Unbiased construction of a temporally consistent morphological atlas of neonatal brain development

Andreas Schuh[a,1], Antonios Makropoulos[a], Emma C. Robinson[c], Lucilio Cordero-Grande[b], Emer Hughes[b], Jana Hutter[b], Anthony Price[b], Maria Murgasova[b], Rui Pedro A. G. Teixeira[b], Nora Tusor[b], Johannes K. Steinweg[b], Suresh Victor[b], Mary A. Rutherford[b], Joseph V. Hajnal[b], A. David Edwards[b], Daniel Rueckert[a]

[a] *Biomedical Image Analysis Group, Department of Computing, Imperial College London, London, United Kingdom*
[b] *Centre for the Developing Brain, King's College London, London, United Kingdom*
[c] *Department of Biomedical Engineering, School of Biomedical Engineering and Imaging Sciences, King's College London, London, United Kingdom*

Abstract

Premature birth increases the risk of developing neurocognitive and neurobehavioural disorders. The mechanisms of altered brain development causing these disorders are yet unknown. Studying the morphology and function of the brain during maturation provides us not only with a better understanding of normal development, but may help us to identify causes of abnormal development and their consequences. A particular difficulty is to distinguish abnormal patterns of neurodevelopment from normal variation. The Developing Human Connectome Project (dHCP) seeks to create a detailed four-dimensional (4D) connectome of early life. This connectome may provide insights into normal as well as abnormal patterns of brain development. As part of this project, more than a thousand healthy fetal and neonatal

*Corresponding author
Email address: andreas.schuh@imperial.ac.uk (Andreas Schuh)

brains will be scanned *in vivo*. This requires computational methods which scale well to larger data sets. We propose a novel groupwise method for the construction of a spatio-temporal model of mean morphology from cross-sectional brain scans at different gestational ages. This model scales linearly with the number of images and thus improves upon methods used to build existing public neonatal atlases, which derive correspondence between all pairs of images. By jointly estimating mean shape and longitudinal change, the atlas created with our method overcomes temporal inconsistencies, which are encountered when mean shape and intensity images are constructed separately for each time point. Using this approach, we have constructed a spatio-temporal atlas from 275 healthy neonates between 35 and 44 weeks post-menstrual age (PMA). The resulting atlas qualitatively preserves cortical details significantly better than publicly available atlases. This is moreover confirmed by a number of quantitative measures of the quality of the spatial normalisation and sharpness of the resulting template brain images.
Keywords: dHCP, spatio-temporal, neonatal MRI, brain atlas

---

1. Introduction

Medical image analysis has made significant progress over the past decades. With advances in acquiring high quality images of the developing human brain using magnetic resonance imaging (MRI), analysing this data to understand brain development is rapidly becoming feasible (Kwon et al., 2014).

Many studies have shown that premature birth increases the risk of developing neurocognitive and neurobehavioural disorders (Rutherford, 2002; Blencowe et al., 2012; Sled and Nossin-Manor, 2013). A known difference to term-born infants is the reduced cortical folding observed in premature-born infants at term-equivalent age (Hüppi et al., 1996; Inder et al., 2005; Kapellou et al., 2006). Other reported abnormalities include: (1) white matter (WM) signal abnormalities such as lesions and diffuse excessive high signal intensities, (2) cortical grey matter (cGM) signal abnormalities, and (3) cerebellar hemorrhages (cf. Kwon et al., 2014, and references therein). However, the mechanisms of these alterations are unknown (Molnár and Rutherford, 2013). Studying the morphology and function of the brain during maturation, provides us not only with a better understanding of normal development, but may help identify causes for these factors.

The Developing Human Connectome Project (dHCP) seeks to create a 4D connectome of early life. Understanding this connectome may provide insights into abnormal patterns of brain development. In this context, an atlas of normal brain development is important, as it facilitates spatial normalisation, voxel-based morphometry (Ashburner and Friston, 2000), deformation-based morphometry (Ashburner et al., 1998), and computational anatomy studies (Grenander and Miller, 1998; Thompson and Toga, 2002).

In previous work (Schuh et al., 2014), we constructed a spatio-temporal neonatal atlas from registrations between all pairs of images, and the Log-Euclidean mean of stationary velocity fields (SVFs) (Arsigny et al., 2006). It improved upon the method of Serag et al. (2012a), which is based on the free-form deformation (FFD) algorithm (Rueckert et al., 1999), and arithmetic mean of displacement fields. A longitudinal model of brain morphology is hereby obtained using Nadaraya-Watson kernel regression similar to Davis et al. (2007), with adaptive kernel size as proposed by Serag et al. (2012a).

