleep naturally in the scanner. Feeds were timed to increase the likelihood of post-prandial sleep, flexible earplugs and neonatal earmuffs (MiniMuffs, Natus) were used for acoustic protection, and a soothing environment was created in terms of light and noise. Pulse oximetry, electrocardiography and temperature were monitored. All scans were supervised by a doctor or nurse trained in neonatal resuscitation. Each acquisition was inspected contemporaneously for motion artefact and repeated if there had been movement but the baby was still sleeping; dMRI acquisitions were repeated if signal loss was seen in 3 or more volumes. The majority of the cohort had one or more sequences repeated in order to acquire the best possible quality data for processing.

Conventional images were reported by an experienced paediatric radiologist (A.J.Q.) using a structured system (Leuchter et al., 2014; Woodward et al., 2006), and none of the images included in the final sample ($N = 105$) showed evidence of focal parenchymal injury (defined as post-haemorrhagic ventricular dilatation, porencephalic cyst or cystic periventricular leukomalacia), or central nervous system malformation.

2.2. Data preprocessing

All the following preprocessing steps, including maps calculation and quality check, were performed using dcm2niix, FSL, MRtrix, MIRTk, ANTs, Connectome Workbench and cuDIMOT (Smith et al., 2004; Avants et al., 2011; Marcus et al., 2011; Makropoulos et al., 2014; Li et al., 2016; Hernandez-Fernandez et al., 2019; Tournier et al., 2019).

First, all DICOM image files (dMRI and sMRI) were converted to NIFTI (Li et al., 2016). Structural data were preprocessed using the developing Human

117 Connectome Project (dHCP) minimal structural processing pipeline for neonatal data (Makropoulos et al., 2018). Briefly, the T1w image was co-registered
118 to the T2w image, both were corrected for bias field inhomogeneities (Tustison
119 et al., 2010) and an initial brain mask was created (Smith, 2002). Following this,
120 the brain was segmented into different tissue types (CSF: cerebrospinal fluid;
121 WM: white matter; cGM: cortical grey matter; GM: subcortical grey matter)
122 using the Draw-EM algorithm (Makropoulos et al., 2014). Twenty manually
123 labelled atlases (Gousias et al., 2012) were then registered to each subject using
124 a multi-channel registration approach, where the different channels of the
125 registration were the original intensity T2-weighted images and GM probability
126 maps. These GM probability maps were derived from an initial tissue segmentation,
127 performed using tissue priors propagated through registration of a preterm
128 probabilistic tissue atlas (Serag et al., 2012). The framework produces several
129 output files, but for this study only the aligned T1w and the T2w images and
130 the parcellation in 87 ROIs were used (Makropoulos et al., 2018). Note that
131 from these 87 ROIs six were removed: the background, the unlabelled brain
132 area (mainly internal capsule), the CSF, the lateral ventricles (left and right)
133 and the corpus callosum (see section 2.4).
134

135 Diffusion MRI processing was performed as follows: for each subject the two
136 dMRI acquisitions were first concatenated and then denoised using a Marchenko-
137 Pastur-PCA-based algorithm (Veraart et al., 2016,b); the eddy current, head
138 movement and EPI geometric distortions were corrected using outlier replacement
139 and slice-to-volume registration with TOPUP and EDDY (Andersson
140 et al., 2003; Smith et al., 2004; Andersson and Sotiropoulos, 2016; Andersson
141 et al., 2016, 2017); bias field inhomogeneity correction was performed by calculating
142 the bias field of the mean b0 volume and applying the correction to all the
143 volumes (Tustison et al., 2010). This framework only differs from the optimal
144 pipeline for diffusion preprocessing presented in Maximov et al. (2019) in that
145 we did not perform the final smoothing or the gibbs-ring removal (Kellner et al.,
146 2016) due to the nature of the data (partial fourier space acquisition).
147

The mean b0 EPI volume of each subject was co-registered to their structural

148 T2w volume using boundary-based registration (Greve and Fischl, 2009), then
 149 the inverse transformation was used to propagate ROI labels to dMRI space,
 150 with a modified bbrslope parameter of 0.5, which is used for neonatal data
 151 (Toulmin et al., 2015).

152 For each ROI, two metrics were computed in structural space: ROI volume
 153 and the mean T1w/T2w signal ratio (Glasser and Van Essen, 2011). The other
 154 ten metrics were calculated in native diffusion space: five metrics derived from
 155 the diffusion kurtosis (DK) model (Jensen et al., 2005) and five derived from the
 156 Neurite Orientation Dispersion and Density Imaging model (NODDI) (Zhang
 157 et al., 2012; Tariq et al., 2016).

158 *2.3. Feature extraction*

159 *2.3.1. Structural metrics*

160 ROI volumes were calculated without normalising for the whole brain vol-
 161 ume, as they are used only to compute inter-regional similarities within subjects.
 162 The mean T1w/T2w signal ratio was calculated before the bias field correction.
 163 The T1w/T2w ratio was used because it enhances myelin contrast and math-
 164 ematically cancels the signal intensity bias related to the sensitivity profile of
 165 radio frequency receiver coils (Glasser and Van Essen, 2011).

166 *2.3.2. Diffusion kurtosis metrics*

167 The diffusion kurtosis (DK) model is an expansion of the diffusion tensor
 168 model. In addition to the diffusion tensor, the DK model quantifies the degree
 169 to which water diffusion in biological tissues is non-Gaussian using the kurtosis
 170 tensor. The reason for this is that the Gaussian displacement assumption un-
 171 derlying the diffusion tensor breaks at high b-values (Jensen et al., 2005). On
 172 the kurtosis component, we only focus on the mean value along all diffusion
 173 directions.

174 The metrics obtained from the DK model for each ROI are the means of: the
 175 fractional anisotropy (FA), mean, axial and radial diffusivity (MD, RD, AD) and
 176 kurtosis (MK). The MK map quantifies the deviation from Gaussianity of water

177 molecule displacement and can reflect different degrees of tissue heterogeneity
178 (Steven et al., 2014).

179 2.3.3. NODDI metrics

180 We included NODDI metrics alongside the more commonly adopted diffu-
181 sion tensor measures as previous studies have shown that NODDI indices are
182 sensitive to underlying biological changes in the brain and provide more spe-
183 cific microstructural characteristics, in agreement with histology (Grussu et al.,
184 2017; Batalle et al., 2018).

185 For the NODDI measures, the Bingham distribution was employed (Tariq
186 et al., 2016) as it allows extra flexibility by describing fibre dispersion along
187 two orthogonal axes. From this NODDI implementation we obtain five metrics:
188 intracellular volume fraction (v_{ic}), isotropic volume fraction (v_{iso}), the orien-
189 tation dispersion index along the primary and secondary directions (ODI_P and
190 ODI_S) and the overall orientation dispersion index (ODI_{TOT}).

191 One limitation of this model is that it requires fixing a value for the diffu-
192 sivity along the axons. However, optimal values for this parameter are region-
193 dependent (Karmacharya et al., 2018) and the default value may be suboptimal
194 for the neonatal population as it has been optimised using an adult cohort
195 (Zhang et al., 2012; Karmacharya et al., 2018). Several studies have been re-
196 porting NODDI values for neonates using default (or unspecified) parameters
197 (Batalle et al., 2018; Bastiani et al., 2018; Karmacharya et al., 2018) or modi-
198 fied ones (Kunz et al., 2014; Jelescu et al., 2015). As our goal was not to report
199 NODDI values for the different areas, and because of the lack of reference val-
200 ues for this population, we calculated NODDI maps using default parameters
201 (Batalle et al., 2018).

202 2.4. Data Quality Control

203 The parcellations obtained after the processing were visually inspected and
204 parcels corresponding to CSF and background parcels were excluded because
205 they do not represent brain tissue. We observed a poor segmentation of the

corpus callosum in part of the subjects, but we did not find any anomalies in the rest of the parcels. This effect could be caused by different factors: a) this area is problematic to segment due to the proximity to CSF and its small thickness (see for example Otsuka et al. (2019)); b) the framework we used was optimised for the dHCP data that have a very high resolution (0.5 mm^3 isotropic) and data quality, making the partial volume effect more noticeable in data with a resolution of 1 mm^3 ; c) or susceptibility artifacts. Instead of removing the subjects with a poor segmentation, we decided to remove the corpus callosum from the model, aiming at maximising the number of subjects. As a result of the whole quality check, we include the whole population ($N = 105$) and each network is composed of 81 nodes (ROIs).

For the dMRI data we use eddy QC (Bastiani et al., 2019). The quality control is performed at subject level and group level. Eddy QC provides several measures related to the rotation, translation and outliers of the images. In addition, it also computes the signal-to-noise (SNR) ratio maps of the b_0 volumes and the contrast-to-noise (CNR) ratio maps for the different b -values. These maps can be used at group level to visualise the quality of the data (Bastiani et al., 2018). The results show that the overall quality of the data-set was good (Fig. 2). For eddy QC to work, we removed the $b\text{-value} = 200 \text{ s/mm}^2$ only from the quality control. This is because the low number of volumes with this b -value sometimes leads the Gaussian process performed by eddy to produce a perfect fit, which makes the CNR maps unrealistic.

Fig. 2 shows two representative subjects, one from the top quartile of the SNR and CNR distributions (green star) and one from the bottom quartile (red star). In the first panel we can see where they are placed in terms of SNR and CNR over the overall population. The second panel shows the SNR maps (for the b_0) and the CNR maps (for the rest of b -values). The bottom panel of the Fig. 2 shows the b_0 before and after the processing of the selected subjects. It is possible to observe the effect of the different steps involved, such as the EPI geometric corrections or the bias field inhomogeneity correction. Supplementary Figs. S8 and S9 report the above results for the term and preterm population

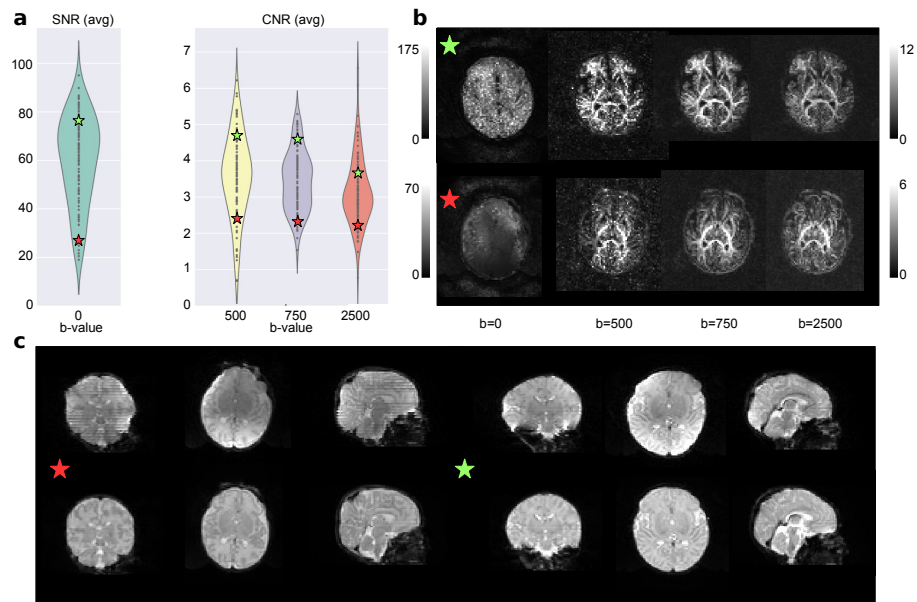


Figure 2: Quality control results. a) Results for the overall population with two selected subjects, one from the top quartile of the SNR and CNR distributions (green star) and the other from the bottom quartile (red star). b) The SNR and CNR maps for the selected subjects. c) The b0 of both subjects before and after the processing pipeline.

237 respectively.

238 Following Bastiani et al. (2019), for each volume, motion is quantified by
 239 averaging voxel displacement across all voxels (computed as 3 translations and
 240 3 rotations around the x, y and z axes). Absolute displacement is computed
 241 with respect to the reference volume, while relative displacement is computed
 242 with respect to the previous volume. A summary measure for each subject is
 243 calculated as the average (absolute or relative) displacement across all volumes.
 244 In Supplementary Fig. S10 we show the distribution of absolute and relative
 245 motion for the term and the preterm groups. We compared the distributions
 246 with a Wilcoxon rank-sum test and found no difference between the relative
 247 motion scores ($W = 1330$, $p = 0.43$) and a significant difference between the
 248 absolute motion scores ($W = 1720$, $p = 0.02$). However, as the violin plot
 249 shows, this difference is driven by the presence of outliers.

