

OPEN

Quantitative assessment of myelination patterns in preterm neonates using T2-weighted MRI

Siying Wang¹, Christian Ledig², Joseph V. Hajnal³, Serena J. Counsell³, Julia A. Schnabel³ & Maria Deprez³

Myelination is considered to be an important developmental process during human brain maturation and closely correlated with gestational age. Quantitative assessment of the myelination status requires dedicated imaging, but the conventional T₂-weighted scans routinely acquired during clinical imaging of neonates carry signatures that are thought to be associated with myelination. In this work, we develop a quantitative marker of progressing myelination for assessment preterm neonatal brain maturation based on novel automatic segmentation method for myelin-like signals on T₂-weighted magnetic resonance images. Firstly we define a segmentation protocol for myelin-like signals. We then develop an expectation-maximization framework to obtain the automatic segmentations of myelin-like signals with explicit class for partial volume voxels whose locations are configured in relation to the composing pure tissues via second-order Markov random fields. The proposed segmentation achieves high Dice overlaps of 0.83 with manual annotations. The automatic segmentations are then used to track volumes of myelinated tissues in the regions of the central brain structures and brainstem. Finally, we construct a spatio-temporal growth models for myelin-like signals, which allows us to predict gestational age at scan in preterm infants with root mean squared error 1.41 weeks.

Human brain maturation involves a complex series of morphological, structural and functional changes. Among these changes is the process of myelin growth and axon ensheathment known as myelination, which facilitates electrical conduction in the neural system¹. Process of myelination is initiated during intra-uterine development, progressing from brainstem and deep structures, which are partly myelinated prior to birth, towards peripheral brain regions (e.g. cerebral white matter tracts) that mostly myelinate during the first year, and continue throughout life¹⁻⁴. Patterns of myelination were initially observed through histological studies^{2,5} and more recently, T₁- and T₂-weighted Magnetic Resonance Imaging (MRI) has been established as the primary method for qualitative assessment of myelination during prenatal and early postnatal period^{1,4}.

Preterm birth and other perinatal insults often lead to injury to immature WM and subsequent hypomyelination^{6,7}. Impaired myelination in posterior limb of internal capsule (PLIC) is considered a sign of developmental delay in the newborn⁸ and is one of the cerebral abnormalities observed on MRI of preterm infants at term-equivalent age linked to very preterm birth⁹. Previous work has also suggested that absolute volume of myelinated white matter (WM) at term equivalent age is decreased in preterm infants¹⁰. Diffusion MRI has shown decrease in fractional anisotropy (FA) in PLIC in preterm infants at term equivalent age¹¹ which may reflect either reduction in myelination or decreased coherence of WM fibers¹². A recent study of a small cohort of preterm infants suggest that decrease in FA is accompanied by increase in T₂ relaxation times in very preterm infants compared to late preterm infants at term equivalent age¹³. White matter abnormalities observed on MRI of preterm infants have been linked to neurodevelopmental impairment, such as motor and cognitive delay¹⁴⁻¹⁶.

Progressing myelination results in shortening of T₁ and T₂ relaxation times^{1,3,17}. However, their time courses are different. T₁ shortening starts already in “pre-myelination” stage and is thought to be linked to the bounding of free water molecules to the accumulating building blocks of myelin (cholesterol and glycolipids). On the other hand, shortening of T₂ relaxation times seems to correspond to tightening of the myelin sheath around the axons and consequent redistribution of free water, reflecting true maturation of the myelin^{3,17}. T₁ and T₂-weighted

¹Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, UK.

²Department of Computing, Imperial College London, London, UK. ³School of Imaging Sciences & Biomedical Engineering, King's College London, London, UK. Correspondence and requests for materials should be addressed to M.D. (email: maria.murgasova@kcl.ac.uk)

weighted MRI remain the most common modalities for imaging neonatal brain^{18,19} and qualitative descriptions of myelination using these modalities^{20,21} are consistent with histological observations². Though specialised sequences for imaging of myelin have been proposed (e.g. myelin water imaging²², magnetization transfer imaging²³), they are not routinely available and their application in preterm and neonatal population would currently be limited to small prospective studies. On the other hand, availability of quantitative markers of myelin using standard T₂-weighted imaging would facilitate quantitative evaluation of mature myelin in large-scale studies of neurodevelopment in preterm and neonatal populations.

In this study, we refer to the tissue that is likely to contain myelin in T₂-weighted neonatal brain MRI as myelin-like signals (MLS). One way of developing a quantitative markers of myelination is to develop an automatic method for segmentation of MLS, that can be further utilised in volumetric studies and spatio-temporal modelling of progressing myelination. Automatic segmentation of MLS is challenging and there is currently a lack of dedicated methods for segmenting MLS during preterm and neonatal period. Recently developed methods for neonatal brain segmentation in the NeoBrainS12 challenge^{24–28} all showed great promise in segmenting a number of brain structures in neonatal brain MRI. However, none of them performed well in segmenting myelination²⁹. Most neonatal brain segmentation methods adapt the approaches developed for adults^{30–34}, and use a probabilistic atlas or manual annotations to obtain prior information on the expected tissue locations. However, myelin is not included in any of the existing neonatal brain atlases^{35–37} or manual annotation database³⁸, and neither it is considered in segmentation protocol in Developing Human Connectome Project³⁹. Another issue that complicates accurate segmentation of MLS in perinatal period is its small volume compared to resolution of the MRI. Partial volume (PV) effect (two or more tissues mixing in a single voxel) therefore needs to be taken into account. Many methods for PV modeling were previously proposed^{40–45}, however they all require prior information about location of mixing tissues, which is currently not available for MLS.

In this paper we propose an automatic method of MLS on T₂-weighted neonatal brain images that does not require any probabilistic atlas or manual annotation of myelin. Similarly to Ledig *et al.*⁴¹ we model PV between MLS and background (BKG) voxels using second-order Markov random fields (MRFs) within an expectation-maximization (EM) framework. We introduce an explicit PV class whose locations are configured in relation to MLS and BKG using a 3D connectivity tensor. We distinguish between MLS and other tissue with similar intensity ranges by defining an anatomical region of interest (ROI) for MLS based on anatomical knowledge. Our method achieves automatic MLS segmentations of high Dice coefficients (DCs)⁴⁶ with respect to the manual annotations for 16 preterm infants at one-week intervals between 29 and 44 weeks gestational age (GA). The proposed segmentation method is applied to T₂-weighted scans of 114 preterm infants to develop quantitative markers of myelination. Firstly, we perform volumetric analysis of the progressing myelination and show that myelination increases in deep brain region while it is stable in brainstem. Secondly, we build a spatio-temporal model of progressing myelination, which we compare to the qualitative studies of myelination^{4,21}. Finally we show that the spatio-temporal atlas of progressing myelination enables us to predict gestational age at scan with high accuracy and is thus a potential quantitative marker of developmental delay.

Results

Volumetric analysis of progressing myelination. Automatic segmentation of the MLS was applied to 114 T₂-weighted neonatal scans. The details of the subjects, acquisition protocol and the segmentation method can be found in the Methods section. The volumes calculated from the resulting segmentations in the deep brain region and brainstem were plotted against GA in Fig. 1. We fitted exponential trend lines with equation $y = ae^{bx}$ and calculated R^2 scores.