A drawback of these techniques is the quadratic computational complexity with respect to the number of images, which means that these approaches are difficult to scale to large datasets. A more significant limitation is the reliance on accurate pairwise correspondences, and the dependence on average spatial correspondences, which can capture complex cortical folding patterns. The less accurate the individual registrations are, the less detail is preserved. An approach which iteratively refines the transformations that relate each anatomy with the average space, may be more robust towards residual misalignment.

Therefore, instead of expressing the spatio-temporal atlas in terms of the transformations relating each brain image to all others, in this work we derive a new recursive and continuous formulation based on temporal kernel regression, the transformations relating each image to the spatio-temporal atlas space, and their respective age-dependent Log-Euclidean mean. This novel approach jointly estimates mean shape and longitudinal change iteratively. A comparison of our methods to the atlas obtained with Serag et al.'s (2012a) technique demonstrates a marked increase in anatomical detail.

1.1. Overview

In section 1.2, we review main approaches to brain atlas construction. In section 1.3, we summarise related work in which an average model of the neonatal brain has been created. Our method for the construction of a spatio-temporal atlas of neonatal brain morphology is outlined in section 2. Here, we first detail the initial affine spatial normalisation of all brain images in section 2.2. The unbiased diffeomorphic registration method, which is the basis for our deformable atlas construction, is described in section 2.5. The atlas construction itself is derived in section 2.6. We compare our method to previous approaches in section 3, and discuss these results in section 4.

1.2. Brain atlas construction

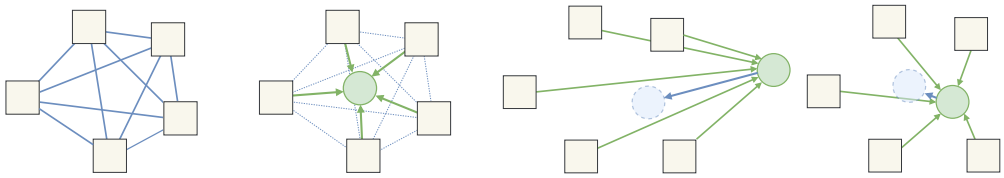
Figure 1a illustrates an atlas construction referred to herein as "pairwise" technique. First, the transformations between all pairs of images (squares) are computed (blue lines). A mean transformation (green arrows) is then derived for each image, which maps it to the average space (circle). Different approaches distinguish themselves by the transformation model, registration method used, and the definition of a mean of pairwise correspondences.

An adult brain atlas using a pairwise method with affine transformations was first constructed by Woods et al. (1998). A non-linear extension was proposed by Seghers et al. (2004) based on the arithmetic mean of displacement fields obtained by non-parametric deformable registration. Similarly, the neonatal atlas construction of Serag et al. (2012a) employs the arithmetic mean of displacement-based FFDs, which are neither inverse consistent nor guaranteed diffeomorphic. These limitations were overcome by our previous pairwise approach (Schuh et al., 2014), which is based on an efficient

diffeomorphic registration method, and the Log-Euclidean mean of SVFs.

Two steps of a "groupwise" approach, on the other hand, are depicted in figure 1b (left to right). At each iteration, the individual anatomies (squares) are registered to a reference anatomy (green circle). The average image is subsequently deformed by the average of the deformations (blue arrow) to form a new mean brain shape and intensity model (dashed blue circle), which is the reference for the next iteration (right). This iteratively minimises the distance of the reference to each anatomy until the residual deformation approaches zero. Such atlas construction resembles groupwise registration techniques (e.g. Studholme, 2003; Studholme and Cardenas, 2004; Bhatia et al., 2004, 2007; Zöllei, 2006; Noblet et al., 2008; Wu et al., 2012).

A groupwise atlas construction, where one of the brain anatomies is used as initial reference, was first proposed by Guimond et al. (1998). Unbiased methods based on large deformation models similar to large deformation diffeomorphic metric mapping (LDDMM) (Beg et al., 2005) were pioneered by Joshi et al. (2004), Lorenzen et al. (2005, 2006), and Avants et al. (2010). Of these, the initial greedy approach of Joshi et al. (2004) and its non-greedy variant of Lorenzen et al. (2005) are based on the sum of squared differences, and therefore unsuitable for the construction of an atlas of the developing brain due to wide intensity variations. The symmetric groupwise normalisation (SyGN) of Avants et al. (2010), on the other hand, is based on the symmetric normalisation (SyN) method (Avants et al., 2008), which uses local normalised cross-correlation (LNCC) as image dissimilarity measure, so it may be more directly applicable to brain images of the developing brain.