250 *2.5. Experimental design and statistical analysis*

251 The models and the analyses described in this section were implemented in
252 Python (v3.6.4) using open source libraries and frameworks for scientific com-
253 puting, including SciPy (v1.0.0), Numpy (v1.14.0), Statsmodels (v0.8.0), Pan-
254 das (v0.22.0), Scikit-learn (v0.19.1) and Matplotlib (v2.1.2) (Jones et al., 2001;
255 Hunter, 2007; Seabold and Perktold, 2010; McKinney et al., 2010; Pedregosa
256 et al., 2011; Van Der Walt et al., 2011).

257 *2.5.1. Network Construction*

258 The MSN for each subject was constructed starting from 81 ROIs; each of the
259 ROI metrics was normalised (z-scored) and Pearson correlations were computed
260 between the vectors of metrics from each pair of ROIs. In this way, the nodes of
261 each network are the ROIs and the edges represent the morphometric similarity
262 between the two related ROIs (Fig. 3). In the following, the terms “edge”,
263 “connection” and “inter-regional similarity” are used interchangeably to refer
264 to the correlation between the regional metrics of a pair of ROIs.

265 *2.5.2. Confounding variables*

266 Early exposure to the extrauterine environment due to preterm birth ex-
267 poses infants to several processes that are known to impact brain maturation
268 (e.g. specific co-morbidities such as bronchopulmonary dysplasia and necrotis-
269 ing enterocolitis (Barnett et al., 2018)), and other processes and diseases that
270 can modify brain maturation (for example gestational age at birth, chorioam-
271 nionitis, fetal growth restriction, nutritional insufficiency, pain and medication
272 exposures (Duerden et al., 2016; Anblagan et al., 2016; Barnett et al., 2018;
273 Schneider et al., 2018; Duerden et al., 2018; Blesa et al., 2019)). In addition,
274 there may be as yet unknown environmental risks to brain structural connec-
275 tivity and genomic and epigenomic factors may interact with gestational age at
276 birth to confer risk (Batalle et al., 2017, 2018b; Boardman et al., 2014; Sparrow
277 et al., 2016; Krishnan et al., 2017). Therefore, it is not possible to define a
278 preterm infant cohort without any exposures to processes that could influence

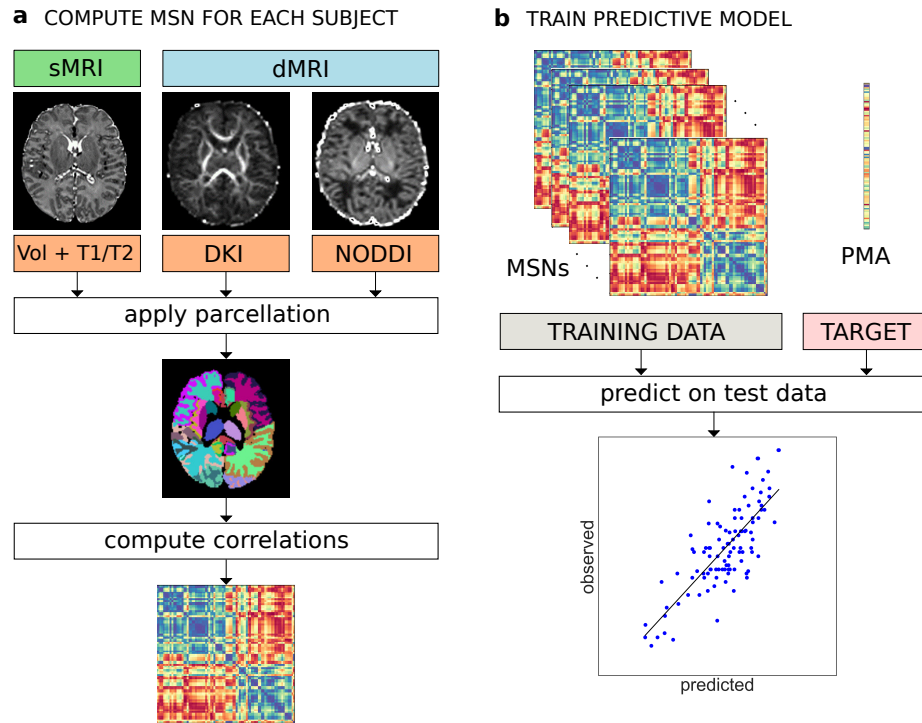


Figure 3: a) Individual MSN construction. Different metrics are extracted from dMRI and sMRI data. The same parcellation is applied to all image types and the average metric values are computed for each ROI. A MSN (represented here as a connectivity matrix) is built by computing the Pearson correlation between the vectors of metrics of each pair of ROIs. b) Training of a predictive model (here for PMA at scan) from individual MSNs. The inter-regional correlations are used as predictor variables in a machine learning model. The performance of the model is evaluated on an independent test set.

279 brain maturation. As our intention was to develop an integrated approach for
280 characterising dysmaturation in a study group representative of the target pop-
281 ulation, rather than to investigate possible drivers of dysmaturation, we did not
282 control for any of the above factors.

283 We did however find that the preterm group was characterised by higher in-
284 scanner motion than the term-group, hence we considered absolute displacement
285 as a confounder (section 2.4). We also observed a positive correlation ($\rho =$
286 $0.27, p = 0.0048$) between PMA at scan and PMA at birth and a negative
287 correlation ($\rho = -0.22, p = 0.0233$) between PMA at scan and gender (coded as
288 a binary variable where 0 indicates female infants and 1 male infants), implying
289 that in our sample term subjects and female subjects tend to have their scan
290 acquired at a later age (see also fig. 1). To control for potential bias, we
291 used these confounders as predictors and compared their predictive performance
292 with our network-based features. We tested the interaction between gender and
293 prematurity in a linear regression model of PMA at scan, but the interaction
294 term was not significant ($p = 0.9634$). Birthweight was not included explicitly
295 as a confounder due to its collinearity with PMA at birth.

296 2.5.3. Regression model for age

297 We trained a linear regression model with elastic net regularisation to pre-
298 dict PMA at scan – i.e. chronological brain age – in both preterm and term
299 infants starting from individual MSNs. This model was chosen for its ability to
300 cope with a high number of features (Zou and Hastie, 2005). For each subject,
301 the edges of the MSN (inter-regional correlations) were concatenated to form a
302 feature vector to be given as input to the regression model. Since the connec-
303 tivity matrix representing the MSN is symmetric, we considered only the upper
304 triangular matrix for each subject. Gender and age at birth were included in the
305 model to control for their possible confounding effects. The prediction perfor-
306 mances were evaluated with a leave-one-out cross-validation (LOOCV) scheme,
307 by computing the mean absolute error (MAE) averaged across subjects. Within
308 each fold of the LOOCV, the parameters of the elastic net were selected with

309 a nested 3-fold cross-validation loop; the folds were stratified in percentiles to
 310 include samples covering the whole age range in each of the folds. Permutation
 311 testing was used for the statistical validation of the model performance: the null
 312 distribution was built by running the age prediction analysis on 1000 random
 313 permutation of the PMA.

314 *2.5.4. Classification model*

315 A Support Vector Machine (SVM) classifier with linear kernel was trained
 316 to discriminate between preterm and term infants. As per the regression model,
 317 the input for each subject consisted of inter-regional connections taken from the
 318 upper triangular connectivity matrix and the performances were evaluated with
 319 LOOCV. Age at the time of scanning, gender and motion were included as ad-
 320 ditional covariates. While in the case of regression the elastic net regularisation
 321 performs automatically a variable selection step, recursive feature elimination
 322 (RFE) was applied in combination with SVM to select the best subset of con-
 323 nections. Model selection was implemented using nested cross validation: an
 324 outer 3-fold cross-validation loop was used to select the SVM parameters and
 325 an inner 4-fold cross-validation loop was used for RFE. Folds were stratified to
 326 include the same proportion of term and preterm subjects. The accuracy of
 327 the model was computed as the number of correctly classified subjects across
 328 the leave-one-out folds over the total number of subjects in the test set. The
 329 null distribution was built by repeating the exact same analysis 1000 times after
 330 randomly assigning subjects to the term and the preterm group.

331 *2.5.5. Feature selection*

332 After the preprocessing phase, twelve different metrics were available for each
 333 ROI. To study which combination of features produced better performance in
 334 the prediction tasks, we implemented a sequential backward-forward feature
 335 selection scheme. Starting from the full set of features, at each iteration we
 336 compare the performances of different models built by removing in turn each of
 337 the features from the current set of candidate features. We then exclude from

the next iteration the feature whose subtraction caused the least increase in prediction error (down to three features, for a total of 73 combinations). The rationale behind this scheme is to explore the space of possible models without enumerating all possible solutions, thus reducing the computational demands compared to an exhaustive search. The procedure was performed separately for the regression and the classification models.

2.5.6. Cross-validation strategy

We adopted LOOCV to select the best performing model in both the age prediction and the classification tasks as this scheme enabled maximum size of the training set and therefore best use of available data, but this strategy might induce high variance in the estimation of prediction accuracy (Kohavi, 1995; Efron, 1983). In the context of brain decoding (i.e. predictions from brain images or signals), LOOCV was shown to produce overly optimistic estimates of prediction accuracy in the within-subject setting (i.e. when all samples are highly correlated because they come from the same subject). In the between-subject setting (as in this work), the performance of LOOCV is similar to schemes involving random splits and mostly determined by sample size (Varoquaux et al., 2017; Varoquaux, 2018). To assess the stability of our results with respect to the chosen cross-validation scheme, we report the prediction accuracy computed with a 10 repeated stratified 5-fold scheme (10-5-fold) for all the models selected with LOOCV.

2.5.7. Comparison with individual metrics and single data modalities models

We compared the performances of the best performing models based on MSNs with three classes of baseline models: a) models based on single global brain metrics (total brain volume and median FA in the WM); b) models based on individual metrics, where instead of similarities, predictors are the concatenation of all regional values for each of the individual metrics used to build MSNs; c) single data modality MSNs, i.e. models built on structural features only (Volume and T1/T2), on DKI features only, and on NODDI features only.

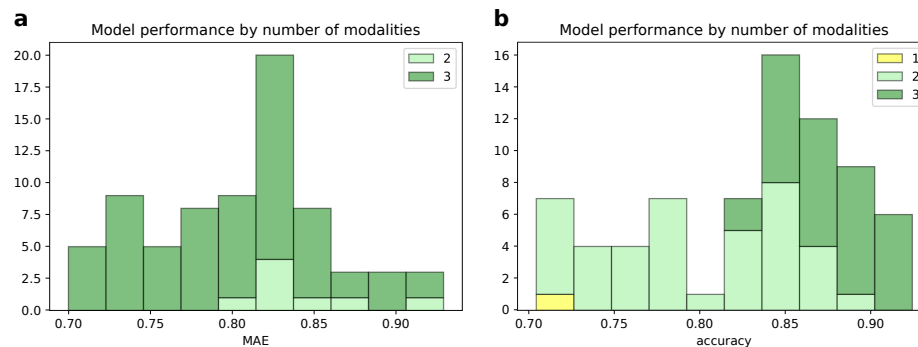


Figure 4: Histograms of the performance of the 73 models compared in the backward feature selection scheme for the age prediction task (a) and for the classification task (b). Bars are grouped by the number of modalities included in the models.

2.6. Data and code availability

Source code implementing the methods described in this paper is available upon request to the corresponding author. The preprocessed and anonymised data used in the analyses can be requested through the Brains Image Bank (<https://www.brainsimagebank.ac.uk/>) (Job et al., 2017).

3. Results

3.1. Feature selection

In Fig. 4 we report two histograms summarising the LOOCV performance of the 73 different models compared per each task in the backward feature selection scheme. In both cases, we can observe that the models based on all three data modalities achieved better results in terms of prediction accuracy. The performances of each of the compared model are reported in Supplementary Figs. S1 and S3 for the age prediction and for the classification models, respectively.

The best performing model for age prediction, which was adopted for all subsequent analyses, was based on seven features (Volume, FA, MD, AD, MK, v_{iso} , ODI_P). Fig. 5a shows the average MSN matrix computed across all subjects for the selected set of features and the matrix of correlation between inter-regional similarities and PMA at scan across subjects. The average MSN matrix shows

four main blocks that correspond roughly to positive correlations between ROIs within GM and between ROIs within WM, and to negative correlation between WM ROIs and GM ROIs, indicating that ROIs within GM (and within WM) share similar structural properties, while GM and WM regional descriptors tend to be anti-correlated. The four-block structure is recognisable also in the matrix reporting correlations with chronological age: with increasing age regions within GM or within WM become more similar with each other, while the dissimilarities between GM and WM ROIs increases.