We found that the MLS volume appeared to grow exponentially with GA in the deep brain region ($R^2 = 0.94$), and this trend persisted in the MLS volume fraction after we corrected for the different deep brain region sizes in individual subjects ($R^2 = 0.87$). In contrast, the MLS volume appeared to grow nearly linearly with GA in the brainstem ($R^2 = 0.53$ for exponential model, $R^2 = 0.52$ for linear model), which resulted in an almost constant MLS volume fraction after correcting for the different brainstem sizes in individual subjects ($R^2 = 0.25$ for exponential model, $R^2 = 0.24$ for linear; very small $b = -0.015$ suggest nearly constant or slightly negative trend). This is consistent with the previous findings that most of the brainstem is myelinated before 29 weeks GA, whereas new regions become myelinated in the deep brain between 29 and 44 weeks GA^{21,47,48}. This suggest that the MLS volume increased in the brainstem mainly due to the overall brain growth, and the increase in the deep brain region was attributed to both brain growth and new myelination.

Spatio-temporal model of progressing myelination. Segmented images were registered into common reference space, and we fitted a modified logistic function to each voxel location to create spatio-temporal model of progressing myelination. The resulting model for MLS in deep brain region is presented in Fig. 2. The details of the spatio-temporal modelling can be found in Methods section and our previous work⁴⁹. Note that because we corrected for the different sizes of each ROI in individual subjects through non-rigid registration⁵⁰, the MLS progression depicted by the spatio-temporal model was entirely due to the appearance of new myelinated brain structures without contribution from the overall brain growth.

We found that there was an evident difference in the onset and rate of myelination in deep brain region and brainstem. The model of MLS in deep brain region (Fig. 2) showed that the ventro-lateral nuclei (VLN) and sub-thalamic nucleus (STN) were myelinated before 29 weeks GA. Moreover, the model captured the arrival of MLS in both of the posterior limbs of internal capsule (PLIC) at 40 weeks GA, which confirmed the clinically observed time point^{4,21}. Myelination of the PLIC tracts is an important landmark for evaluating neonatal brain development⁵¹. The brainstem, however, appeared to be well developed by the time of 29 weeks GA, and there was

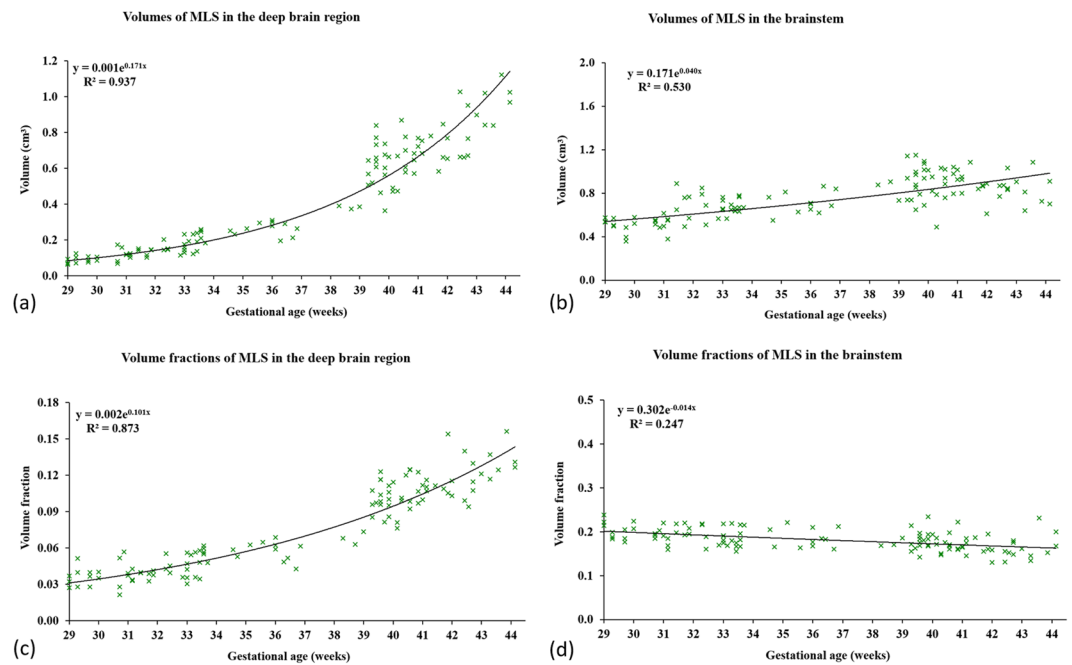


Figure 1. (a,b) Volumes of the automatic segmentations for myelin-like signals (MLS) in the deep brain region and brainstem obtained using the proposed segmentation method, (c,d) volume fractions of MLS plotted against the gestational ages at scan (GAs) of 114 subjects between 29 and 44 weeks GA. In deep brain region the MLS volume appears to grow exponentially with GA, and this trend persists in the MLS volume fraction after correcting for the different sizes of the deep brain region in individual subjects. In brainstem the trend for MLS volume appears nearly linear with GA, which results in an almost constant MLS volume fraction after correcting for the different sizes of the brainstem in individual subjects.

no new myelination in the age range of our investigation. We summarise these milestones in Table 1, and compare them to the previous studies by Counsell *et al.*²¹ and Barkovich *et al.*⁴.

Estimation of gestational ages. The spatio-temporal models of MLS were utilised for estimation of gestational ages at scan (GA), to evaluate whether MLS were suitable markers of GA. The individual segmentations registered to the reference space were compared with models for each time-point using sum of squared errors and the GA with smallest error between the model and the segmentation was chosen as the final GA estimate for each subject.

The estimated GAs were plotted against the nominal values in Fig. 3. We obtained root mean squared errors (RMSEs) of 1.41 weeks and 2.56 weeks in the deep brain region and brainstem respectively. Therefore, each ROI has a different predictive power that best assesses a particular period of brain development. The the deep brain region with the more prominent MLS growth between 29 and 44 weeks GA produced the more accurate age estimates for preterm infants in this age range.

Discussion

In this work we proposed quantitative markers of the myelination patters in preterm neonates with GA between 29 and 44 weeks using T_2 -weighted MRI. We have proposed a novel dedicated method for segmentation of myelin-like signals with high reproducibility comparable to the intra-rater variability. The achieved Dice overlap with manual segmentations 0.83 can be considered particularly high given small size of myelinated structures compared to the voxels size. This was achieved by defining a clear segmentation protocol and designing a dedicated method to reproduce the protocol reliably. We have used Gaussian Mixture Model (GMM) to estimate the intensity profiles of myelinated and non-myelinated tissues to overcome the problem of qualitative nature T_2 -weighted MRI. Modelling of partial volume effect proved to be particularly important, as GMM alone did not produce reliable results. On the other hand, modeling the partial volume by an extra class and constraining its location using Markov Random Fields resulted in a stable segmentation method. We have demonstrated, that resulting segmentations can be used to produce quantitative markers of myelination by calculating volumes of myelinated tissues in absolute terms (e.g. in mL) or as a proportion of a pre-defined region of interest, to account for individual differences in brain sizes. Additionally, modelling of typical patterns of myelination during this age-range and accurate method for estimation of myelination status in form of predicted GA has been proposed. This method might have potential in uncovering developmental delays in this age group. Other studies have explored myelination patters in infancy and childhood^{52–56}, but this is the first such detailed study during pre-term neonatal period.