(a) "pairwise" (b) "groupwise"

Figure 1: Illustration of common approaches to single-time point brain atlas construction. A spatio-temporal model is obtained using kernel regression along the time dimension.

1.3. Neonatal brain atlases

Kazemi et al. (2007) created the first population average from seven neonatal brains scanned at 39 to 42 weeks PMA for spatial normalisation using the Statistical Parameter Mapping (SPM) toolbox (Ashburner and Friston, 1999). Validation of spatial normalisation using this template instead of the adult template created at the Montreal Neurological Institute (MNI) on another seven brain scans showed that a neonatal template results in better co-registration. The same research group also provided a symmetric template built from all 14 term-born neonates for asymmetry studies of language development (Noorizadeh et al., 2013). Recently, a bimodal neonatal brain template consisting of an average magnetic resonance (MR) and computed tomography (CT) image was added (Ghadimi et al., 2017).

A set of 14 T1-weighted (T1w), 20 T2-weighted (T2w), and 20 diffusion tensor imaging (DTI) scans acquired within 38 to 41 weeks PMA were used by Oishi et al. (2011) to create different templates after affine and non-rigid alignment. Non-rigid registration was done using dual-channel LDDMM (Ceritoglu et al., 2009) to match the mean diffusivity (MD) and fractional anisotropy (FA) maps derived from the DTI scans.

A parcellation of the cortex of 10 healthy neonates born at term, and scanned between 40 and 43 weeks PMA, which is consistent with the cortical and subcortical adult brain atlas used in FreeSurfer (Desikan et al., 2006), was done by Alexander et al. (2016). Structural T1w and T2w templates for spatial normalisation to this cortical parcellation atlas were created using SyGN (Avants et al., 2010) of the Advanced Normalization Tools (ANTs).

Kuklisova-Murgasova et al. (2011) created an affine spatio-temporal atlas consisting of T2w average images of 142 preterm-born neonates and corresponding tissue probability maps for 29 to 44 weeks PMA. Using the non-rigid pairwise approach, Serag et al. (2012b) built a spatio-temporal model of average T1w and T2w intensities from 204 preterm-born neonates, also with accompanying tissue probability maps. Makropoulos et al. (2016) furthermore non-rigidly registered automatic whole brain segmentations of 87 brain structures of 420 mostly preterm-born neonates to the age-matched T2w templates created by Serag et al. (2012b), in order to built a fine-granular probabilistic atlas. These automatic segmentations are based on 20 single-subject atlases of 15 preterm-born neonates scanned at term-equivalent age, and 5 term-born neonates, referred to as ALBERTs. Each image was manually segmented as described in Gousias et al. (2012) to create an atlas of 50 anatomical regions. These were further subdivided into cortical and subcortical parts by Makropoulos et al. (2016).

Table 1 summarises publications which created a neonatal brain atlas. Of these, only Kuklisova-Murgasova et al. (2011), Serag et al. (2012a,b), and Schuh et al. (2014) constructed a 4D model of mean shape and intensity. Atlases of the developing brain were also reviewed by Benkarim et al. (2017).

	mod.	no	age	temp.	registration
Kazemi et al. (2007)	T1w	7	39–42	–	non-rigid, SPM2
Noorizadeh et al. (2013)	T1w	14	39–42	–	non-rigid, SPM8
Ghadimi et al. (2017)	T1w	7	39–42	–	non-rigid, SPM8
	CT				
Oishi et al. (2011)	T1w	14	38-41	–	dual-channel
	T2w	20			LDDMM
	DTI				
Alexander et al. (2016)	T1w	10	40–43	–	SyGN, ANTs,
	T2w				groupwise
Gousias et al. (2012)	T2w	15	26–35	–	–
Gousias et al. (2012)	T2w	5	39–45	–	–
Shi et al. (2011)	T1w	95	39–46	–	HAMMER,
	T2w				groupwise
Kuklisova-Murgasova et al. (2011)	T2w	142	29–44	KR	affine, IRTK
Serag et al. (2012b)	T1w	204	28–44	AKR	FFD, IRTK,
	T2w				pairwise
Schuh et al. (2014)	T2w	118	28–44	AKR	SVF, pairwise
Makropoulos et al. (2016)	T2w	420	28–44	AKR	FFD, IRTK

Table 1: Atlases of the neonatal human brain. Columns from left to right: Citation, imaging modalities (mod.), number of subjects (no), post-menstrual age in weeks (age), method used for temporal regression (temp.), and spatial normalisation approach. Acronyms:

KR: kernel regression with constant kernel width; AKR: adaptive kernel regression; HAMMER: Hierarchical Attribute Matching Mechanism for Elastic Registration (Shen and Davatzikos, 2002); IRTK: Image Registration ToolKit (Rueckert et al., 1999).