The best classifier model was based on eleven out of the twelve features (all except ODI_S), so compared to the age prediction model, four additional features were included: T1/T2, RD, v_{ic} and ODI_{TOT} . The average MSN computed with the selected features and the matrix of correlation with PMA at birth is shown in Fig. 5 (panels b and c). Comparing panel b and d of Fig. 5, it is apparent that while the patterns of correlation with PMA at scan and at birth are similar within GM and WM, subcortical ROIs show an opposite trend: with increasing PMA at scan subcortical ROIs tend to become more similar to WM ROIs and more dissimilar to GM ROIs, but the similarity between subcortical ROIs and cortical GM is positively correlated to age at birth.

3.2. Prediction results

The best regression model selected with LOOCV predicted chronological age (PMA at scan) with a MAE of 0.70 ± 0.56 weeks on the test data, and a correlation between the predicted and the actual age equal to $r = 0.78$ ($p = 1.71 \times 10^{-22}$) (Supplementary Fig. S5). The results of the permutation test are shown in Fig. 6 and Supplementary Fig. S6. The confounding variables (gender and age at birth) were not selected by the internal feature selection procedure, hence the predictions were based on network features alone. To test whether there was any systematic difference in the predicted age between the term and the preterm group, we compared the error distributions with a Wilcoxon rank-sum test, but the result was not significant ($W = 1108$, $p = 0.1085$). For comparison, we evaluated the predictive performance of a linear regression model using

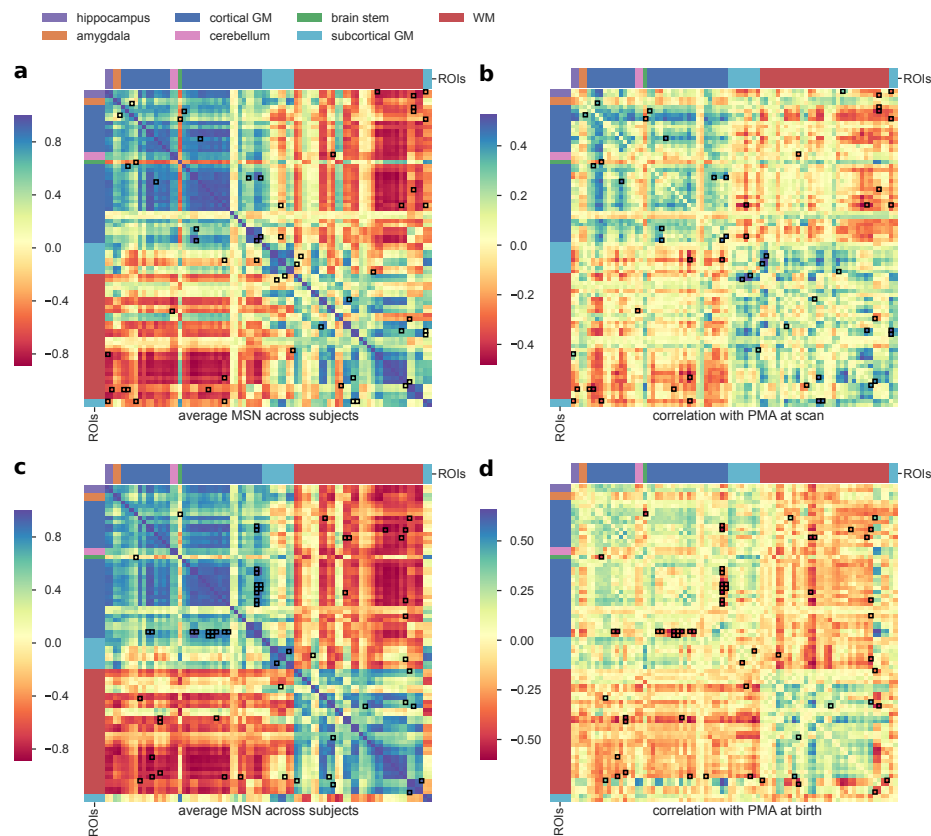


Figure 5: a) Average MSN computed across all subjects using the combination of features selected through the backward feature selection scheme for the age prediction task (Volume, FA, MD, AD, MK, v_{iso} , ODI_P). b) Correlation between each connection weight (inter-regional similarity) shown in (a) and PMA at scan across subjects. c) Average MSN computed across all subjects using the combination of features selected through the backward feature selection scheme for the classification task (Volume, T1/T2, FA, MD, AD, RD, MK, v_{ic} , v_{iso} , ODI_P , ODI_{TOT}). d) Correlation between each connection weight (inter-regional similarity) shown in (c) and PMA at birth across subjects. Connections that were identified as predictive features by the models are highlighted in black. ROIs are ordered as in Supplementary Table S1.

only gender and PMA at birth as independent variables, that achieved a MAE of 1.03 ± 0.88 weeks. A Wilcoxon signed-rank test confirmed that the latter model achieved a significantly greater error ($W = 1633$, $p = 0.0001$). Also models based on single global metrics and single-modality MSNs models provided poorer predictive performance than the selected multi-modality MSNs model (brain volume: MAE= 0.93 ± 0.68 , $R = 0.58$; median FA: MAE= 0.88 ± 0.63 , $R = 0.58$; structural: MAE= 1.08 ± 0.79 , $R = 0.32$; DKI: MAE= 0.94 ± 0.70 , $R = 0.57$; NODDI: MAE= 0.88 ± 0.69 , $R = 0.61$) and this was confirmed by a Wilcoxon signed-rank test (brain volume: $W = 1813$, $p = 0.0019$; median FA: $W = 2045$, $p = 0.0184$; structural: $W = 1361$, $p = 2.76 \times 10^{-06}$; DKI: $W = 1734$, $p = 0.0004$; NODDI: $W = 1811$, $p = 0.0009$). Conversely, the baseline model based on the ensemble on individual metrics used to build the best performing MSN model achieved similar performances (MAE: 0.72 ± 0.56 , $R = 0.77$). A scatter plot of the residuals of the two models (Supplementary Fig. S11) showed a linear trend, indicating that the two models share a similar information content.

Supplementary Fig. S2 shows the results computed with 10-5-fold cross-validation in. All compared models performed similarly under the 10-5-fold scheme, and in general worse than with the LOOCV scheme, with the selected model achieving a MAE of 1 ± 0.2 weeks (Supplementary Fig. S7).

To study which connections contributed the most to chronological age prediction, we selected only edges which were assigned a non-zero coefficient in at least 99% of cross-validation folds. These edges are shown in the chord diagram in Fig. 7 (realised with Circos, Krzywinski et al. (2009)), and are colour coded to distinguish between inter-regional similarities that increase or decrease with age, to highlight networks of regions whose morphological properties are converging (gray) or that tend to differentiate with increasing age (red). Intuitively, these edges connect ROIs whose anatomical and micro-structural properties are changing more than others between 38 and 45 weeks PMA, making the ROIs more or less similar. In other words, it is the relative timing of maturation of different brain tissues to determine the relevance of a connection in the age

prediction task. The selected connections are located in both cortical (frontal, temporal, parietal and occipital lobes; insular and posterior cingulate cortex) and subcortical regions (thalamus, subthalamic and lentiform nuclei), in the brain stem and in the cerebellum. These areas have been previously associated with age-related changes and preterm birth (Boardman et al., 2006; Ball et al., 2013a; Bataille et al., 2017). For comparison, we report in Supplementary Table S2 the regional metrics selected as most predictive of age in the baseline model based on individual metrics.

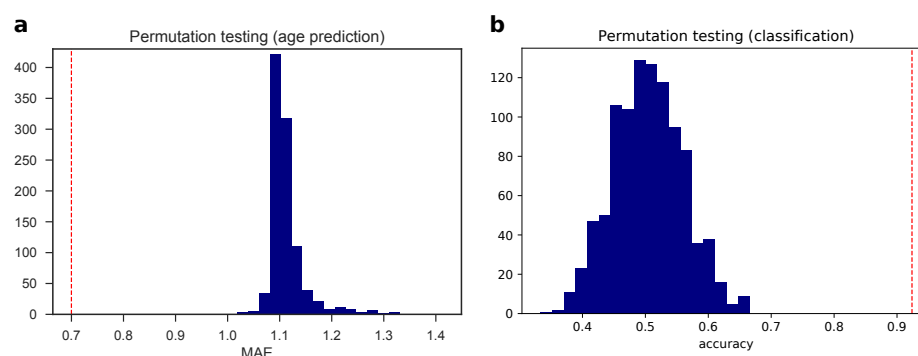


Figure 6: Null distributions computed over 1000 random permutations of the target variable for the age prediction (a) and the classification tasks (b). The red dotted lines indicate the performances of our models.

The best classifier discriminated between term and preterm infants with a 92% LOOCV accuracy (Fig. 6). None of the confounders were included among the selected features. A logistic regression model built on age at scan and gender did not achieve significant accuracy (56%, $p = 0.091$), while adding motion to the predictors produced a 61% accuracy, slightly above chance level ($p = 0.03$), but it should be noted that a model based on motion only was 59% accurate ($p = 0.02$). Models based on global features achieved 55% accuracy for total brain volume and 56% accuracy for median FA. Models built on single data modalities attained 65% accuracy for structural features only, 89% accuracy for DKI features only, and 88% accuracy for NODDI features only. Results computed with 10-5-fold cross-validation are shown in Supplementary Fig. S4.

465 The best classifier selected with LOOCV also achieved top accuracy with 10-5-
466 fold (accuracy 90%, Supplementary Fig. S7).

467 The network of regions that showed the most divergent pattern of structural
468 brain properties in preterm versus term infants comprised the brain stem, the
469 thalamus and the subthalamic nucleus; WM regions in the frontal and insu-
470 lar lobes; GM regions in the occipital lobe; both WM and GM regions in the
471 temporal and parietal lobes and in the posterior cingulate cortex. The chord
472 diagram of edges selected by 99% of the models is shown in Fig. 8, in red where
473 inter-regional similarities are greater in the term group and in gray where they
474 are greater in the preterm group. For comparison, Supplementary Table S3 lists
475 the regional metrics selected by the baseline model based on individual metrics,
476 that obtained a 94% accuracy.

477 3.3. Testing for asymmetry

478 In both chord diagrams (Figs. 7 and 8) we observed more edges in the right
479 hemisphere than in the left one. Both elastic net and SVM models perform a
480 feature selection step to exclude features that are correlated and that carry re-
481 dundant information in order to improve prediction performance, hence it might
482 be the case that the models selected the right connections and discarded the
483 left ones precisely because they had a similar information content. Additionally,
484 in the leave-one-out cross-validation scheme the training sets only differ by two
485 samples in each fold, hence models might be similar across folds.