Inherent limitation of any study of myelination patters stems from lack of specificity of MRI sequences towards myelin. Histological studies consistently demonstrate that many MRI sequences (including quantitative

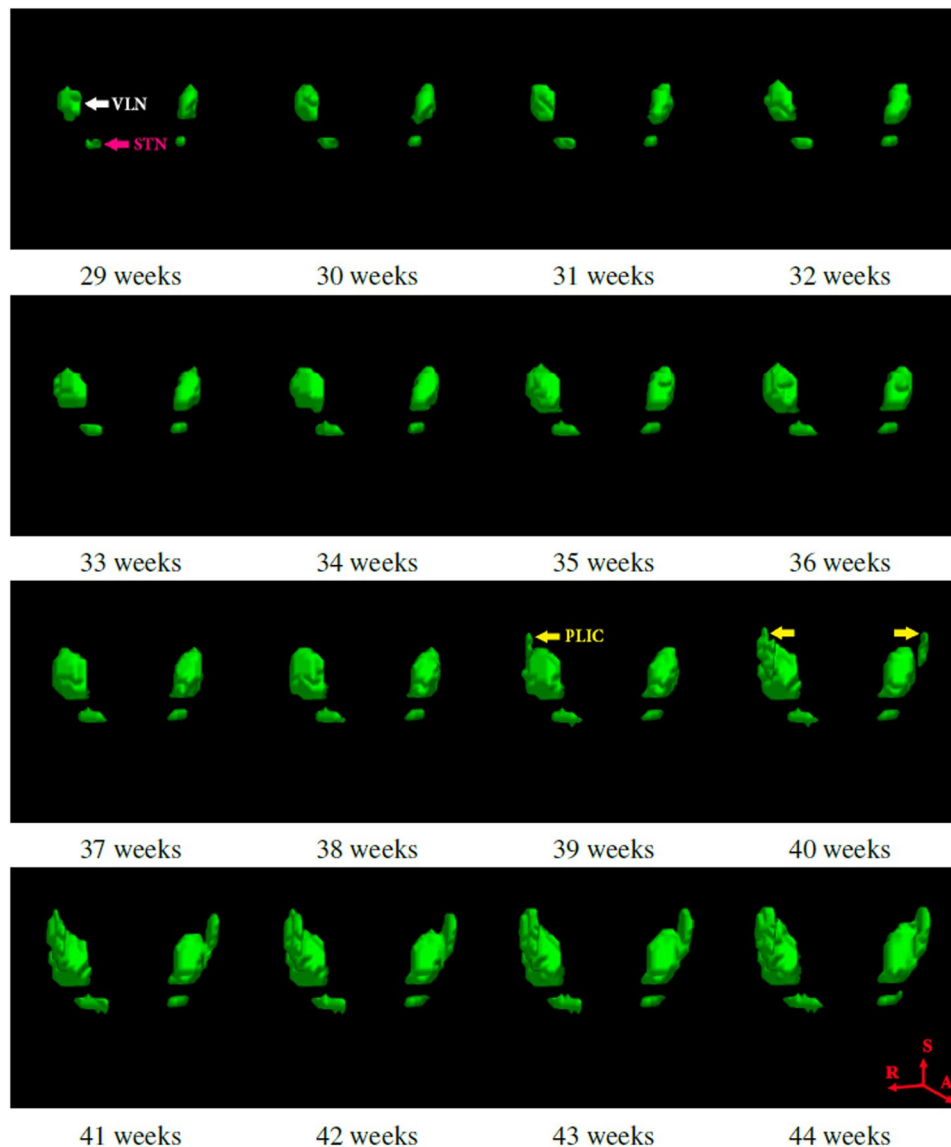


Figure 2. Spatio-temporal growth model for myelin-like signals (MLS) in the deep brain region between 29 and 44 weeks GA. The ventrolateral nuclei (VLN) and subthalamic nuclei (STN) appear to be myelinated before 29 weeks GA. MLS becomes evident in both tracts of the posterior limbs of the internal capsule (PLIC) at approximately 40 weeks GA. This is an important landmark for evaluating neonatal brain development. As the different deep brain region sizes in individual subjects have been corrected through non-rigid registration, the progression of MLS is completely due to the appearance of new myelination without contribution from the overall brain growth.

MRI such as relaxometry and MTI) are sensitive to myelin, but none is specific, which means that intensity patterns linked to degree of myelination could also be attributed to other underlying biological processes¹. We have overcome this issue by defining a segmentation protocol based on the available literature²¹ and creating a ROI “deep brain region” to separate myelinating structures from the structures that are more likely to be composed of densely packed grey matter. Inherently, uncertainties in the existing anatomical knowledge also exist in the proposed segmentation methodology. Nevertheless, gradual increase of the MLS in our spatio-temporal model supports the hypothesis that we are indeed detecting developing myelin.

Interesting direction for future work would be to quantify myelination progress using combination of T_1 and T_2 -weighted imaging. Given that shortening of T_1 and T_2 relaxation times determine different stages of myelination process¹⁷, their joint analysis could provide more detailed staging of myelination of different brain structures⁵⁶, resulting in further insights into brain maturation during perinatal period.

Though our segmentation method offers quantitative measure of volume of myelinated tissue, it cannot determine proportion of myelin within each voxel. Quantitative methods such as relaxometry and MTI could produce even more accurate quantification of myelin maturation once large databases in this age-range become available. Our results however show that the proposed methods can be nevertheless a useful marker of brain maturation

Region	structures	Counsell <i>et al.</i>	Barkovich <i>et al.</i>	Myelin atlas
Brainstem	SCP, DSCP, IC, LL, ML, MLE, VN	25–30 weeks GA	25–30 weeks GA	≤29 weeks GA
Deep brain	VLN	25 weeks GA	—	≤29 weeks GA
Deep brain	STN	28 weeks GA	—	≤29 weeks GA
Deep brain	PLIC	40 weeks GA	40 weeks GA	39–40 weeks GA

Table 1. Comparison of myelination milestones between the spatio-temporal atlas (Myelin atlas) and the qualitative description^{4,21}. The milestone refers to first time-point in gestation when the myelin is identified in a new site. The full names of the structures are provided in the section “Segmentation Protocol”.

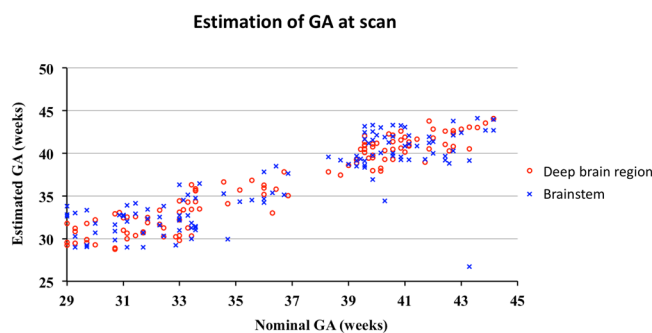


Figure 3. Gestational ages (GAs) predicted for 114 preterm infants and compared to the nominal GAs using the logistic growth models for myelin-like signals (MLS) in the deep brain region and brainstem. The predictions based on MLS in the brainstem display larger deviations from the nominal GAs than in the deep brain region. This is because most of the MLS spatio-temporal changes between 29 and 44 weeks GA occur in the deep brain region, whereas the brainstem provides much less information on progressing myelination during this particular age period.

during this period of brain development and has particular applications in large-scale studies to determine consequences of preterm birth on developing myelin.