2. Material and methods

2.1. Data acquisition

The image data was collected at St. Thomas Hospital, London, on a Philips 3T scanner using a dedicated 32 channel neonatal head coil (Hughes et al., 2017). T2w images were obtained using a turbo spin echo (TSE) sequence, acquired in two stacks of 2D slices in sagittal and axial planes, using parameters: TR = 12 s, TE = 156 ms, SENSE factor 2.11 (axial) and 2.58 (sagittal). Overlapping slices with resolution $0.8 \times 0.8 \times 1.6$ mm^3 were acquired. During motion corrected reconstruction (Cordero-Grande et al., 2017), the images are resampled to an isotropic voxel size of 0.5 mm. T1w images were acquired using an inversion recovery TSE sequence at the same resolutions with TR = 4.8 s, TE = 8.7 ms, SENSE factor 2.26 (axial) and 2.66 (sagittal). Babies were imaged without sedation during natural sleep. All images were checked for abnormalities by a paediatric neuroradiologist.

2.2. Image preprocessing

The reconstructed volumes were preprocessed using a pipeline designed for neonatal brain images, including an initial rough brain extraction (Smith, 2002) and bias field correction (Tustison et al., 2010). The bias corrected images are segmented into 87 regions of interest (ROIs) using an extension of the developing region annotation with expectation maximisation (Draw-EM) algorithm developed by Makropoulos et al. (2014). For details, including references to the software, please refer to Makropoulos et al. (2017).

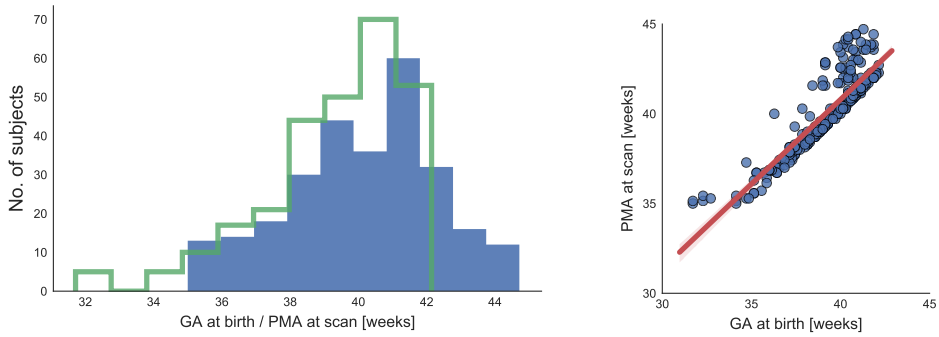


Figure 2: The histogram plot on the left shows the distribution of gestational age (GA) at birth (steps) and PMA at scan (bars) of the used dataset. The scatter plot on the right is overlayed with the regression line mapping from GA at birth to PMA at scan. Most neonates were scanned shortly after birth, with an average time to scan of 0.8 weeks.

2.3. Image selection

Out of 495 processed brain scans of 474 participants, we selected images of 275 neonates. Their age distributions are shown in figure 2. This set of images was obtained via a process of exclusion based on (a) manual image quality score between 1 (poor) and 4 (good) (Makropoulos et al., 2017), excluding images with score 1 or 2, (b) availability of both T1w and T2w images, (c) exclusion of follow up scans, (d) maximum time to scan after birth of 4 weeks, and (e) minimum age at scan of 35 weeks PMA.

2.4. Unbiased global normalisation

Pose differences of the imaged anatomies are removed by a rigid alignment to the 40 weeks template of the atlas constructed by Serag et al. (2012a). Differences in brain size are also removed at this global normalisation stage, such that the template construction based on deformable registration can

focus on local shape differences. The so constructed template brain images are rescaled afterwards to reflect global volume changes.

The global normalisation and age-dependent rescaling is similar to the atlas construction proposed by Kuklisova-Murgasova et al. (2011), where only affine registration was employed to create average brain images and tissue probability maps of the preterm-born neonate. For this, the authors use a common reference image for the affine registrations. Bias towards the reference anatomy is hereby accounted for by composing each affine transformation with the inverse of their Log-Euclidean mean (Arsigny et al., 2007). Under the assumption of ideal, and thereby inverse-consistent and transitive affine transformations, the Log-Euclidean mean is independent of the chosen reference (Pennec and Arsigny, 2013). Bias towards the reference is thus removed by mapping it to the barycentre of all observed anatomies.