486 To test the hypothesis that the two hemispheres carry a different information
487 content, we performed two experiments. First, we repeated the same analyses
488 extracting inter-regional similarities from either the right or the left hemisphere.
489 We compared the performance obtained with the regression and classification
490 models on the different subsets of features used in the backward feature se-
491 lection scheme in the main analyses. We found that for the age prediction
492 model a Wilcoxon signed-rank test testing the hypothesis that the prediction
493 error was higher using only connections from the left hemisphere was significant
494 ($W = 156, p = 2.57 \times 10^{-11}$), while there was no statistically significant differ-

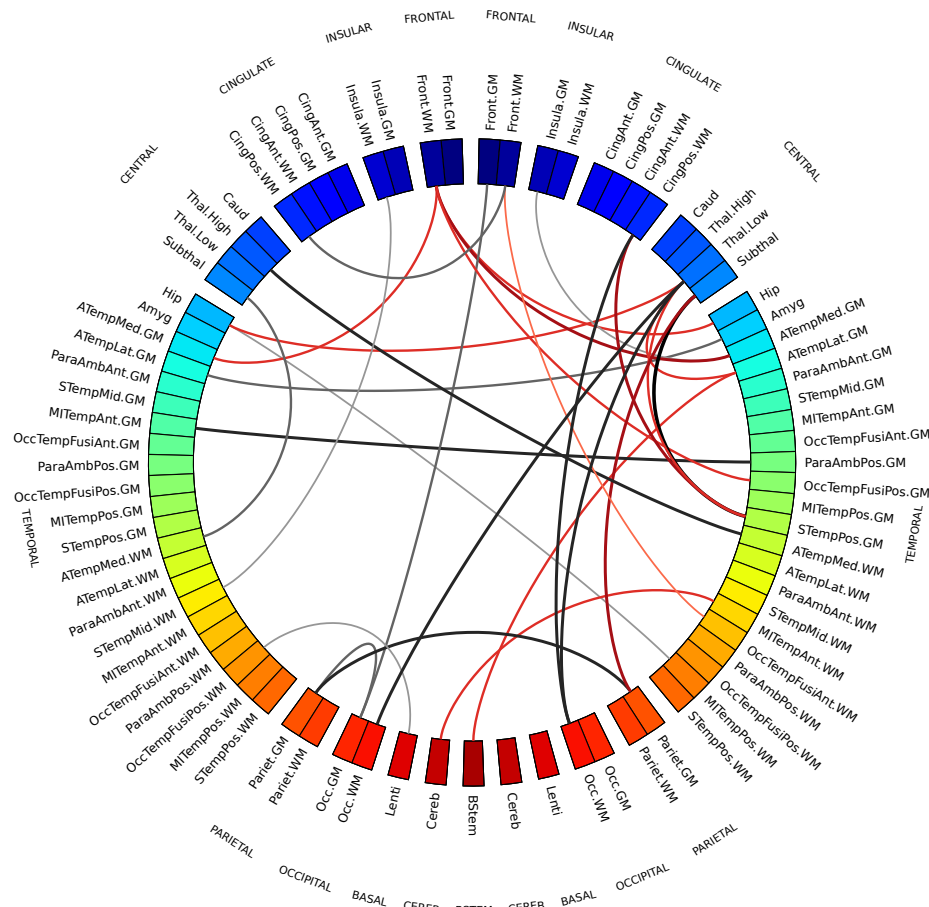


Figure 7: Chord diagram showing MSN edges used for age prediction in at least 99% of regression models in the cross-validation folds. Connections shown in gray are inter-regional similarities that increase with chronological age, while connections in red are inter-regional similarities that decrease with chronological age. The edge width is proportional to the correlation between inter-regional similarities and PMA. The left side of the diagram corresponds to the left side of the brain. Abbreviations for ROI names are explained in Supplementary Table S1.

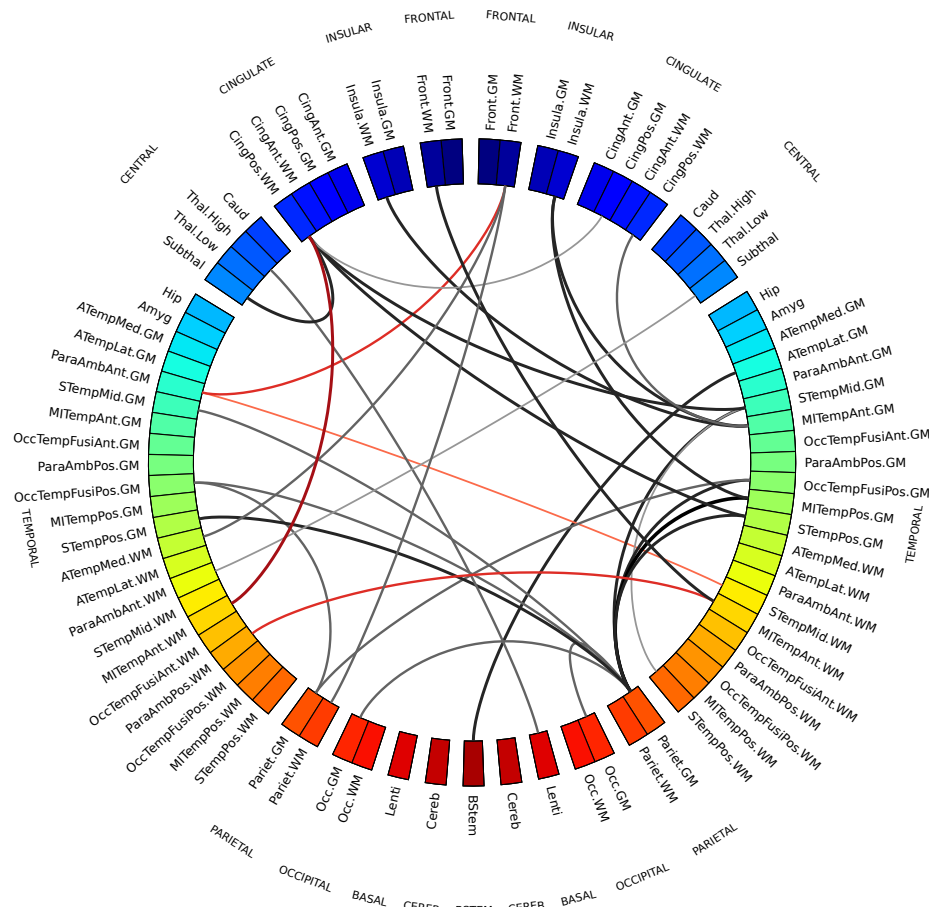


Figure 8: MSN edges showing a divergent pattern of morphological properties in term and preterm infants in at least 99% of classification models in the cross-validation folds. Gray connections indicate inter-regional similarities that are greater in the preterm group, while red connections are greater in the term group. The edge width is proportional to the correlation between inter-regional similarities and prematurity. The left side of the diagram corresponds to the left side of the brain. Abbreviations for ROI names are explained in Supplementary Table S1.

ence in the case of the classification model. These results replicated also when using 10-5-fold cross-validation (age prediction: $W = 160, p = 2.98 \times 10^{-11}$; no significant difference in classification). We also compared the residuals obtained using either the right or the left hemisphere for age prediction with the set of features selected with backward feature selection (Supplementary Fig. S11) and found that the residuals of the fitted models are linearly correlated, suggesting that the two hemispheres do carry a similar information content, but one presents clearer signal than the other. We then used permutation testing to test the “interchangeability” of right and left regions: starting from the subsets of imaging metrics selected in the main analyses for the age prediction and classification models, we generated two null distributions by randomly swapping a subset of homotopic brain regions between the right and left hemisphere, and then repeating the exact same analyses 1000 times. We then counted how many times in the random models there was a disproportion of inter-regional similarities selected in the right hemisphere equal or greater than the one we observed with our models. If the right and left are “interchangeable”, the number of inter-regional similarities selected should remain the same on average. We found that in the age prediction task, under the null distribution, the disproportion of predictive connections in the right hemisphere was associated with a $p = 0.036$, while in the classification task the disproportion was not significant ($p = 0.166$). This implies that at least for age prediction the two hemispheres are not interchangeable, suggesting again that the right hemisphere has a stronger signal. A similar trend was observed under the 10-5-fold cross-validation scheme, but in this case we could not reject the null hypothesis that inter-regional similarities are selected with the same frequency from both hemispheres ($p = 0.098$).

4. Discussion

These results show that the information encoded in MSNs is predictive of chronological brain age in the neonatal period and that MSNs provide a novel data-driven method for investigating neuroanatomic variation associated with

preterm birth. MSNs were built by combining features from different imaging sequences that describe complementary aspects of brain structure that have been previously studied in isolation (Makropoulos et al., 2016; Batalle et al., 2017) and the resulting predictive models achieved a high accuracy for age prediction and classification. By comparing the performance of MSNs features with basic demographic information (age at birth and gender) and simple metrics such as total brain volume and median white matter FA, we also showed that integrating imaging data provides relevant additional information to characterise brain age. Although we cannot exclude the possibility that some of the variability shared with age at birth, gender or brain volume is encoded in the imaging variables, the comparative analysis and the permutation testing results showed that the observed variance cannot be completely explained by demographic variables or simpler metrics alone. However, a high accuracy is not the only goal of the proposed method: once we have determined that the model is able to learn a relationship between the MSN features and age or prematurity, we can interrogate it to find out which features, regions and structures are involved in the predictions, thus allowing for further inferences.

We anticipate that the main clinical and research utilities of MSNs will be to investigate divergent maturational patterns in the context of perinatal environmental, genetic and clinical exposures, leading to improved understanding of antecedents to, and consequences of, atypical brain development. For these purposes a prediction tool with average 5 days error is highly precise compared with other methods for assessing brain maturation, which usually rely upon simple linear regression, use single image features, or broad classifications of prematurity (Toews et al., 2012; Brown et al., 2017; Batalle et al., 2018; Deprez et al., 2018; Bouyssi-Kobar et al., 2018; Ouyang et al., 2018).

The regions identified as most predictive have been previously associated with age-related changes and preterm birth (Boardman et al., 2006; Ball et al., 2013a; Batalle et al., 2017; Bouyssi-Kobar et al., 2018). These data suggest that to fully describe morphological variation in the developing brain it may be advantageous to adopt a holistic approach, leveraging the additional information

that can be derived from integrating multi-contrast MRI data. The main motivation for using a network-based approach is to obtain a whole-brain description of a developmental pattern. By using topologically integrated features instead of single metrics it is possible to access an additional layer of information that is not explicitly encoded in the individual metrics, i.e. how the relationships between metrics vary in different parts of the brain. Working with correlations instead of an ensemble of heterogeneous metrics also aids interpretation, as the focus is shifted from the values of single metrics across the brain, each influenced by disparate factors, to similarities between brain regions, which is a more relatable concept. Additionally, the adoption of a network model has proven to be a useful abstraction to capture the modular organisation of the brain: in the original work introducing MSNs to study microscale cortical organization in adults, the authors demonstrated that regions that were similar in MSNs were more likely to belong to the same cytoarchitectonic class, to be axonally connected and to have high levels of co-expressions of genes specialised for neural functions (Seidlitz et al., 2018). Another reason for working with similarities instead of single regional metrics is methodological: computing edge weights as inter-regional similarities enables an integrated representation of several metrics in a single network; to work with the original features directly would mean either working with several networks (thus requiring a further step to integrate them and aggravating the problems related with the “curse of dimensionality”) or concatenating all the features in a single predictive model (thus excluding the interactions between metrics from the model).

Table 2: Results from previous works in the age prediction task.

	Age span	Model	Error/Accuracy
Brown et al. 2017	27-45 weeks PMA	FA-weighted structural connectivity	MAE = 1.6 weeks
Ouyang et al. 2019	31.5-41.7 weeks PMA	cortical FA and MK (mean kurtosis)	FA: $r = .92$; MK: $r = .63$
Deprez et al. 2018	29-44 weeks PMA	spatio-temporal growth models for myelin-like signals in the thalami and brainstem	Thalami: MAE = 1.41 weeks Brainstem: MAE = 2.56 weeks
Toews et al. 2012	8-590 days from birth	scale-invariant T1w features	MAE = 72 days
Wu et al. 2019	14-48 days from birth	cortical measures	MAE = 11.1 ± 0.3 days

PMA = postmenstrual age, MAE = mean absolute error, r = Pearson’s coefficient between actual and predicted age.

Our data are consistent with previous studies of perinatal brain age prediction based on a single type of data or a single metric. For example, Brown et al. (2017) used dMRI tractography to predict brain dysmaturation in preterm infants with brain injury and abnormal developmental outcome and found that altered connectivity in the posterior cingulate gyrus and the inferior orbitofrontal cortex were associated with a delayed maturation; both of these regions are included in the networks identified by our model. Regional FA, MD, MK, and v_{ic} are each predictive of age (Genc et al., 2017; Karmacharya et al., 2018; Ouyang et al., 2019), and the first three measures were selected in our age prediction model. Growth of the thalami and brainstem, defined in terms of myelin-like signals from T2-weighted images, successfully predicted age between 29 and 44 weeks (Deprez et al., 2018) and these regions are included in the networks most predictive of age in the current study. In Toews et al. (2012), scale-invariant image features were extracted from T1-weighted MRI data of 92 subjects over an age range of 8-590 days to build a developmental model that was used to predict age of new subjects; and Ceschin et al. (2018) proposed a deep learning approach to detect subcortical brain dysmaturation from T2-weighted fast spin echo images in infants with congenital heart disease. Wu et al. (2019) used cortical features extracted from structural images to predict age of 50 healthy subjects with 251 longitudinal MRI scans from 14 to 797 days; in accordance with our results, the regions reported to be important for age prediction were bilateral medial orbitofrontal, parahippocampal, temporal pole, right superior parietal and posterior cingulate cortex. Although our results are not directly comparable with the above works because of the heterogeneity of employed models, validation techniques and population variation (different age ranges), our prediction error is among the lowest reported (see Table 2 for a summary of previous results), but it should be noted that there is a strong positive correlation between the reported MAEs and the age range of the samples. In addition, many works have identified imaging biomarkers associated with preterm birth, such as brain tissue volume (Alexander et al., 2018; Gui et al., 2019), myelin content (Melbourne et al., 2016), and diffusion tensor metrics (Anjari et al.,

2007; Bouyssi-Kobar et al., 2018).