Methods

Subjects and MR acquisition. 114 T₂-weighted scans of 95 preterm infants were acquired between 29 and 44 weeks GA at Hammersmith Hospital, London, UK. Ethical permission for this study was granted by the Hammersmith and Queen Charlotte’s and Chelsea Research Ethics Committee (07/H0707/101). Informed written parental consent was obtained prior to imaging. All methods and experiments were performed in accordance with relevant guidelines and regulations. T₂-weighted fast spin-echo brain images were acquired on a 3 T Philips Intera system with repetition time = 8700 ms, echo time = 160 ms and voxel sizes = 0.86 mm × 0.86 mm × 1 mm. Full information about the subjects is given in Table 2.

Segmentation protocol. Our segmentation protocol focuses on central brain structures (thalamus, basal ganglia and internal capsule) and brainstem for assessing MLS because in these regions the myelination is present between 29 and 44 weeks GA^{21,48}. Cerebellum was not included in this study. Since the myelinated nuclei can have similar intensities to the densely organized GM in the basal ganglia which are thought not to contain myelin between 29 and 44 weeks GA²¹, we create a custom ROI containing the posterior limbs of the internal capsule (PLIC), ventrolateral nuclei (VLN) and subthalamic nuclei (STN), referred to as deep brain region in this paper. The MLS segmentation protocol is defined based on both image intensities and anatomical knowledge²¹. We delineate three structures, PLIC, VLN and STN in the deep brain region. In the brainstem we define the segmentation protocol mainly based on intensities because the image resolution is insufficient to identify the detailed anatomical structures. An example of the segmentation protocol is shown in Fig. 4.

Segmentation pipeline. The overall pipeline proposed for MLS segmentation on T₂-weighted neonatal brain MR images is shown in Fig. 5. The images are first pre-processed to achieve brain extraction and bias field correction, and to obtain the binary ROI masks of the deep brain region and brainstem for individual subjects. We remove the non-brain tissues using label propagation³¹ of manually annotated brain masks, and segment the skull-stripped images using the Statistical Parametric Mapping (SPM) software (www.fil.ion.ucl.ac.uk/spm, version SPM8)³⁰. The spatial priors are provided by a publicly available (brain-development.org/brain-atlases) 4D probabilistic neonatal brain atlas³⁵. Subsequently, we obtain the bias-corrected images as well as the segmentations of the cortical GM, WM, deep gray matter (DGM), brainstem, cerebellum and CSF.

We use the automatic SPM segmentations of the brainstem directly as the ROI mask. The DGM segmentations include the basal ganglia, thalami, hippocampi and amygdalae. Since the myelinated nuclei can have similar intensities to the densely organized GM in the basal ganglia which is considered not to contain myelin between

	Subjects	All scans	Preterm scans	Term scans
Number	95	114	58	56
GA scan (weeks)		36.6 ± 4.6	32.4 ± 2.2	40.9 ± 1.5
GA birth (weeks)	29.6 ± 2.6	29.5 ± 2.6	29.4 ± 2.3	29.7 ± 2.9
Birth weight (kg)	1.27 ± 0.52	1.26 ± 0.48	1.20 ± 0.33	1.33 ± 0.61
Female	48	58	32	26
Male	47	56	26	30
Singletons	37	50	27	23
Twins	58	64	31	33
WM lesions	1	2	1	1
WM cysts	2	2	1	1
Cerebellar haem orrage	1	1	1	0

Table 2. Subjects and scans in the study. Some subjects had two or three scans. The first column refers to numbers and averages over all subjects. The other columns refer to numbers or averages over all scans, scans at preterm period (GA at scan between 27 and 37 weeks) and term equivalent age (GA at scan between 37 and 44 weeks).

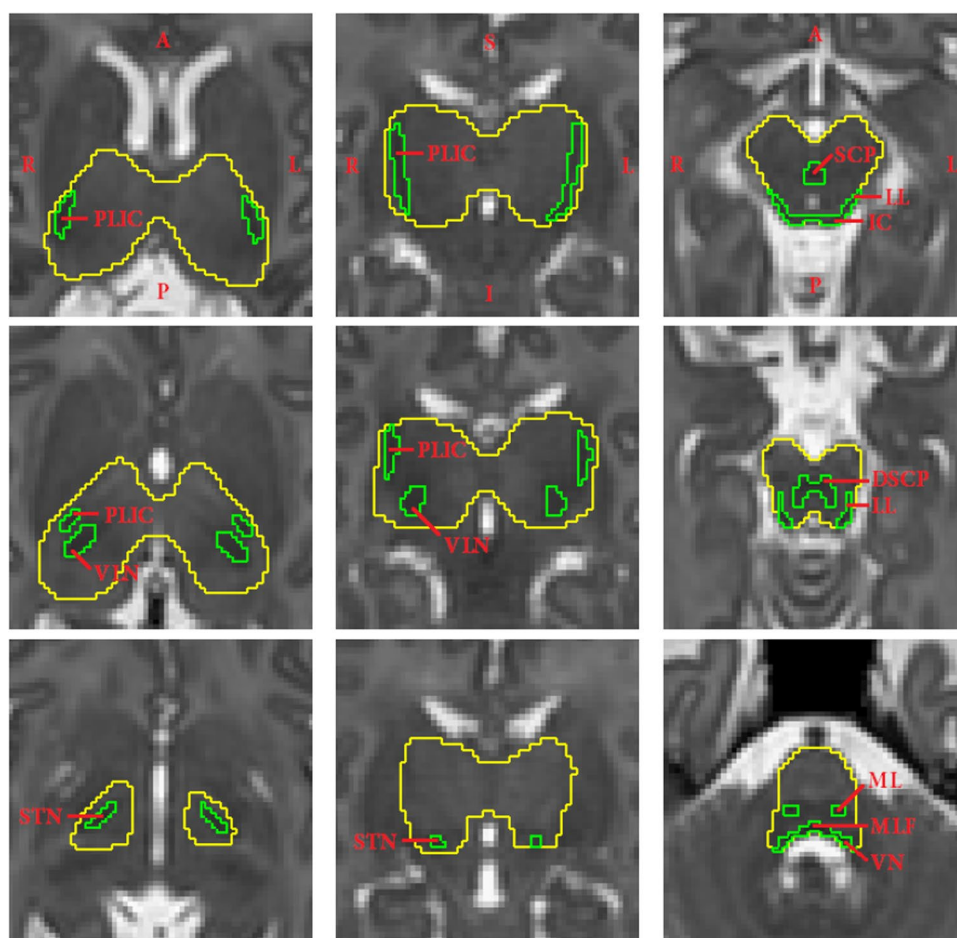


Figure 4. Manual annotations of myelin-like signals (green label) in the deep brain region and brainstem of a subject at 42 gestational weeks, delineated according to the defined segmentation protocol. The regions of interest are labeled in yellow. The columns from left to right show the deep brain region in the axial and coronal views, and the brainstem in the axial view. Abbreviations: PLIC–posterior limb of the internal capsule, VLN–ventrolateral nucleus, STN–subthalamic nucleus, SCP–superior cerebellar peduncle, DSCP–decussation of the superior cerebellar peduncle, IC–inferior colliculus, LL–lateral lemniscus, ML–medial lemniscus, MLF–medial longitudinal fasciculus, VN–vestibular nucleus.