In practice, however, affine registration is more accurate for anatomies that are more similar to the reference anatomy. The choice of which of the two images is to remain fixed, and which is being transformed and resampled, further affects a pairwise registration (c.f. Cachier and Rey, 2000). These biases result in inconsistent transformations that are neither inverse-consistent nor transitive (Woods et al., 1998). Consequently, some anatomical bias is retained when a reference image is used for affine pre-alignment.

To eliminate this remaining bias, we perform affine registrations between all pairs of images, i.e., select each image in turn as reference. A mean transformation different from the Log-Euclidean mean employed in our work, which is derived from affine transformations between all pairs of images, was first employed by Woods et al. (1998) to create an adult brain template. Due

to lack of inverse-consistency, some residual global misalignment remains, however. We therefore repeat the steps detailed below for another iteration.

Note that, while it would have been possible to refine through less costly, direct registrations of all images to the initial average, the benefits of a robust and unbiased pre-alignment of the images, which in turn impacts the subsequent deformable registrations, outweigh the added computational cost.

Formally, the k-th homogeneous coordinate transformation, which maps image I_i to the unbiased global reference space, is given by the recursion

$$\boldsymbol{A}_i^{(0)} = \begin{pmatrix} s_i \boldsymbol{R}_i & \boldsymbol{d}_i \\ \boldsymbol{0}^T & 1 \end{pmatrix}, \quad \boldsymbol{A}_i^{(k)} = \exp\left(\frac{1}{n}\sum_{j=1}^{n} \log \boldsymbol{A}_{ij}^{(k)}\right) \boldsymbol{A}_i^{(k-1)} \tag{1}$$

where $s_i \boldsymbol{R}_i \in \mathbb{R}^{3\times 3}$ is a rotation followed by isotropic scaling, and $\boldsymbol{d}_i \in \mathbb{R}^3$ is a translation. These are found by initial rigid alignment of the i-th image to a selected reference image. The pairwise transformations $\boldsymbol{A}_{ij}^{(k)}$ are found by affine registration, as detailed in Appendix B. The matrix exponential and logarithm are denoted by exp and log (Alexa, 2002; Arsigny et al., 2007).

After K iterations,

$$\bar{I}_{\text{global}}(\boldsymbol{x}) = \frac{1}{n}\sum_{i=1}^{n} \tilde{I}_i \tag{2}$$

is the unbiased population-specific average image, where $\boldsymbol{A}_i \coloneqq \boldsymbol{A}_i^{(K)}$, and

$$\tilde{I}_i(\boldsymbol{x}) = \frac{I_i\left(\boldsymbol{A}_i^{-1}\boldsymbol{x}\right) - \mu_i}{\sigma_i} \tag{3}$$

are the brain images after global spatial normalisation, with zero mean intensity and unit variance, whereby the original mean intensity and standard deviation are denoted by μ_i and σ_i, respectively. The normalisation of intensities by z-score compensates for global differences in intensity distributions.

2.5. Diffeomorphic image registration

The atlas construction is based on a fast and unbiased stationary velocity free-form deformation (SVFFD) algorithm detailed in the following.

2.5.1. Parametric deformation model

The diffeomorphic mapping between a pair of images is given by the group exponential of a SVF, which in our work is a parametric spline function $\boldsymbol{v}(\boldsymbol{x})$ as defined in Appendix A. A spatial mapping $\boldsymbol{\Phi} : \mathbb{R}^3 \to \mathbb{R}^3$ is then given by the group exponential, i.e., $\boldsymbol{\Phi} = \exp(\boldsymbol{v})$, which is the solution at time $t = 1$ of the stationary ordinary differential equation (ODE) (Arsigny et al., 2006)

$$\frac{\partial \boldsymbol{\varphi}(\boldsymbol{x}, t)}{\partial t} = \boldsymbol{v} \circ \boldsymbol{\varphi}(\boldsymbol{x}, t) \tag{4}$$

with initial condition $\boldsymbol{\varphi}(\boldsymbol{x}, 0) = \boldsymbol{x}$. Hence, a SVF is the infinitesimal generator of a one-parameter subgroup of diffeomorphisms $\boldsymbol{\varphi}_t = \exp(t \cdot \boldsymbol{v})$.