The connections most predictive of age revealed that brain maturation is characterised by morphological convergence of some networks and divergence of others (Fig. 7). These connections mostly involve fronto-temporal and subcortical ROIs, which suggests that the micro- and macro-structural properties of these regions are highly dynamic between 38-45 weeks. Among these, inter-regional similarities within GM and WM increase with age, similarities between cortical GM and WM decrease, while subcortical ROIs become more similar to WM and more dissimilar to cortical GM. This is consistent with previous findings on the different trends in development of the thalamus and the cortex (Eaton-Rosen et al., 2015). Additionally, in a study of early development of structural networks (Batalle et al., 2017), connections to and from deep grey matter are reported to show the most rapid developmental changes between 25-45 weeks, while intra-frontal, frontal to cingulate, frontal to caudate and inter-hemispheric connections are reported to mature more slowly.

Conversely, the inter-regional similarities selected by the SVM classifier to discriminate between term and preterm (Fig. 5) are more distributed across cortical GM and WM and are for the most part greater in the preterm group. The fact that in the term group these cortical ROIs are less homogeneous in terms of structural properties could be interpreted as a sign that in term infants these regions are at a different stage of maturation where their morphological profile is consolidating along specialised developmental trajectories. It has been previously suggested that the rapid maturation of cortical structures occurring in the perinatal period is vulnerable to the effects of preterm birth (Kostović and Jovanov-Milošević, 2006; Ball et al., 2011, 2013b; Smyser et al., 2016b).

The differences between networks identified for age prediction and for preterm classification indicate that atypical brain development after preterm birth is not solely a problem of delayed maturation, but it is characterised by a specific signature. Indeed, while the age prediction networks capture changes occurring in both the preterm and the term group, the classification networks highlight where there are group-wise differences, and they do not match: in the case of a delayed

maturation we would have observed differences in the same regions undergoing age-related changes. MSN variations associated with preterm birth affected brain stem, thalami, sub-thalamic nuclei, WM regions in the frontal and insular lobes, GM regions in the occipital lobe, and WM and GM regions in the temporal and parietal lobes and in the posterior cingulate cortex. This distribution of structural variation is consistent with previous reports of regional alteration in brain volume and dMRI parameters based on single contrasts (Boardman et al., 2006; Bonifacio et al., 2010; Ball et al., 2013a; Brown et al., 2017; Batalle et al., 2017; Alexander et al., 2018; Thompson et al., 2018b; Bouyssi-Kobar et al., 2018). Furthermore, compared to the age prediction model, the MSNs used for preterm classification are based on four additional metrics: T1/T2, related to myelination; RD, measuring water dispersion; v_{ic} describing neurite density; and ODI_{TOT} , associated with the fanning of WM tracts. All these metrics contribute to characterise the micro-structural alterations associated with preterm birth (Eaton-Rosen et al., 2015; Melbourne et al., 2016; Batalle et al., 2018; Thompson et al., 2018b; Bouyssi-Kobar et al., 2018).

We observed a disproportion in the distribution of the connections selected by our models, with a preference for the right hemisphere, hinting at the existence of lateralization in the maturational process. An asymmetry in the development of the right hemisphere in neonates was previously reported in Dubois et al. (2010); Yap et al. (2011); Wu et al. (2019), and our experiments (section 3.3) partially supported the hypothesis that the right hemisphere plays a relevant role in the context of age prediction.

4.1. Limitations

This work has some limitations. First, compared with the original work on MSNs (Seidlitz et al., 2018), we did not have a multi-parametric mapping sequence (Weiskopf et al., 2013); however, because the model is extensible, information from other contrasts could be added and evaluated for their effect on prediction. The MSN model could also be applied to study the properties of cortical gray matter (such as thickness, sulcal depth or curvature), that have been

670 previously reported to be predictive of age in children (Brown et al., 2012) and
 671 could contribute significantly in characterising the newborn brain. However,
 672 metrics that only apply to selected structures (e.g. the cortex) cannot be used
 673 in a whole brain analysis, as to compute inter-regional similarities each region
 674 needs to be described by the same set of metrics. This particular study was
 675 designed based on prior knowledge that typical development and atypical devel-
 676 opment associated with preterm birth are characterised by global changes (Ball
 677 et al., 2013a; Anderson, 2014; Eaton-Rosen et al., 2015; Melbourne et al., 2014),
 678 and MSNs integrating dMRI and sMRI data were chosen to study generalised
 679 processes across the whole brain.

680 Second, we used a motion correction technique that attenuates the impact
 681 of head motion on structural connectivity (Andersson and Sotiropoulos, 2016;
 682 Baum et al., 2018), and we found that scanner motion was not contributing
 683 significantly to prediction accuracy; however we cannot rule out a possible con-
 684 founding effect of motion on the estimation of regional metrics.

685 Third, the preterm study population was representative of survivors of mod-
 686 ern neonatal intensive care in terms of gestational age range and prevalence of
 687 co-morbidities of preterm birth that may influence brain maturation, but it is
 688 still possible that the results were influenced by biological variability specific
 689 to the cohort. A replication study will be required to determine whether the
 690 patterns of dysmaturation we found are generalisable.

691 Finally, we assessed the performance of our models with both LOOCV and
 692 10-5-fold schemes in order to investigate the stability of our findings with respect
 693 to the chosen cross-validation scheme and we observed some variability in the
 694 general trends of the results. The disagreement we found might derive from
 695 the limited size of the training set in the case of the repeated-5-fold scheme (all
 696 models tended to perform worse, suggesting there were not enough samples for
 697 learning), and this was indeed the reason why our first choice was the leave-
 698 one-out scheme. As it is always the case when working with machine learning,
 699 increasing the sample size would increase the power of the models, thereby
 700 reducing the margin of error and the risk of overfitting, with the result that

both schemes should converge to similar findings.

4.2. Conclusions

Combining multiple imaging features in a single model enabled a detailed description of the morphological properties of the developing brain that was used inside a predictive framework to identify two networks of regions: the first, predominantly located in subcortical and fronto-temporal areas, that contributed most to age prediction: the second, comprising mostly frontal, parietal, temporal and insular regions, that discriminated between preterm and term born infant brains. Both predictive models performed best when structural, diffusion tensor-derived and NODDI metrics were combined, which demonstrates the importance of integrating different biomarkers to generate a global picture of the developing human brain. The achieved accuracy supports the hypothesis that studying the interaction between regional metrics can shed light on the mechanics of development.

Morphology, structural connectivity and maturation are all influenced by genetics, co-morbidities of preterm birth, and nutrition (Boardman et al., 2014; Anblagan et al., 2016; Sparrow et al., 2016; Krishnan et al., 2016; Ball et al., 2017; Alexander et al., 2018; Blesa et al., 2019). In future work MSNs could offer new understanding of the impact of these factors on integrated measures of brain development, and the relationship between neonatal MSNs and functional outcome could bring novel insights into the neural bases of cognition and behaviour, by identifying networks of regions associated with later development. MSNs could also enable a direct comparison with functional networks extracted from fMRI, to explore how structure and function interplay in the neonatal period, and study how well the two network models together explain individual variability in developmental outcome.

Conflicts of interest

Authors declare no conflict of interests.

729 Acknowledgements

730 We are grateful to the families who consented to take part in the study. This
731 work was supported by Theirworld (www.theirworld.org) and was undertaken
732 in the MRC Centre for Reproductive Health, which is funded by MRC Centre
733 Grant (MRC G1002033). MJT was supported by NHS Lothian Research and
734 Development Office. Participants were scanned in the University of Edinburgh
735 Imaging Research MRI Facility at the Royal Infirmary of Edinburgh which was
736 established with funding from The Wellcome Trust, Dunhill Medical Trust, Ed-
737 inburgh and Lothians Research Foundation, Theirworld, The Muir Maxwell
738 Trust and many other sources; we thank the University's imaging research staff
739 for providing the infant scanning.

740 References

- 741 Alexander, B., Kelly, C.E., Adamson, C., Beare, R., Zannino, D., Chen, J.,
742 Murray, A.L., Loh, W.Y., Matthews, L.G., Warfield, S.K., Anderson, P.J.,
743 Doyle, L.W., Seal, M.L., Spittle, A.J., Cheong, J.L., Thompson, D.K., 2018.
744 Changes in neonatal regional brain volume associated with preterm birth and
745 perinatal factors. *NeuroImage* doi:10.1016/j.neuroimage.2018.07.021.
- 746 Alexander-Bloch, A., Giedd, J.N., Bullmore, E., 2013. Imaging structural co-
747 variance between human brain regions. *Nature Reviews Neuroscience* 14,
748 322–336. doi:10.1038/nrn3465.
- 749 Anblagan, D., Pataky, R., Evans, M.J., Telford, E.J., Serag, A., Sparrow, S.,
750 Piyasena, C., Semple, S.I., Wilkinson, A.G., Bastin, M.E., Boardman, J.P.,
751 2016. Association between preterm brain injury and exposure to chorioam-
752 nionitis during fetal life. *Scientific Reports* 6, 37932. doi:10.1038/srep37932.
- 753 Anderson, P.J., 2014. Neuropsychological outcomes of children born very
754 preterm. *Seminars in Fetal and Neonatal Medicine* 19, 90–96. doi:10.1016/
755 J.SINY.2013.11.012.

756 Andersson, J.L., Graham, M.S., Drobniak, I., Zhang, H., Filippini, N., Bastiani,
757 M., 2017. Towards a comprehensive framework for movement and distortion
758 correction of diffusion MR images: Within volume movement. *NeuroImage*
759 152, 450–466. doi:10.1016/j.neuroimage.2017.02.085.

760 Andersson, J.L., Graham, M.S., Zsoldos, E., Sotiropoulos, S.N., 2016. Incorporating outlier detection and replacement into a non-parametric framework
761 for movement and distortion correction of diffusion MR images. *NeuroImage*
762 141, 556–572. doi:10.1016/j.neuroimage.2016.06.058.

763 Andersson, J.L., Skare, S., Ashburner, J., 2003. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor
764 imaging. *NeuroImage* 20, 870 – 888. doi:[https://doi.org/10.1016/S1053-8119\(03\)00336-7](https://doi.org/10.1016/S1053-8119(03)00336-7).

765 Andersson, J.L., Sotiropoulos, S.N., 2016. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging.
766 *NeuroImage* 125, 1063–1078. doi:10.1016/j.neuroimage.2015.10.019.

767 Anjari, M., Srinivasan, L., Allsop, J.M., Hajnal, J.V., Rutherford, M.A., Edwards, A.D., Counsell, S.J., 2007. Diffusion tensor imaging with tract-based
768 spatial statistics reveals local white matter abnormalities in preterm infants. *NeuroImage* 35, 1021–1027. doi:10.1016/J.NEUROIMAGE.2007.01.035.

769 Avants, B.B., Tustison, N.J., Song, G., Cook, P.A., Klein, A., Gee, J.C., 2011. A reproducible evaluation of ants similarity metric performance in brain image
770 registration. *NeuroImage* 54, 2033 – 2044. doi:<https://doi.org/10.1016/j.neuroimage.2010.09.025>.

771 Back, S.A., Miller, S.P., 2014. Brain injury in premature neonates: A primary cerebral dysmaturation disorder? *Annals of Neurology* 75, 469–486.

772 Ball, G., Aljabar, P., Arichi, T., Tusor, N., Cox, D., Merchant, N., Nongena, P., Hajnal, J., Edwards, A., Counsell, S., 2016. Machine-learning to characterise

783 neonatal functional connectivity in the preterm brain. *NeuroImage* 124, 267–
784 275. doi:10.1016/J.NEUROIMAGE.2015.08.055.

785 Ball, G., Aljabar, P., Nongena, P., Kennea, N., Gonzalez-Cinca, N., et al., 2017.
786 Multimodal image analysis of clinical influences on preterm brain develop-
787 ment. *Annals of Neurology* 82, 233–246. doi:10.1002/ana.24995.

788 Ball, G., Boardman, J.P., Aljabar, P., Pandit, A., Arichi, T., Merchant, N.,
789 Rueckert, D., Edwards, A.D., Counsell, S.J., 2013a. The influence of preterm
790 birth on the developing thalamocortical connectome. *Cortex* 49, 1711 – 1721.
791 doi:<https://doi.org/10.1016/j.cortex.2012.07.006>.