29 and 44 weeks GA²¹, we extract the ROI masks for the deep brain region from the DGM segmentations in order to prevent misclassifications by the intensity-based segmentation model. We first manually delineate the deep

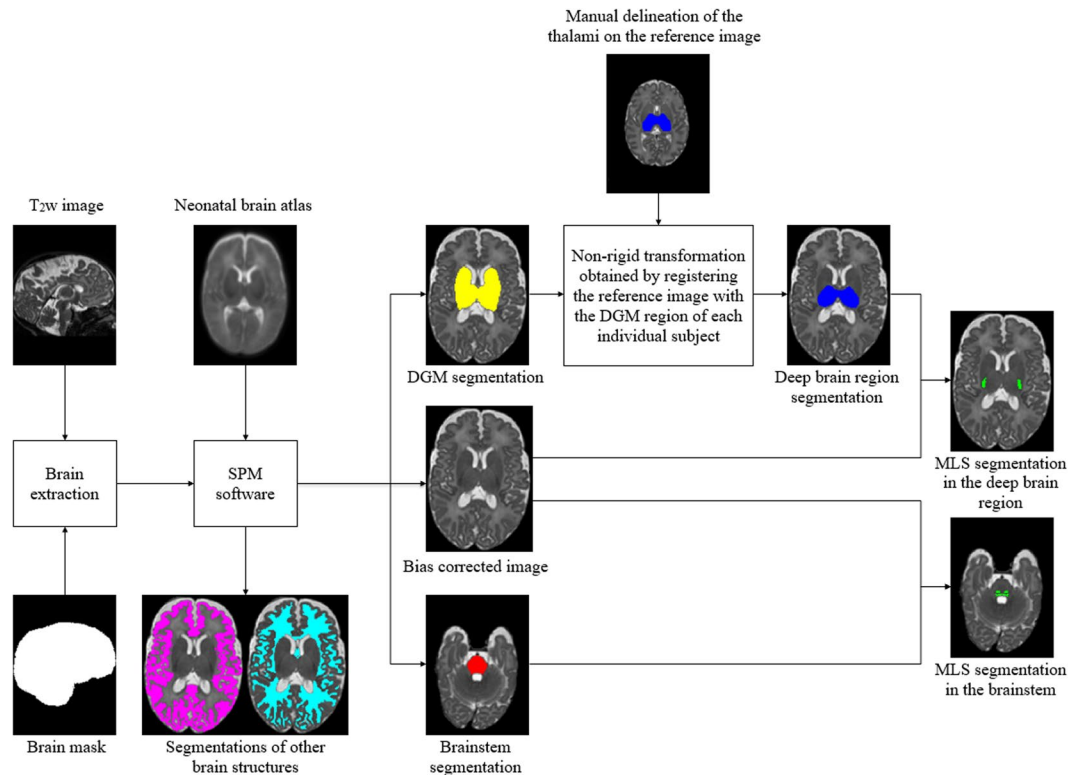


Figure 5. Schematic flow of the proposed segmentation approach for myelin-like signals (MLS) on T_2 -weighted neonatal brain images. Through image preprocessing, we achieve brain extraction and bias field correction, and obtain the binary masks of the deep brain region and brainstem for individual subjects. The automatic segmentations of the deep gray matter (DGM), obtained using the Statistical Parametric Mapping (SPM) software, contain the basal ganglia, thalami, hippocampi and amygdalae. We extract the masks for the deep brain region from the DGM segmentations via label propagation.

brain region on the T_2 -weighted image of a single reference subject at 36 weeks GA. We register the reference image with the dilated DGM region of each subject using free-form deformation (FFD) non-rigid registration⁵⁰, and then transform the manual delineation of the deep brain region from the reference space to each individual subject's space.

Segmentation of myelin like signals. The segmentation of myelin like signals is performed in Expectation Maximisation framework. We model three classes, MLS, BKG and PV, with their intensities defined by a Gaussian distribution $G(y, \mu_k, \sigma_k)$ with means μ_k and variances σ_k^2 , where y is the voxel intensity and k is the class index. We approximate the PV class mean as the arithmetic mean of the composing pure tissue means μ_{MLS} and μ_{BKG} , and assume that the variances of all three classes are identical:

$$\mu_{PV} = \frac{1}{2}(\mu_{MLS} + \mu_{BKG}) \tag{1}$$

$$\sigma_{MLS}^2 = \sigma_{PV}^2 = \sigma_{BKG}^2 \tag{2}$$

By substituting Eqs. 1 and 2 into the objective function of log likelihood⁵⁷, the ML estimations for the means of the MLS and BKG classes result in the following system of linear equations (see more details in the appendix):

$$\begin{bmatrix} \mu_{MLS}^{(m+1)} \\ \mu_{BKG}^{(m+1)} \end{bmatrix} = \begin{bmatrix} \frac{1}{4}B_{PV} & B_{BKG} + \frac{1}{4}B_{PV} \\ B_{MLS} + \frac{1}{4}B_{PV} & \frac{1}{4}B_{PV} \end{bmatrix}^{-1} \begin{bmatrix} A_{BKG} + \frac{1}{2}A_{PV} \\ A_{MLS} + \frac{1}{2}A_{PV} \end{bmatrix} \tag{3}$$

where

$$A_k = \sum_{i=1}^N y_i p_{ik}^{(m+1)} \quad \text{and} \quad B_k = \sum_{i=1}^N p_{ik}^{(m+1)}$$

The variances σ_k^2 of class k are updated as

$$(\sigma_k^{(m+1)})^2 = (\sigma^{(m+1)})^2 = \frac{1}{N} \sum_{i=1}^K \sum_{l=1}^N p_{il}^{(m+1)} (y_i - \mu_l^{(m+1)})^2 \tag{4}$$

where p_{ik} is the posterior probability of voxel i belonging to class k , y_i is the observed intensity of voxel i , K the number of classes, and m the iteration number. The above equations comprise the M-step of Expectation Maximisation algorithm.

The E-step of the expectation maximisation algorithm calculates the posterior probabilities, which represent probabilistic segmentation. The probabilistic segmentation is regularised with the spatial prior $p(\mathbf{z}_i = \mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z)$ determined from the posteriors of the neighboring voxels as described in⁵⁷:

$$p_{ik}^{(m+1)} = \frac{G(y_i, \mu_k^{(m)}, \sigma_k^{(m)}) p(\mathbf{z}_i = \mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z)}{\sum_{k'=1}^K G(y_i, \mu_{k'}^{(m)}, \sigma_{k'}^{(m)}) p(\mathbf{z}_i = \mathbf{e}_{k'} | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z)} \tag{5}$$

where

$$p(\mathbf{z}_i = \mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z) = \frac{e^{-U_{\text{MRF}}(\mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z)}}{\sum_{k''=1}^K e^{-U_{\text{MRF}}(\mathbf{e}_{k''} | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z)}} \tag{6}$$