The cubic B-spline parametrisation of the SVF was first proposed for brain image registration by Modat et al. (2012). It inherits some of the benefits of the not-guaranteed-diffeomorphic FFD model (Sederberg and Parry, 1986; Rueckert et al., 1999), namely an implicit smooth interpolation of the velocity field, and a parametric representation, which reduces the number of parameters and allows for an analytic derivation. Note that even though the support of the basis functions is limited, a control point may still affect the deformation of a point outside its local support region. This can be explained by considering the trajectory of a particle subjected to the flow induced by the SVF. Even when this trajectory originates at a position outside the local support region of a control point, it may enter it at a later time point $t > 0$.

The flow $\boldsymbol{\varphi}_t$, and therefore the group exponential, can be computed using a numerical integration scheme. For example, Pai et al. (2015) use a forward Euler integration. The advantage of a SVF over a time-varying velocity field is due to a method known as scaling and squaring (SS), i.e.,

$$\exp(\boldsymbol{v}) = \exp\left(2^{-m}\boldsymbol{v}\right)^{2^m} \tag{5}$$

where $\exp(\boldsymbol{v})^2 = \exp(\boldsymbol{v}) \circ \exp(\boldsymbol{v})$ and $\exp(2^{-m}\boldsymbol{v}) \approx \mathrm{Id} + 2^{-m}\boldsymbol{v}$ (Arsigny et al., 2006): Instead of m Euler steps, only $\log_2 m$ squaring steps are needed. To further reduce the computational cost of the integration, a SS on the coarser control point lattice of $\boldsymbol{v}$ was proposed by Modat et al. (2012). Different methods for computing $\exp(\boldsymbol{v})$ were compared by Bossa et al. (2008).

In Modat et al. (2012), it is noted that the SVF must be sufficiently smooth for the exponential to be a diffeomorphic mapping. Regularisation terms based on the Jacobian determinant of the SVF are therefore used. Given the cubic B-spline parametrisation, we can always find a minimum number of squaring steps, which guarantees positive Jacobian determinants of the scaled velocity field. This is because the Jacobian determinant of $\boldsymbol{\Phi}$ at $\boldsymbol{x}$ is the product of determinants along the trajectory of a particle originating at this point. By choosing the integration step such that the maximum scaled coefficient is less than 0.4 times the minimum control point spacing, all factors are positive, and a diffeomorphic mapping is ensured (c.f. Choi and Lee, 2000; Rueckert et al., 2006). Folding and inconsistencies between $\exp(\boldsymbol{v})$ and $\exp(-\boldsymbol{v})$ may still be observed due to the numerical integration, i.e., the errors introduced during the composition of discrete deformation fields at each squaring step. Numerical errors caused by the integration are reduced by avoiding extreme Jacobian determinant values (Modat et al., 2012).

2.5.2. Symmetric energy function

Having defined the SVFFD model, we now detail our unbiased pairwise image registration method based on a symmetric energy formulation. Unlike the closely related method of Modat et al. (2012), where forward and backward mappings are parametrised by two separate SVFFDs, our formulation is based on a single spline function. Similar to the consistent-midpoint-cost proposed for non-stationary LDDMM by Beg and Khan (2007), we define

$$\begin{aligned}\mathcal{E}(I_0, I_1; \boldsymbol{v}) = \; & \mathcal{D}\left(I_0 \circ \exp\left(-\boldsymbol{v}/2\right), I_1 \circ \exp\left(\boldsymbol{v}/2\right)\right) \\ & + w_{\mathrm{L}} \; \mathcal{S}_{\mathrm{LE}}\left(\boldsymbol{v}\right) + w_{\mathrm{B}} \; \mathcal{S}_{\mathrm{BE}}\left(\boldsymbol{v}\right) + \mathcal{P}_{\mathrm{JAC}}\left(\boldsymbol{v}\right)\end{aligned} \tag{6}$$

where $I_{\{0,1\}}$ are the images to be registered. For neonatal brain image registration, suitable choices of image dissimilarity measures $\mathcal{D}$ are normalised mutual information (NMI), and LNCC. In case of LNCC, our symmetric data term is equivalent to the data matching term of the non-parametric LCC LogDemons approach (Lorenzi et al., 2013). To explicitly enforce smoothness of the SVF, regularisation terms based on first and second order derivatives are employed. These are the elastic potential of the spline-based SVF (similar to the non-parametric viscous-fluid model (Christensen et al., 1996)), i.e.,

$$\mathcal{S}_{\mathrm{LE}}(\boldsymbol{v}) = \int_{\Omega} \frac{\mu}{4} \sum_{j,k=1}^{3} \left(\frac{\partial v_k}{\partial x_j} + \frac{\partial v_j}{\partial x_k}\right)^2 + \frac{\lambda}{2} \left(\sum_{j=1}^{3} \frac{\partial v_j}{\partial x_j}\right)^2 d\boldsymbol{x} \tag{7}$$

and the bending energy used in the FFD algorithm (Rueckert et al., 1999), but applied here to the SVF instead of a displacement field, i.e.,

$$\mathcal{S}_{\mathrm{BE}}(\boldsymbol{v}) = \frac{1}{2} \int_{\Omega} \sum_{i,j,k=1}^{3} \left(\frac{\partial^2 v_k(\boldsymbol{x})}{\partial x_i \partial x_j}\right)^2 d\boldsymbol{x} \tag{8}$$