792 Ball, G., Boardman, J.P., Rueckert, D., Aljabar, P., Arichi, T., Merchant, N.,
793 Gousias, I.S., Edwards, A.D., Counsell, S.J., 2011. The effect of preterm birth
794 on thalamic and cortical development. *Cerebral Cortex* 22, 1016–1024.

795 Ball, G., Srinivasan, L., Aljabar, P., Counsell, S.J., Durighel, G., Hajnal, J.V.,
796 Rutherford, M.A., Edwards, A.D., 2013b. Development of cortical microstruc-
797 ture in the preterm human brain. *Proceedings of the National Academy of*
798 *Sciences* 110, 9541–9546. doi:10.1073/PNAS.1301652110.

799 Barnett, M.L., Tusor, N., Ball, G., Chew, A., Falconer, S., Aljabar, P., Kimpton,
800 J.A., Kennea, N., Rutherford, M., David Edwards, A., Counsell, S.J., 2018.
801 Exploring the multiple-hit hypothesis of preterm white matter damage using
802 diffusion MRI. *NeuroImage: Clinical* 17, 596–606. doi:10.1016/J.NICL.
803 2017.11.017.

804 Bastiani, M., Andersson, J., Cordero-Grande, L., Murgasova, M., Hutter, J.,
805 Price, A.N., Makropoulos, A., Fitzgibbon, S.P., Hughes, E., Rueckert, D.,
806 Victor, S., Rutherford, M., Edwards, A.D., Smith, S., Tournier, J.D., Hajnal,
807 J.V., Jbabdi, S., Sotiropoulos, S.N., 2018. Automated processing pipeline for
808 neonatal diffusion MRI in the developing human connectome project. *Neu-*
809 *roImage* doi:<https://doi.org/10.1016/j.neuroimage.2018.05.064>.

810 Bastiani, M., Cottaar, M., Fitzgibbon, S.P., Suri, S., Alfaro-Almagro, F.,
811 Sotiropoulos, S.N., Jbabdi, S., Andersson, J.L., 2019. Automated quality con-
812 trol for within and between studies diffusion MRI data using a non-parametric
813 framework for movement and distortion correction. *NeuroImage* 184, 801 –
814 812. doi:<https://doi.org/10.1016/j.neuroimage.2018.09.073>.

815 Batalle, D., Edwards, A.D., O’Muircheartaigh, J., 2018b. Annual research re-
816 view: Not just a small adult brain: understanding later neurodevelopment
817 through imaging the neonatal brain. *Journal of Child Psychology and Psy-*
818 *chiatry* 59, 350–371. doi:10.1111/jcpp.12838.

819 Batalle, D., Hughes, E.J., Zhang, H., Tournier, J.D., Tusor, N., others., 2017.
820 Early development of structural networks and the impact of prematurity on
821 brain connectivity. *NeuroImage* 149, 379–392. doi:10.1016/j.neuroimage.
822 2017.01.065.

823 Batalle, D., O’Muircheartaigh, J., Makropoulos, A., Kelly, C.J., Dimitrova,
824 R., Hughes, E.J., Hajnal, J.V., Zhang, H., Alexander, D.C., Edwards, A.D.,
825 Counsell, S.J., 2018. Different patterns of cortical maturation before and after
826 38 weeks gestational age demonstrated by diffusion MRI in vivo. *NeuroImage*
827 doi:<https://doi.org/10.1016/j.neuroimage.2018.05.046>.

828 Baum, G.L., Roalf, D.R., Cook, P.A., Ciric, R., Rosen, A.F., Xia, C., Elliott,
829 M.A., Ruparel, K., Verma, R., Tunc, B., et al., 2018. The impact of in-
830 scanner head motion on structural connectivity derived from diffusion mri.
831 *Neuroimage* 173, 275–286.

832 Blesa, M., Sullivan, G., Anblagan, D., Telford, E.J., Quigley, A.J., Sparrow,
833 S.A., Serag, A., Semple, S.I., Bastin, M.E., Boardman, J.P., 2019. Early
834 breast milk exposure modifies brain connectivity in preterm infants. *NeuroIm-*
835 *age* 184, 431 – 439. doi:[https://doi.org/10.1016/j.neuroimage.2018.09.](https://doi.org/10.1016/j.neuroimage.2018.09.045)
836 045.

837 Boardman, J., Craven, C., Valappil, S., Counsell, S., Dyet, L., Rueckert, D.,
838 Aljabar, P., Rutherford, M., Chew, A., Allsop, J., Cowan, F., Edwards, A.,

2010. A common neonatal image phenotype predicts adverse neurodevelopmental outcome in children born preterm. *NeuroImage* 52, 409 – 414. doi:<https://doi.org/10.1016/j.neuroimage.2010.04.261>.

Boardman, J.P., Counsell, S.J., Rueckert, D., Kapellou, O., Bhatia, K.K., Aljabar, P., Hajnal, J., Allsop, J.M., Rutherford, M.A., Edwards, A.D., 2006. Abnormal deep grey matter development following preterm birth detected using deformation-based morphometry. *NeuroImage* 32, 70–78. doi:<https://doi.org/10.1016/j.neuroimage.2006.03.029>.

Boardman, J.P., Walley, A., Ball, G., Takousis, P., Krishnan, M.L., Hughes-Carre, L., Aljabar, P., Serag, A., King, C., Merchant, N., Srinivasan, L., Froguel, P., Hajnal, J., Rueckert, D., Counsell, S., Edwards, A.D., 2014. Common genetic variants and risk of brain injury after preterm birth. *Pediatrics* 133, e1655–e1663. doi:10.1542/peds.2013–3011.

Bonifacio, S.L., Glass, H.C., Chau, V., Berman, J.I., Xu, D., Brant, R., Barkovich, A.J., Poskitt, K.J., Miller, S.P., Ferriero, D.M., 2010. Extreme premature birth is not associated with impaired development of brain microstructure. *The Journal of Pediatrics* 157, 726–732.e1. doi:10.1016/J.JPeds.2010.05.026.

Bouyssi-Kobar, M., Brossard-Racine, M., Jacobs, M., Murnick, J., Chang, T., Limperopoulos, C., 2018. Regional microstructural organization of the cerebral cortex is affected by preterm birth. *NeuroImage: Clinical* 18, 871–880. doi:10.1016/j.nicl.2018.03.020.

Brown, C.J., Miller, S.P., Booth, B.G., Andrews, S., Chau, V., Poskitt, K.J., Hamarneh, G., 2014. Structural network analysis of brain development in young preterm neonates. *NeuroImage* 101, 667 – 680. doi:<https://doi.org/10.1016/j.neuroimage.2014.07.030>.

Brown, C.J., Moriarty, K.P., Miller, S.P., Booth, B.G., et al., 2017. Prediction of brain network age and factors of delayed maturation in very preterm

867 infants, in: Lecture Notes in Computer Science, pp. 84–91. doi:10.1007/
868 978-3-319-66182-7_10.

869 Brown, T., Kuperman, J., Chung, Y., Erhart, M., McCabe, C., Hagler, D.,
870 Venkatraman, V., Akshoomoff, N., Amaral, D., Bloss, C., Casey, B., Chang,
871 L., Ernst, T., Frazier, J., Gruen, J., Kaufmann, W., Kenet, T., Kennedy,
872 D., Murray, S., Sowell, E., Jernigan, T., Dale, A., 2012. Neuroanatomical
873 Assessment of Biological Maturity. *Current Biology* 22, 1693–1698. doi:10.
874 1016/J.CUB.2012.07.002.

875 Caldinelli, C., Froudust-Walsh, S., Karolis, V., Tseng, C.E., Allin, M.P., Walshe,
876 M., Cuddy, M., Murray, R.M., Nosarti, C., 2017. White matter alterations to
877 cingulum and fornix following very preterm birth and their relationship with
878 cognitive functions. *NeuroImage* 150, 373 – 382. doi:[https://doi.org/10.](https://doi.org/10.1016/j.neuroimage.2017.02.026)
879 1016/j.neuroimage.2017.02.026.

880 Cao, M., Huang, H., He, Y., 2017. Developmental connectomics from infancy
881 through early childhood. *Trends in neurosciences* 40, 494 – 506.

882 Caruyer, E., Lenglet, C., Sapiro, G., Deriche, R., 2013. Design of multishell sam-
883 pling schemes with uniform coverage in diffusion MRI. *Magnetic Resonance*
884 *in Medicine* 69, 1534–1540. doi:10.1002/mrm.24736.

885 Ceschin, R., Zahner, A., Reynolds, W., Gaesser, J., Zuccoli, G., Lo, C.W.,
886 Gopalakrishnan, V., Panigrahy, A., 2018. A computational framework for
887 the detection of subcortical brain dysmaturity in neonatal MRI using 3D
888 Convolutional Neural Networks. *NeuroImage* 178, 183–197. doi:10.1016/J.
889 NEUROIMAGE.2018.05.049.

890 Counsell, S.J., Edwards, A.D., Chew, A.T.M., Cowan, F.M., Boardman, J.P.,
891 Allsop, J.M., Hajnal, J.V., Srinivasan, L., Dyet, L.E., Rutherford, M.A.,
892 Anjari, M., 2008. Specific relations between neurodevelopmental abilities and
893 white matter microstructure in children born preterm. *Brain* 131, 3201–3208.

894 Deprez, M., Wang, S., Ledig, C., Hajnal, J., Counsell, S., Schnabel, J., 2018.
895 Segmentation of myelin-like signals on clinical MR images for age estimation
896 in preterm infants. bioRxiv doi:10.1101/357749.

897 Dubois, J., Benders, M., Lazeyras, F., Borradori-Tolsa, C., Leuchter, R.H.V.,
898 Mangin, J., Hüppi, P., 2010. Structural asymmetries of perisylvian regions in
899 the preterm newborn. *NeuroImage* 52, 32–42. doi:10.1016/J.NEUROIMAGE.
900 2010.03.054.

901 Duerden, E.G., Grunau, R.E., Guo, T., Foong, J., Pearson, A., Au-
902 Young, S., Lavoie, R., Chakravarty, M.M., Chau, V., Synnes, A.,
903 Miller, S.P., 2018. Early procedural pain is associated with regionally-
904 specific alterations in thalamic development in preterm neonates.
905 *Journal of Neuroscience* 38, 878–886. URL: <https://www.jneurosci.org/content/38/4/878>,
906 doi:10.1523/JNEUROSCI.0867-17.2017,
907 arXiv:<https://www.jneurosci.org/content/38/4/878.full.pdf>.

908 Duerden, E.G., Guo, T., Doddiba, L., Chakravarty, M.M., Chau, V., Poskitt,
909 K.J., Synnes, A., Grunau, R.E., Miller, S.P., 2016. Midazolam dose corre-
910 lates with abnormal hippocampal growth and neurodevelopmental outcome in
911 preterm infants. *Annals of Neurology* 79, 548–559. doi:10.1002/ana.24601.

912 Eaton-Rosen, Z., Melbourne, A., Orasanu, E., Cardoso, M.J., Modat, M., Bain-
913 bridge, A., Kendall, G.S., Robertson, N.J., Marlow, N., Ourselin, S., 2015.
914 Longitudinal measurement of the developing grey matter in preterm sub-
915 jects using multi-modal MRI. *NeuroImage* 111, 580–589. doi:10.1016/J.
916 NEUROIMAGE.2015.02.010.

917 Efron, B., 1983. Estimating the error rate of a prediction rule: improvement on
918 cross-validation. *Journal of the American statistical association* 78, 316–331.

919 Genc, S., Malpas, C.B., Holland, S.K., Beare, R., Silk, T.J., 2017. Neurite
920 density index is sensitive to age related differences in the developing brain.
921 *NeuroImage* 148, 373–380. doi:10.1016/j.neuroimage.2017.01.023.

922 Glasser, M.F., Van Essen, D.C., 2011. Mapping human cortical areas in vivo
923 based on myelin content as revealed by T1- and T2-Weighted MRI. *Journal*
924 *of Neuroscience* 31, 11597–11616. doi:10.1523/JNEUROSCI.2180-11.2011.

925 Gousias, I.S., Edwards, A.D., Rutherford, M.A., Counsell, S.J., Hajnal, J.V.,
926 Rueckert, D., Hammers, A., 2012. Magnetic resonance imaging of the newborn
927 brain: Manual segmentation of labelled atlases in term-born and preterm
928 infants. *NeuroImage* 62, 1499 – 1509. doi:[https://doi.org/10.1016/j.](https://doi.org/10.1016/j.neuroimage.2012.05.083)
929 [neuroimage.2012.05.083](https://doi.org/10.1016/j.neuroimage.2012.05.083).