The symbol \mathbf{z}_i indicates to which class voxel i belongs, and \mathbf{e}_k is a unit vector of length K whose components are equal to zero except the k th component. The probabilities in the 26-neighborhood \mathcal{N}_i of voxel i are denoted as $\mathbf{p}_{\mathcal{N}_i}$, and the MRF parameters as Φ_z . We compute the energy function $U_{\text{MRF}}(\mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}, \Phi_z)$ for second-order MRFs as follows:

$$U_{\text{MRF}}(\mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z) = \sum_{k_1=1}^K v_{k_1}^{(m)} \sum_{k_2=1}^K \mathbf{T}_k(k_1, k_2) v_{k_2}^{(m)} \\ = [v_1 \ v_2 \ \dots \ v_K] \begin{bmatrix} T_k(1, 1) & T_k(1, 2) & \dots & T_k(1, K) \\ T_k(2, 1) & T_k(2, 2) & \dots & T_k(2, K) \\ \vdots & \vdots & \ddots & \vdots \\ T_k(K, 1) & T_k(K, 2) & \dots & T_k(K, K) \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_K \end{bmatrix} \tag{7}$$

where

$$v_k^{(m)} = \sum_{j \in \mathcal{N}_i} \frac{p_{jk}^{(m)}}{d_{ij}} \tag{8}$$

Here p_{jk} is the probability that neighbor j belongs to class k , and d_{ij} the Euclidean distance between voxel i and neighbor j . The element of the connectivity tensor, denoted as $\mathbf{T}_k(k_1, k_2)$, indicates the penalty when classes k_1 and k_2 are both present in the neighborhood of class k . By assigning a large value to $\mathbf{T}_k(k_1, k_2)$ in Eq. 7, we are able to penalize a particular configuration when the probabilities of classes k_1 and k_2 are both high in the neighborhood. This would result in a large value of the energy function $U_{\text{MRF}}(\mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}, \Phi_z)$, making voxel i less probable to belong to class k .

We assign the penalties $\mathbf{T}_k(k_1, k_2)$ according to the following rules based on the triplet that contains class k of the current voxel and classes k_1 and k_2 of a pair of neighboring voxels:

- $\mathbf{T}_k(\text{MLS}, \text{MLS})$ and $\mathbf{T}_k(\text{BKG}, \text{BKG})$: We set the penalty to 0 if class k is the same as the pair. We assign a penalty t_1 if k is a pure tissue class different from the pair to encourage the PV class as MLS and BKG are both in the triplet. We assign a penalty t_3 if k is PV as there is no evidence for PV in the triplet. This effectively prevents over-estimation of the PV class.
- $\mathbf{T}_k(\text{MLS}, \text{BKG})$: We assign a penalty t_2 if k is MLS or BKG, and set the penalty to 0 if k is PV. We aim to encourage the PV class by deliberately penalizing the MLS and BKG classes if they are both present in the neighborhood.
- $\mathbf{T}_k(\text{MLS}, \text{PV})$, $\mathbf{T}_k(\text{PV}, \text{PV})$ and $\mathbf{T}_k(\text{BKG}, \text{PV})$: We set the penalty to 0 regardless of the class of k .

The values of penalties t_1 , t_2 and t_3 needed to be determined empirically and were set to 0.05, 0.03 and 0.01 by optimising the agreement with manual segmentation using Dice score⁴⁶.

The EM algorithm is initialized by assigning 6 percent voxels with lowest intensities to MLS and rest to BKG class in deep brain region, while 25th percentile is used for brainstem. These values were estimated from manual segmentations. Tissue classification is achieved by interleaving the E-step and M-step until the relative change in the objective function of log likelihood⁵⁷ is less than 0.01%. The resultant PPMs are converted into hard segmentations using the maximum-vote rule. In addition, we reclassify the PV voxels as one of the composing pure tissues by calculating the fraction of the MLS class, denoted as f_i , at PV voxel i :

Method	Average DC	<i>p</i> -value
Thresholding at the 6th percentile	0.789 ± 0.091	0.043
GMM	0.762 ± 0.092	0.001
GMM-PV	0.662 ± 0.192	0.002
GMM-PV-MRF	0.837 ± 0.057	—

Table 3. Average Dice coefficients (DCs) (±standard deviations) for automatic segmentations of myelin-like signals in deep brain region compared with manual segmentations for 16 test subjects aged between 29 and 44 gestational weeks. GMM-PV-MRF is compared with all the other methods using two-tailed Student's *t*-tests at the 5% significance level.

Method	Average DC	<i>p</i> -value
Thresholding at the 25th percentile	0.827 ± 0.035	0.020
GMM	0.675 ± 0.206	0.010
GMM-PV	0.762 ± 0.057	5.449 × 10 ⁻⁴
GMM-PV-MRF	0.831 ± 0.038	—

Table 4. Average Dice coefficients (DCs) (±standard deviations) for automatic segmentations of myelin-like signals in brainstem compared with manual segmentations for 16 test subjects aged between 29 and 44 gestational weeks. GMM-PV-MRF is compared with all the other methods using two-tailed Student's *t*-tests at the 5% significance level.

Method	Manual set 1		Manual set 2		Intra-rater reliability
	Average DC	<i>p</i> -value	Average DC	<i>p</i> -value	
GMM	0.785 ± 0.085	0.257	0.754 ± 0.054	0.023	0.834 ± 0.043
GMM-PV	0.662 ± 0.210	0.060	0.626 ± 0.217	0.028	
GMM-PV-MRF	0.843 ± 0.051	0.624	0.816 ± 0.046	0.413	

Table 5. Average Dice coefficients (DCs) (±standard deviations) for automatic segmentation of myelin-like signals in deep brain region of eight test subjects with respect to two sets of manual annotations compared to the intra-rater reliability. The *p*-values indicate the outcomes of two-tailed Student's *t*-tests with the intra-rater DCs at the 5% significance level. The performance of GMM-PV-MRF is close to the intra-rater reliability.

$$f_i = \frac{\mu_{\text{BKG}} - y_i}{\mu_{\text{BKG}} - \mu_{\text{MLS}}} \quad (9)$$

where μ_{MLS} and μ_{BKG} are the means of the MLS and BKG classes, and y_i the observed intensity. PV voxels with fractions above 0.5 are reclassified as MLS and combined with the voted hard segmentation of the MLS class to form the final segmentation.

Evaluation of automatic segmentation. The automatic segmentation method of MLS was evaluated against manual annotations of MLS for 16 preterm infants, each at a different time point, between 29 and 44 weeks GA at one-week intervals. Eight of the 16 subjects, from 30 weeks GA onward at two-week intervals, have a repeated manual segmentation delineated by the same rater in order to assess the intra-rater reliability. The quality of these manual annotations was confirmed by a clinical expert.

We compared agreement with manual segmentations for three different methods: Gaussian Mixture Model with 2 classes MLS and BKG (GMM), Gaussian Mixture Model with explicit PV modelling but no spatial regularisation (GMM-PV) and the proposed three class model with spatial regularisation using MRF (GMM-PV-MRF). The proposed method clearly outperformed the other methods as well as the initial thresholding as shown in Tables 3 and 4. Additionally, the proposed method GMM-PV-MRF was found to be insensitive to the initialisation, while other tested method depended on correct initial threshold.

We also assessed the intra-rater reliability as DCs between two sets of manual annotations that were available for eight of the 16 test subjects at two-week intervals between 30 and 44 weeks GA. We evaluated the automated model-based methods with respect to each manual set. The average DCs over the eight subjects are shown in Tables 5 and 6 along with the *p*-values of two-tailed Student's *t*-tests with the intra-rater DCs at the 5% significance level. It can be seen that only GMM-PV-MRF performed equally well as a human rater in both ROIs, and demonstrated no significant differences with respect to the repeated manual annotations.