The non-negative weights w_{L} and w_{B} determine the trade-off between the smoothness imposed on the SVF and the image matching term. In addition

to the explicit smoothness constraints, penalty terms based on the Jacobian determinants of the SVF can be added. A weighted sum of Jacobian-based penalty terms is denoted in equation (6) by $\mathcal{P}_{\text{JAC}}(\boldsymbol{v})$. It should be noted that the bending energy and linear elasticity constraints are symmetric due to the squaring of derivatives, i.e., $\mathcal{S}_{\text{BE}}(\boldsymbol{v}) = \mathcal{S}_{\text{BE}}(-\boldsymbol{v})$ and $\mathcal{S}_{\text{LE}}(\boldsymbol{v}) = \mathcal{S}_{\text{LE}}(-\boldsymbol{v})$. Jacobian-based penalties are generally not symmetric, however, and must be evaluated for both the positive and negative SVF. Therefore,

$$\begin{aligned}\mathcal{P}_{\text{JAC}}(\boldsymbol{v}) = \frac{w_{\text{J}}}{2}\Big(\mathcal{P}_{\text{VP}}(\text{Id}+\boldsymbol{v}) + \mathcal{P}_{\text{VP}}(\text{Id}-\boldsymbol{v})\Big) \\ + \frac{w_{\text{T}}}{2}\Big(\mathcal{P}_{\text{TP}}(\text{Id}+\boldsymbol{v}) + \mathcal{P}_{\text{TP}}(\text{Id}-\boldsymbol{v})\Big)\end{aligned} \tag{9}$$

where $\mathcal{P}_{\text{VP}}$ and $\mathcal{P}_{\text{TP}}$ are penalties of the deformation induced by the instantaneous velocity field and its approximate inverse, respectively, i.e.,

$$\mathcal{P}_{\text{VP}}(\boldsymbol{\Psi}) = \int_{\Omega} \log\left(J_{\boldsymbol{\Psi}}(\boldsymbol{x})\right)^2 \, d\boldsymbol{x} \tag{10}$$

$$\mathcal{P}_{\text{TP}}(\boldsymbol{\Psi}) = \int_{\Omega} \mathbb{1}_{\{\boldsymbol{x} \mid J_{\boldsymbol{\Psi}}(\boldsymbol{x}) < \gamma\}}(\boldsymbol{x}) \left(\frac{J_{\boldsymbol{\Psi}}(\boldsymbol{x})^2}{\gamma^2} + \frac{\gamma^2}{J_{\boldsymbol{\Psi}}(\boldsymbol{x})^2} - 2\right) d\boldsymbol{x} \tag{11}$$

where the Jacobian determinant of a mapping $\boldsymbol{\Psi}$ is denoted by $J_{\boldsymbol{\Psi}}(\boldsymbol{x})$, and the indicator function $\mathbb{1}_{\mathcal{X}}(\boldsymbol{x})$ is one for $\boldsymbol{x} \in \mathcal{X}$, and zero otherwise. While $\mathcal{P}_{\text{VP}}$ penalises both excessive contraction and expansion, $\mathcal{P}_{\text{TP}}$ is active for determinants $J < \gamma < 1$. It was used by Rueckert et al. (2006) as topology-preservation constraint in case of the FFD algorithm. Note that $\mathcal{P}_{\text{VP}}$ is undefined for $J \leq 0$. Similarly, $\mathcal{P}_{\text{TP}}$ is undefined for $J = 0$, and has the undesired property that it can prevent already negative determinants from being corrected for, since it is monotone increasing for $J \in [-\gamma, 0)$.

Different implementations may handle the case of non-positive Jacobian determinants differently. An approach found in the Image Registration ToolKit (IRTK)[1] is to simply clamp the values below a threshold ϵ, e.g., $\epsilon = 10^{-6}$. This has the disadvantage that the value of the energy function does not change as long as the Jacobian determinant is less than ϵ. This may result in a premature termination of the iterative optimisation, because the total energy may not have improved when stepping along the gradient direction. A different strategy to avoid undefined values of $\mathcal{P}_{\mathrm{VP}}$ is taken by NiftyReg. After every gradient step, non-positive Jacobian determinants are detected, and an iterative unfolding, which aims to remove these based on a separate maximisation of their values, is performed (Modat et al., 2012). This strategy may fail to resolve all non-positive values, which then results in undefined penalties and may in consequence prematurely terminate the registration.