930 Greve, D.N., Fischl, B., 2009. Accurate and robust brain image alignment
931 using boundary-based registration. *NeuroImage* 48, 63–72. doi:10.1016/j.
932 [neuroimage.2009.06.060](https://doi.org/10.1016/j.neuroimage.2009.06.060).

933 Grussu, F., Schneider, T., Tur, C., Yates, R.L., Tachrount, M., Ianuș, A., Yian-
934 nakas, M.C., Newcombe, J., Zhang, H., Alexander, D.C., et al., 2017. Neurite
935 dispersion: a new marker of multiple sclerosis spinal cord pathology? *Annals*
936 *of clinical and translational neurology* 4, 663–679.

937 Gui, L., Loukas, S., Lazeyras, F., Hüppi, P., Meskaldji, D., Borradori Tolsa, C.,
938 2019. Longitudinal study of neonatal brain tissue volumes in preterm infants
939 and their ability to predict neurodevelopmental outcome. *NeuroImage* 185,
940 728–741. doi:10.1016/J.NEUROIMAGE.2018.06.034.

941 Hernandez-Fernandez, M., Reguly, I., Jbabdi, S., Giles, M., Smith, S.,
942 Sotiropoulos, S.N., 2019. Using gpus to accelerate computational diffu-
943 sion MRI: From microstructure estimation to tractography and connectomes.
944 *NeuroImage* 188, 598 – 615. doi:[https://doi.org/10.1016/j.neuroimage.](https://doi.org/10.1016/j.neuroimage.2018.12.015)
945 [2018.12.015](https://doi.org/10.1016/j.neuroimage.2018.12.015).

946 Hunter, J.D., 2007. Matplotlib: A 2d graphics environment. *Computing in*
947 *science & engineering* 9, 90.

948 Jelescu, I.O., Veraart, J., Adisetiyo, V., Milla, S.S., Novikov, D.S.,
949 Fieremans, E., 2015. One diffusion acquisition and different white

950 matter models: How does microstructure change in human early de-
 951 velopment based on wmti and noddi? *NeuroImage* 107, 242
 952 – 256. URL: [http://www.sciencedirect.com/science/article/pii/](http://www.sciencedirect.com/science/article/pii/S1053811914010015)
 953 [S1053811914010015](http://www.sciencedirect.com/science/article/pii/S1053811914010015), doi:[https://doi.org/10.1016/j.neuroimage.2014.](https://doi.org/10.1016/j.neuroimage.2014.12.009)
 954 12.009.

955 Jensen, J., Helpert, J., Ramani, A., Lu, H., Kaczynski, K., 2005. Diffusional
 956 kurtosis imaging: The quantification of nongaussian water diffusion by means
 957 of magnetic resonance imaging. *Magnetic Resonance in Medicine* 53, 1432–
 958 1440. doi:10.1002/mrm.20508.

959 Job, D.E., Dickie, D.A., Rodriguez, D., Robson, A., Danso, S., Per-
 960 net, C., Bastin, M.E., Boardman, J.P., Murray, A.D., Ahearn, T.,
 961 Waiter, G.D., Staff, R.T., Deary, I.J., Shenkin, S.D., Wardlaw, J.M.,
 962 2017. A brain imaging repository of normal structural mri across the
 963 life course: Brain images of normal subjects (brains). *NeuroImage* 144,
 964 299 – 304. URL: [http://www.sciencedirect.com/science/article/pii/](http://www.sciencedirect.com/science/article/pii/S1053811916000331)
 965 [S1053811916000331](http://www.sciencedirect.com/science/article/pii/S1053811916000331), doi:[https://doi.org/10.1016/j.neuroimage.2016.](https://doi.org/10.1016/j.neuroimage.2016.01.027)
 966 01.027. data Sharing Part II.

967 Jones, E., Oliphant, T., Peterson, P., et al., 2001. SciPy: Open source scientific
 968 tools for Python,. <http://www.scipy.org/>.

969 Karmacharya, S., Gagoski, B., Ning, L., Vyas, R., Cheng, H.H., Soul, J., New-
 970 berger, J.W., Shenton, M.E., Rathi, Y., Grant, P.E., 2018. Advanced diffu-
 971 sion imaging for assessing normal white matter development in neonates and
 972 characterizing aberrant development in congenital heart disease. *NeuroImage:*
 973 *Clinical* 19, 360–373. doi:10.1016/j.nicl.2018.04.032.

974 Kellner, E., Dhital, B., Kiselev, V.G., Reiser, M., 2016. Gibbs-ringing artifact
 975 removal based on local subvoxel-shifts. *Magnetic Resonance in Medicine* 76,
 976 1574–1581. doi:10.1002/mrm.26054.

977 Keunen, K., Benders, M.J., Leemans, A., Fieret-Van Stam, P.C., Scholtens,
 978 L.H., Viergever, M.A., Kahn, R.S., Groenendaal, F., de Vries, L.S., van den

979 Heuvel, M.P., 2017. White matter maturation in the neonatal brain is pre-
 980 dictive of school age cognitive capacities in children born very preterm. *De-*
 981 *velopmental Medicine & Child Neurology* 59, 939–946.

982 Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy esti-
 983 mation and model selection, in: *Proceedings of the 14th international joint*
 984 *conference on Artificial intelligence-Volume 2*, Morgan Kaufmann Publishers
 985 Inc.. pp. 1137–1143.

986 Kostović, I., Jovanov-Milošević, N., 2006. The development of cerebral connec-
 987 tions during the first 2045 weeks' gestation. *Seminars in Fetal and Neonatal*
 988 *Medicine* 11, 415–422. doi:10.1016/J.SINY.2006.07.001.

989 Krishnan, M.L., Van Steenwinckel, J., Schang, A.L., Yan, J., Arnadottir, J.,
 990 Le Charpentier, T., Csaba, Z., Dournaud, P., Cipriani, Constance Auvynet,
 991 S., Titomanlio, L., Pansiot, J., Ball, G., Boardman, J.P., Walley, A.J., Saxena,
 992 A., Mirza, G., Fleiss, B., Edwards, A.D., Petretto, E., Gressens, P., 2017.
 993 Integrative genomics of microglia implicates *dlg4* (*psd95*) in the white matter
 994 development of preterm infants. *Nature Communications* 8.

995 Krishnan, M.L., Wang, Z., Silver, M., Boardman, J.P., Ball, G., Coun-
 996 sell, S.J., Walley, A.J., Montana, G., Edwards, A.D., 2016. Possible re-
 997 lationship between common genetic variation and white matter develop-
 998 ment in a pilot study of preterm infants. *Brain and behavior* 6, e00434.
 999 doi:10.1002/brb3.434.

1000 Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D.,
 1001 Jones, S.J., Marra, M.A., 2009. Circos: an information aesthetic for compar-
 1002 ative genomics. *Genome research* 19, 1639–45. doi:10.1101/gr.092759.109.

1003 Kulikova, S., Hertz-Pannier, L., Dehaene-Lambertz, G., Buzmakov, A., Poupon,
 1004 C., Dubois, J., 2015. Multi-parametric evaluation of the white matter
 1005 maturation. *Brain Structure and Function* 220, 3657–3672. doi:10.1007/
 1006 s00429-014-0881-y.

1007 Kunz, N., Zhang, H., Vasung, L., O'Brien, K.R., Assaf, Y., Lazeyras, F., Alexan-
1008 der, D.C., Hüppi, P.S., 2014. Assessing white matter microstructure of the
1009 newborn with multi-shell diffusion MRI and biophysical compartment models.
1010 *NeuroImage* 96, 288–299. doi:10.1016/j.neuroimage.2014.03.057.

1011 Leuchter, R.H.V., Gui, L., Poncet, A., Hagmann, C., Lodygensky, G.A., Martin,
1012 E., Koller, B., Darqu, A., Bucher, H.U., Hüppi, P.S., 2014. Association Between
1013 Early Administration of High-Dose Erythropoietin in Preterm Infants and
1014 Brain MRI Abnormality at Term-Equivalent Age. *JAMA* 312, 817–824.

1015 Li, W., Yang, C., Shi, F., Wu, S., Wang, Q., Nie, Y., Zhang, X., 2017. Construc-
1016 tion of individual morphological brain networks with multiple morphometric
1017 features. *Frontiers in Neuroanatomy* 11. doi:10.3389/fnana.2017.00034.

1018 Li, X., Morgan, P.S., Ashburner, J., Smith, J., Rorden, C., 2016. The first step
1019 for neuroimaging data analysis: Dicom to nifti conversion. *Journal of Neuro-*
1020 *science Methods* 264, 47 – 56. doi:[https://doi.org/10.1016/j.jneumeth.](https://doi.org/10.1016/j.jneumeth.2016.03.001)
1021 2016.03.001.

1022 Mahjoub, I., Mahjoub, M.A., Rekik, I., 2018. Brain multiplexes reveal mor-
1023 phological connectional biomarkers fingerprinting late brain dementia states.
1024 *Scientific Reports* 8, 4103. doi:10.1038/s41598-018-21568-7.

1025 Makropoulos, A., Aljabar, P., Wright, R., Hüning, B., Merchant, N., et al., 2016.
1026 Regional growth and atlasing of the developing human brain. *NeuroImage*
1027 125, 456–478. doi:10.1016/j.neuroimage.2015.10.047.

1028 Makropoulos, A., Gousias, I.S., Ledig, C., Aljabar, P., Serag, A., Hajnal, J.V.,
1029 Edwards, A.D., Counsell, S.J., Rueckert, D., 2014. Automatic whole brain
1030 MRI segmentation of the developing neonatal brain. *IEEE Transactions on*
1031 *Medical Imaging* 33, 1818–1831. doi:10.1109/TMI.2014.2322280.

1032 Makropoulos, A., Robinson, E.C., Schuh, A., Wright, R., Fitzgibbon, S., et al.,
1033 2018. The developing human connectome project: A minimal processing

1034 pipeline for neonatal cortical surface reconstruction. *NeuroImage* 173, 88–
1035 112. doi:10.1016/j.neuroimage.2018.01.054.

1036 Marcus, D., Harwell, J., Olsen, T., Hodge, M., Glasser, M., Prior, F., Jenkinson,
1037 M., Laumann, T., Curtiss, S., Van Essen, D., 2011. Informatics and data
1038 mining tools and strategies for the human connectome project. *Frontiers in*
1039 *Neuroinformatics* 5, 4.

1040 Mathewson, K., Chow, C., Dobson, K., Pope, E., Schmidt, L., Van Lieshout,
1041 R., 2017. Mental health of extremely low birth weight survivors: A systematic
1042 review and meta-analysis. *Psychological Bulletin* 143, 347 – 383.

1043 Maximov, I.I., Alnaes, D., Westlye, L.T., 2019. Towards an optimised processing
1044 pipeline for diffusion MRI data: Effects of artefact corrections on diffusion
1045 metrics and their age associations in UK Biobank. *bioRxiv* .

1046 McKinney, W., et al., 2010. Data structures for statistical computing in python,
1047 in: *Proceedings of the 9th Python in Science Conference*, Austin, TX. pp. 51–
1048 56.

1049 Melbourne, A., Eaton-Rosen, Z., Orasanu, E., Price, D., Bainbridge, A., Car-
1050 doso, M.J., Kendall, G.S., Robertson, N.J., Marlow, N., Ourselin, S., 2016.
1051 Longitudinal development in the preterm thalamus and posterior white mat-
1052 ter: MRI correlations between diffusion weighted imaging and T2 relaxome-
1053 try. *Human Brain Mapping* 37, 2479–2492.

1054 Melbourne, A., Kendall, G.S., Cardoso, M.J., Gunny, R., Robertson, N.J., Mar-
1055 low, N., Ourselin, S., 2014. Preterm birth affects the developmental synergy
1056 between cortical folding and cortical connectivity observed on multimodal
1057 MRI. *NeuroImage* 89, 23–34. doi:10.1016/J.NEUROIMAGE.2013.11.048.

1058 Nosarti, C., Reichenberg, A., Murray, R.M., Cnattingius, S., Lambe, M.P.,
1059 Yin, L., MacCabe, J., Rifkin, L., Hultman, C.M., 2012. Preterm Birth and
1060 Psychiatric Disorders in Young Adult Life. *Archives of General Psychiatry*
1061 69, 610–617.

1062 Otsuka, Y., Chang, L., Kawasaki, Y., Wu, D., Ceritoglu, C., Oishi, K., Ernst,
1063 T., Miller, M., Mori, S., Oishi, K., 2019. A multi-atlas label fusion tool for
1064 neonatal brain mri parcellation and quantification. *Journal of Neuroimaging*
1065 .