We achieved high Dice overlaps of 0.83 comparable to the intra-rater variability. This proves that our method is reliable and reproducible across the target age-range. In contrast, the Dice overlap achieved by the methods that

Method	Manual set 1		Manual set 2		Intra-rater reliability
	Average DC	p-value	Average DC	p-value	
GMM	0.702 ± 0.171	0.046	0.695 ± 0.156	0.026	0.842 ± 0.034
GMM-PV	0.763 ± 0.047	0.002	0.766 ± 0.077	0.031	
GMM-PV-MRF	0.832 ± 0.034	0.608	0.815 ± 0.024	0.087	

Table 6. Average Dice coefficients (DCs) (\pm standard deviations) for automatic segmentation of myelin-like signals in brainstem of eight test subjects with respect to two sets of manual annotations compared to the intra-rater reliability. The p -values indicate the outcomes of two-tailed Student's t -tests with the intra-rater DCs at the 5% significance level. The performance of GMM-PV-MRF is close to the intra-rater reliability.

entered Neonatal Segmentation Challenge²⁹, achieved Dice overlap 0.16–0.69 for 30 weeks GA and 0.23–0.68 for 40 weeks GA. None of the method achieved Dice overlap over 0.6 for both age groups.

Spatio-temporal growth models for MLS. We constructed MLS growth models in the deep brain region and brainstem using voxelwise logistic regression based on the automatic segmentations computed using GMM-PV-MRF for the 114 subjects⁴⁹. First we registered the T_2 -weighted image of each individual subject with the dilated ROI of the single reference subject at 36 weeks GA using FFD non-rigid registration⁵⁰. This was the same reference subject used to create the ROI masks of the deep brain region. We used normalized mutual information (NMI) as the similarity measure and 10 mm B-spline control point spacing. The automatic MLS segmentations were then transformed accordingly from each subject's space to the common reference space. Lastly, we constructed the growth model in each ROI by fitting a voxelwise logistic function to the transformed segmentations. Details of constructing of the spatio-temporal models can be found in our previous work⁴⁹.

We further used the MLS growth models in the deep brain region and brainstem to predict GAs of the 114 preterm infants. The age estimates were determined by minimizing the sum of squared differences (SSD) between each individual transformed segmentation and the average growth model constructed in a leave-one-out procedure.

References

1. Laule, C. *et al.* Magnetic resonance imaging of myelin. *Neurotherapeutics* **4**, 460–484 (2007).
2. Yakolev, P. & Lecours, A. The myelogenetic cycles of regional maturation of the brain. In Mankowski, A. (ed.) *Regional Development of the Brain in Early Life*, 3–70 (Blackwell, Oxford, 1967).
3. Barkovich, A. J., Lyon, G. & Evrard, P. Formation, maturation, and disorders of white matter. *American Journal of Neuroradiology* **13**, 447–461, <http://www.ajnr.org/content/13/2/447.full.pdf> (1992).
4. Barkovich, A. J. Magnetic resonance techniques in the assessment of myelin and myelination. *Journal of Inherited Metabolic Disease* **28**, 311–343, <https://doi.org/10.1007/s10545-005-5952-z> (2005).
5. Rorke, L. & Riggs, H. *Myelination of the brain in the newborn*. (Lippincott, Philadelphia, 1969).
6. Volpe, J. J., Kinney, H. C., Jensen, F. E. & Rosenberg, P. A. The developing oligodendrocyte: Key cellular target in brain injury in the premature infant. *International Journal of Developmental Neuroscience*, <https://doi.org/10.1016/j.ijdevneu.2011.07.008> (2011).
7. Ortinau, C. & Neil, J. The neuroanatomy of prematurity: Normal brain development and the impact of preterm birth. *Clinical Anatomy*, <https://doi.org/10.1002/ca.22430> (2015).
8. McArdle, C. B. *et al.* Developmental features of the neonatal brain: MR imaging. Part I. Gray-white matter differentiation and myelination. *Radiology* **162**, 223–229 (1987).
9. Inder, T. E., Wells, S. J., Mogridge, N. B., Spencer, C. & Volpe, J. J. Defining the nature of the cerebral abnormalities in the premature infant: A qualitative magnetic resonance imaging study. *Journal of Pediatrics*, [https://doi.org/10.1067/S0022-3476\(03\)00357-3](https://doi.org/10.1067/S0022-3476(03)00357-3) (2003).
10. Inder, T. E. Abnormal Cerebral Structure Is Present at Term in Premature Infants. *Pediatrics*, <https://doi.org/10.1542/peds.2004-0326> (2005).
11. Hüppi, P. S. *et al.* Microstructural development of human newborn cerebral white matter assessed *in vivo* by diffusion tensor magnetic resonance imaging. *Pediatric Research*, <https://doi.org/10.1203/00006450-199810000-00019> (1998).
12. Tsur, N., Arichi, T., Counsell, S. J. & Edwards, A. D. Brain development in preterm infants assessed using advanced MRI techniques. *Clinics in Perinatology* **41**, 25–45, <https://doi.org/10.1016/j.clp.2013.10.001>, *Advances in Neonatal Neurology* (2014).
13. Knight, M. J., Smith-Collins, A., Newell, S., Denbow, M. & Kauppinen, R. A. Cerebral White Matter Maturation Patterns in Preterm Infants: An MRI T2 Relaxation Anisotropy and Diffusion Tensor Imaging Study. *Journal of Neuroimaging*, <https://doi.org/10.1111/jon.12486> (2018).
14. Woodward, L. J., Anderson, P. J., Austin, N. C., Howard, K. & Inder, T. E. Neonatal MRI to Predict Neurodevelopmental Outcomes in Preterm Infants. *New England Journal of Medicine*, <https://doi.org/10.1056/nejmoa053792> (2006).
15. Ment, L. R., Hirtz, D. & Hüppi, P. S. Imaging biomarkers of outcome in the developing preterm brain. *The Lancet Neurology*, [https://doi.org/10.1016/S1474-4422\(09\)70257-1](https://doi.org/10.1016/S1474-4422(09)70257-1) (2009).
16. Woodward, L. J., Clark, C. A. C., Bora, S. & Inder, T. E. Neonatal White Matter Abnormalities an Important Predictor of Neurocognitive Outcome for Very Preterm Children. *PLoS One*, <https://doi.org/10.1371/journal.pone.0051879> (2012).
17. Dubois, J. *et al.* The early development of brain white matter: A review of imaging studies in fetuses, newborns and infants. *Neuroscience* **276**, 48–71 (2014).
18. Rutherford, M. *et al.* Magnetic resonance imaging in perinatal brain injury: Clinical presentation, lesions and outcome. *Pediatr. Radiol.* **36**, 582–592 (2006).
19. Srinivasan, L. & Rutherford, M. MRI of the newborn brain. *Paediatr. Child Health* **18**, 183–195 (2008).
20. Barkovich, A. MR of the normal neonatal brain: Assessment of deep structures. *Am. J. Neuroradiol.* **19**, 1397–1403 (1998).
21. Counsell, S. *et al.* MR imaging assessment of myelination in the very preterm brain. *Am. J. Neuroradiol.* **23**, 872–881 (2002).
22. Alonso-Ortiz, E., Levesque, I. & Pike, G. MRI-based myelin water imaging: A technical review. *Magn. Reson. Med.* **73**, 70–81 (2015).
23. Henkelman, R., Stanisz, G. & Graham, S. Magnetization transfer in MRI: A review. *NMR Biomed.* **14**, 57–64 (2001).
24. Gui, L. *et al.* Morphology-driven automatic segmentation of MR images of the neonatal brain. *Med. Image Anal.* **16**, 1565–1579 (2012).
25. Makropoulos, A. *et al.* Automatic whole brain MRI segmentation of the developing neonatal brain. *IEEE Trans. Med. Imag.* **33**, 1818–1831 (2014).

26. Melbourne, A. *et al.* NeobrainS12 challenge: Adaptive neonatal MRI brain segmentation with myelinated white matter class and automated extraction of ventricles I–IV. In *Proceedings of the MICCAI Grand Challenge: Neonatal Brain Segmentation*, 16–21 (2012).
27. Wang, S., Kuklisova-Murgasova, M. & Schnabel, J. An atlas-based method for neonatal MR brain tissue segmentation. In *Proceedings of the MICCAI Grand Challenge: Neonatal Brain Segmentation*, 28–35 (2012).
28. Wu, J. & Avants, B. Automatic registration-based segmentation for neonatal brains using ANTs and Atropos. In *Proceedings of the MICCAI Grand Challenge: Neonatal Brain Segmentation*, 36–43 (2012).
29. Isgum, I. *et al.* Evaluation of automatic neonatal brain segmentation algorithms: The NeoBrainS12 challenge. *Med. Image Anal.* **20**, 135–151 (2015).
30. Ashburner, J. & Friston, K. Unified segmentation. *NeuroImage* **26**, 839–851 (2005).
31. Heckemann, R., Hajnal, J., Aljabar, P., Rueckert, D. & Hammers, A. Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. *NeuroImage* **33**, 115–126 (2006).
32. Van Leemput, K., Maes, F., Vandermeulen, D. & Suetens, P. Automated model-based bias field correction of MR images of the brain. *IEEE Trans. Med. Imag.* **18**, 885–896 (1999).
33. Wells, W., Grimson, W., Kikinis, R. & Jolesz, F. Adaptive segmentation of MRI data. *IEEE Trans. Med. Imag.* **15**, 429–442 (1996).
34. Zhang, Y., Brady, M. & Smith, S. Segmentation of brain MR images through a hidden Markov random field and the expectation-maximization algorithm. *IEEE Trans. Med. Imag.* **20**, 45–57 (2001).
35. Kuklisova-Murgasova, M. *et al.* A dynamic 4D probabilistic atlas of the developing brain. *NeuroImage* **54**, 2750–2763 (2011).
36. Serag, A. *et al.* Construction of a consistent high-definition spatio-temporal atlas of the developing brain using adaptive kernel regression. *NeuroImage* **59**, 2255–2265 (2012).
37. Makropoulos, A. *et al.* Regional growth and atlas of the developing human brain. *NeuroImage* **125**, 456–478 (2016).
38. Gousias, I. *et al.* Magnetic resonance imaging of the newborn brain: Manual segmentation of labelled atlases in term-born and preterm infants. *NeuroImage* **62**, 1499–1509 (2012).
39. Makropoulos, A. *et al.* The developing human connectome project: A minimal processing pipeline for neonatal cortical surface reconstruction. *NeuroImage* **173**, 88–112 (2018).
40. Cardoso, M. *et al.* LoAd: A locally adaptive cortical segmentation algorithm. *NeuroImage* **56**, 1386–1397 (2011).
41. Ledig, C., Wright, R., Serag, A., Aljabar, P. & Rueckert, D. Neonatal brain segmentation using second order neighborhood information. In *Proceedings of the MICCAI Workshop: Perinatal and Paediatric Imaging*, 33–40 (2012).
42. Manjon, J., Tohka, J. & Robles, M. Improved estimates of partial volume coefficients from noisy brain MRI using spatial context. *NeuroImage* **53**, 480–490 (2010).
43. Shattuck, D., Sandor-Leahy, S., Schaper, K. A., Rottenberg, D. & Leahy, R. Magnetic resonance image tissue classification using a partial volume model. *NeuroImage* **13**, 856–876 (2001).
44. Tohka, J., Zijdenbos, A. & Evans, A. Fast and robust parameter estimation for statistical partial volume models in brain MRI. *NeuroImage* **23**, 84–97 (2004).
45. Van Leemput, K., Maes, F., Vandermeulen, D. & Suetens, P. A unifying framework for partial volume segmentation of brain MR images. *IEEE Trans. Med. Imag.* **22**, 105–119 (2003).
46. Dice, L. Measures of the amount of ecologic association between species. *Ecology* **23**, 297–302 (1945).
47. Chi, J., Dooling, E. & Gilles, F. Gyral development of the human brain. *Ann. Neurol.* **1**, 86–93 (1977).
48. Rutherford, M. *MRI of the Neonatal Brain*. (Saunders, Edinburgh, 2002).
49. Wang, S., Kuklisova-Murgasova, M., Hajnal, J., Ledig, C. & Schnabel, J. Regression analysis for assessment of myelination status in preterm brains with magnetic resonance imaging. In *Proceedings of the 13th International Symposium on Biomedical Imaging (ISBI)*, 278–281 (2016).
50. Rueckert, D. *et al.* Nonrigid registration using free-form deformations: Application to breast MR images. *IEEE Trans. Med. Imag.* **18**, 712–721 (1999).
51. Bird, C. *et al.* MR assessment of myelination in infants and children: Usefulness of marker sites. *Neuroradiology* **10**, 731–740 (1989).
52. Devi, C. N., Chandrasekharan, A., Sundararaman, V. K. & Alex, Z. C. Automatic segmentation of infant brain MR images: With special reference to myelinated white matter. *Biocybernetics and Biomedical Engineering* **37**, 143–158 (2017).
53. Deoni, S. C. L. *et al.* Mapping infant brain myelination with magnetic resonance imaging. *Journal of Neuroscience* **31**, 784–791 (2011).
54. Dean, D. C. *et al.* Modeling healthy male white matter and myelin development: 3 through 60 months of age. *NeuroImage* **84**, 742–752 (2014).
55. Dean, D. C. III. *et al.* Estimating the age of healthy infants from quantitative myelin water fraction maps. *Human Brain Mapping* **36**, 1233–1244 (2015).
56. Kulikova, S. *et al.* Multi-parametric evaluation of the white matter maturation. *Brain Structure and Function* **220**, 3657–3672 (2015).
57. Van Leemput, K., Maes, F., Vandermeulen, D. & Suetens, P. Automated model-based tissue classification of MR images of the brain. *IEEE Trans. Med. Imag.* **18**, 897–908 (1999).

Acknowledgements

S. Wang was a recipient of the Oxford Clarendon Fund Scholarship which is provided by the Oxford University Press through the Life Sciences Interface Doctoral Training Center. The work was supported by MRC Strategic Funds (MR/K006355/1), ERC developing Human Connectome Project (Grant Agreement no. 319456), Academy of Medical Sciences Springboard Award (SBF004040), Wellcome EPSRC CME (WT 203148/Z/16/Z) and by the NIHR Biomedical Research Centre. The views expressed are those of the authors and not necessarily of the NHS, NIHR or the Department of Health.

Author Contributions

S.W. conducted the study and experiments, M.D. and J.A.S. conceived the study, S.W., M.D., J.A.S. and C.L. conceived the techniques, S.W., M.D. and J.A.S. conceived the experiments, S.J.C. advised on anatomy of the neonatal brain, J.V.H. advised on acquisition aspects. All authors reviewed the manuscript.

Additional Information

Competing Interests: The authors declare no competing interests.

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019