1066 Ouyang, M., Dubois, J., Yu, Q., Mukherjee, P., Huang, H., 2018. Delineation
1067 of early brain development from fetuses to infants with diffusion MRI and
1068 beyond. *NeuroImage* doi:10.1016/j.neuroimage.2018.04.017.

1069 Ouyang, M., Jeon, T., Sotiras, A., Peng, Q., Mishra, V., Halovanic, C., Chen,
1070 M., Chalak, L., Rollins, N., Roberts, T.P.L., Davatzikos, C., Huang, H.,
1071 2019. Differential cortical microstructural maturation in the preterm human
1072 brain with diffusion kurtosis and tensor imaging. *Proceedings of the National*
1073 *Academy of Sciences* , 201812156doi:10.1073/PNAS.1812156116.

1074 Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O.,
1075 Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al., 2011. Scikit-
1076 learn: Machine learning in python. *Journal of machine learning research* 12,
1077 2825–2830.

1078 Schneider, J., Fischer Fumeaux, C.J., Duerden, E.G., Guo, T., Foong, J.,
1079 Graz, M.B., Hagmann, P., Chakravarty, M.M., Hüppi, P.S., Beauport, L.,
1080 Truttmann, A.C., Miller, S.P., 2018. Nutrient intake in the first two weeks
1081 of life and brain growth in preterm neonates. *Pediatrics* 141. doi:10.1542/
1082 *peds*.2017–2169.

1083 Seabold, S., Perktold, J., 2010. Statsmodels: Econometric and statistical mod-
1084 eling with python, in: *Proceedings of the 9th Python in Science Conference*,
1085 *Scipy*. p. 61.

1086 Seidlitz, J., Váša, F., Shinn, M., Romero-Garcia, R., Whitaker, K.J., et al.,
1087 2018. Morphometric similarity networks detect microscale cortical organiza-
1088 tion and predict inter-individual cognitive variation. *Neuron* 97, 231–247.e7.
1089 doi:10.1016/j.neuron.2017.11.039.

1090 Serag, A., Aljabar, P., Ball, G., Counsell, S.J., Boardman, J.P., Rutherford,
1091 M.A., Edwards, A.D., Hajnal, J.V., Rueckert, D., 2012. Construction of a
1092 consistent high-definition spatio-temporal atlas of the developing brain using
1093 adaptive kernel regression. *NeuroImage* 59, 2255 – 2265. doi:<https://doi.org/10.1016/j.neuroimage.2011.09.062>.
1094

1095 Shi, F., Yap, P.T., Gao, W., Lin, W., Gilmore, J.H., Shen, D., 2012. Altered
1096 structural connectivity in neonates at genetic risk for schizophrenia: A com-
1097 bined study using morphological and white matter networks. *NeuroImage* 62,
1098 1622–1633. doi:10.1016/j.neuroimage.2012.05.026.

1099 Smith, S.M., 2002. Fast robust automated brain extraction. *Human Brain*
1100 *Mapping* 17, 143–155. doi:10.1002/hbm.10062.

1101 Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.,
1102 Johansen-Berg, H., Bannister, P.R., Luca, M.D., Drobnjak, I., Flitney,
1103 D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y., Stefano, N.D.,
1104 Brady, J.M., Matthews, P.M., 2004. Advances in functional and struc-
1105 tural mr image analysis and implementation as fsl. *NeuroImage* 23, S208
1106 – S219. doi:<https://doi.org/10.1016/j.neuroimage.2004.07.051>. math-
1107 ematics in Brain Imaging.

1108 Smyser, C.D., Dosenbach, N.U., Smyser, T.A., Snyder, A.Z., Rogers, C.E.,
1109 Inder, T.E., Schlaggar, B.L., Neil, J.J., 2016a. Prediction of brain maturity
1110 in infants using machine-learning algorithms. *NeuroImage* 136, 1–9. doi:10.
1111 1016/J.NEUROIMAGE.2016.05.029.

1112 Smyser, T.A., Smyser, C.D., Rogers, C.E., Gillespie, S.K., Inder, T.E., Neil, J.J.,
1113 2016b. Cortical gray and adjacent white matter demonstrate synchronous
1114 maturation in very preterm infants. *Cerebral Cortex* 26, 3370–3378. doi:10.
1115 1093/cercor/bhv164.

1116 Soussia, M., Rekik, I., 2018. Unsupervised manifold learning using high-order
1117 morphological brain networks derived from T1-w MRI for autism diagnosis.
1118 *Frontiers in Neuroinformatics* 12, 70. doi:10.3389/fninf.2018.00070.

1119 Sparrow, S., Manning, J.R., Cartier, J., Anblagan, D., Bastin, M.E., Piyasena,
1120 C., Pataky, R., Moore, E.J., Semple, S.I., Wilkinson, A.G., Evans, M., Drake,
1121 A.J., Boardman, J.P., 2016. Epigenomic profiling of preterm infants reveals
1122 DNA methylation differences at sites associated with neural function. *Trans-*
1123 *lational psychiatry* 6, e716. doi:10.1038/tp.2015.210.

1124 Steven, A.J., Zhuo, J., Melhem, E.R., 2014. Diffusion kurtosis imaging: an
1125 emerging technique for evaluating the microstructural environment of the
1126 brain. *AJR. American journal of roentgenology* 202 1, W26–33.

1127 Tariq, M., Schneider, T., Alexander, D.C., Wheeler-Kingshott, C.A.G., Zhang,
1128 H., 2016. Bingham-NODDI: Mapping anisotropic orientation dispersion of
1129 neurites using diffusion MRI. *NeuroImage* 133, 207 – 223. doi:<https://doi.org/10.1016/j.neuroimage.2016.01.046>.

1131 Telford, E.J., Cox, S.R., Fletcher-Watson, S., Anblagan, D., Sparrow, S.,
1132 Pataky, R., Quigley, A., Semple, S.I., Bastin, M.E., Boardman, J.P., 2017.
1133 A latent measure explains substantial variance in white matter microstruc-
1134 ture across the newborn human brain. *Brain Structure and Function* 222,
1135 4023–4033.

1136 Thompson, D.K., Chen, J., Beare, R., Adamson, C.L., Ellis, R., Ahmadzai,
1137 Z.M., Kelly, C.E., Lee, K.J., Zalesky, A., Yang, J.Y., Hunt, R.W., Cheong,
1138 J.L., Inder, T.E., Doyle, L.W., Seal, M.L., Anderson, P.J., 2016. Structural
1139 connectivity relates to perinatal factors and functional impairment at 7years
1140 in children born very preterm. *NeuroImage* 134, 328 – 337. doi:<https://doi.org/10.1016/j.neuroimage.2016.03.070>.

1142 Thompson, D.K., Kelly, C.E., Chen, J., Beare, R., Alexander, B., Seal, M.L.,
1143 Lee, K., Matthews, L.G., Anderson, P.J., Doyle, L.W., Spittle, A.J., Cheong,
1144 J.L., 2018a. Early life predictors of brain development at term-equivalent
1145 age in infants born across the gestational age spectrum. *NeuroImage* doi:10.
1146 1016/j.neuroimage.2018.04.031.

1147 Thompson, D.K., Kelly, C.E., Chen, J., Beare, R., Alexander, B., Seal, M.L.,
1148 Lee, K.J., Matthews, L.G., Anderson, P.J., Doyle, L.W., Cheong, J.L., Spit-
1149 tle, A.J., 2018b. Characterisation of brain volume and microstructure at
1150 term-equivalent age in infants born across the gestational age spectrum. *Neu-
1151 roImage: Clinical* , 101630.

1152 Toews, M., Wells, W.M., Zöllei, L., 2012. A feature-based developmental model
1153 of the infant brain in structural MRI, in: *International Conference on Medical
1154 Image Computing and Computer-Assisted Intervention*, Springer. pp. 204–
1155 211.

1156 Toulmin, H., Beckmann, C.F., O’Muircheartaigh, J., Ball, G., Nongena, P.,
1157 Makropoulos, A., Ederies, A., Counsell, S.J., Kennea, N., Arichi, T., Tu-
1158 sor, N., Rutherford, M.A., Azzopardi, D., Gonzalez-Cinca, N., Hajnal, J.V.,
1159 Edwards, A.D., 2015. Specialization and integration of functional thalamocor-
1160 tical connectivity in the human infant. *Proceedings of the National Academy
1161 of Sciences* 112, 6485–6490.

1162 Tournier, J.D., Smith, R.E., Raffelt, D.A., Tabbara, R., Dhollander, T., Pietsch,
1163 M., Christiaens, D., Jeurissen, B., Yeh, C.H., Connelly, A., 2019. Mrtrix3: A
1164 fast, flexible and open software framework for medical image processing and
1165 visualisation. *bioRxiv* doi:10.1101/551739.

1166 Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A.,
1167 Gee, J.C., 2010. N4ITK: Improved N3 bias correction. *IEEE Transactions on
1168 Medical Imaging* 29, 1310–1320. doi:10.1109/TMI.2010.2046908.

1169 Van Den Heuvel, M.P., Kersbergen, K.J., De Reus, M.A., Keunen, K., et al.,
1170 2015. The neonatal connectome during preterm brain development. *Cerebral
1171 Cortex* 25, 3000–3013. doi:10.1093/cercor/bhu095.

1172 Van Der Walt, S., Colbert, S.C., Varoquaux, G., 2011. The numpy array:
1173 a structure for efficient numerical computation. *Computing in Science &
1174 Engineering* 13, 22.

1175 Van Lieshout, R.J., Ferro, M.A., Schmidt, L.A., Boyle, M.H., Saigal, S., Mor-
1176 rison, K.M., Mathewson, K.J., 2018. Trajectories of psychopathology in ex-
1177 tremely low birth weight survivors from early adolescence to adulthood: a
1178 20-year longitudinal study. *Journal of Child Psychology and Psychiatry* 59,
1179 1192–1200.

1180 Varoquaux, G., 2018. Cross-validation failure: small sample sizes lead to large
1181 error bars. *Neuroimage* 180, 68–77.

1182 Varoquaux, G., Raamana, P.R., Engemann, D.A., Hoyos-Idrobo, A., Schwartz,
1183 Y., Thirion, B., 2017. Assessing and tuning brain decoders: cross-validation,
1184 caveats, and guidelines. *NeuroImage* 145, 166–179.

1185 Veraart, J., Fieremans, E., Novikov, D.S., 2016. Diffusion MRI noise mapping
1186 using random matrix theory. *Magnetic Resonance in Medicine* 76, 1582–1593.
1187 doi:10.1002/mrm.26059.

1188 Veraart, J., Novikov, D.S., Christiaens, D., Ades-aron, B., Sijbers, J., Fiere-
1189 mans, E., 2016b. Denoising of diffusion MRI using random matrix theory.
1190 *NeuroImage* 142, 394–406. doi:10.1016/j.neuroimage.2016.08.016.

1191 Weiskopf, N., Suckling, J., Williams, G., Correia, M.M., Inkster, B., Tait, R.,
1192 Ooi, C., Bullmore, E.T., Lutti, A., 2013. Quantitative multi-parameter map-
1193 ping of R1, PD*, MT, and R2* at 3T: a multi-center validation. *Frontiers in*
1194 *Neuroscience* 7. doi:10.3389/fnins.2013.00095.

1195 Woodward, L.J., Anderson, P.J., Austin, N.C., Howard, K., Inder, T.E., 2006.
1196 Neonatal mri to predict neurodevelopmental outcomes in preterm infants.
1197 *New England Journal of Medicine* 355, 685–694.

1198 Wu, Z., Li, G., Shen, D., Hu, D., Lin, W., 2019. Hierarchical rough-to-fine
1199 model for infant age prediction based on cortical features. *IEEE Journal of*
1200 *Biomedical and Health Informatics* , 1–1doi:10.1109/jbhi.2019.2897020.

- 1201 Yap, P.T., Fan, Y., Chen, Y., Gilmore, J.H., Lin, W., Shen, D., 2011. Develop-
1202 ment trends of white matter connectivity in the first years of life. PLoS ONE
1203 6, e24678. doi:10.1371/journal.pone.0024678.
- 1204 Zhang, H., Schneider, T., Wheeler-Kingshott, C.A., Alexander, D.C., 2012.
1205 NODDI: Practical in vivo neurite orientation dispersion and density imaging
1206 of the human brain. NeuroImage 61, 1000–1016. doi:10.1016/j.neuroimage.
1207 2012.03.072.
- 1208 Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic
1209 net. Journal of the Royal Statistical Society: Series B (Statistical Methodol-
1210 ogy) 67, 301–320.