We instead define both Jacobian-based penalty terms as piecewise functions, where the functions above a small positive threshold ϵ are identical to previously defined penalties, but the penalty below ϵ corresponds to a linear continuation, which monotonically increases with decreasing determinant. In order to obtain a $\mathbb{C}^1$-continuous penalty, the slope of this linear penalty function is equal to $\nabla_J \mathcal{P}_\star$ at $J = \epsilon$, and the intercept with the vertical axis at $J = 0$ is determined such that the two pieces below and above ϵ attain the same value for $J = \epsilon$. Using $\mathcal{P}_{\mathrm{TP}}$ avoids a separate unfolding step, and the registration task is determined by a single well-defined energy minimisation.

[1] https://github.com/BioMedIA/IRTK (last accessed: Oct 13, 2017)

2.5.3. Energy minimisation

The energy given by equation (6) is minimised using a conjugate gradient descent with a greedy line search. The derivatives of the penalty terms with respect to the coefficients of the SVFFD are equivalent to the derivations for the displacement-based FFD.

The gradients of the Jacobian penalties are

$$\frac{\partial \mathcal{P}_\star}{\partial \boldsymbol{c}} = \int_\Omega Df\left(J_{\mathrm{Id}\pm\boldsymbol{v}}(\boldsymbol{x})\right) \frac{\partial J_{\mathrm{Id}\pm\boldsymbol{v}}(\boldsymbol{x})}{\partial \boldsymbol{c}}\, d\boldsymbol{x} \tag{12}$$

where

$$Df(J) = \begin{cases} \frac{2}{J}\log(J) & \text{for } \mathcal{P}_\star = \mathcal{P}_{\mathrm{VP}} \\ 2\left(\frac{J}{\gamma^2} - \frac{\gamma^2}{J^3}\right) & \text{for } \mathcal{P}_\star = \mathcal{P}_{\mathrm{TP}} \end{cases} \tag{13}$$

and $\frac{\partial J_{\mathrm{Id}\pm\boldsymbol{v}}}{\partial c}$ is obtained using Jacobi's formula as in Modat et al. (2012).

The derivatives of the image dissimilarity term can be written as

$$\frac{\partial \mathcal{D}}{\partial \boldsymbol{c}} = \int_\Omega \left(\frac{\partial \mathcal{D}}{\partial I_0}\frac{\partial I_0}{\partial \boldsymbol{\Phi}_0}\frac{\partial \boldsymbol{\Phi}_0}{\partial \boldsymbol{v}} + \frac{\partial \mathcal{D}}{\partial I_1}\frac{\partial I_1}{\partial \boldsymbol{\Phi}_1}\frac{\partial \boldsymbol{\Phi}_1}{\partial \boldsymbol{v}} \right) \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{c}}\, d\boldsymbol{x} \tag{14}$$

where $\boldsymbol{\Phi}_0 = \exp(-\boldsymbol{v}/2)$, and $\boldsymbol{\Phi}_1 = \exp(\boldsymbol{v}/2)$. The factor in parenthesis corresponds to the sum of voxel-wise non-parametric gradients of the image dissimilarity measure with respect to a change in the SVF for each reference point $\boldsymbol{x} \in \Omega$. The partial derivatives $\nabla_{I_{\{0,1\}}}\mathcal{D}$ are given for NMI based on the Parzen window estimation of the joint probability distribution using cubic B-spline kernels in Modat et al. (2009), and for LNCC in Avants et al. (2008).

The numerical approximation of the derivatives of the exponential maps $\nabla_{\boldsymbol{v}}\boldsymbol{\Phi}_{\{0,1\}}$ depends on the integration method. The formulas for the Euler integration are equivalent to the temporal diffeomorphic free-form deformation

(TDFFD) algorithm of De Craene et al. (2012), which is a variant of LDDMM based on a B-spline parametrisation. While more accurate than SS, a Runge-Kutta integration is too time consuming for an efficient gradient evaluation. Faster approximations based on SS are outlined in Ashburner (2007) and Modat et al. (2012). Here, we use an even simpler approximation based on the assumption that the gradient can be considered to be an instantaneous velocity field. The same is assumed by the viscous fluid method (Christensen et al., 1996) and diffeomorphic Demons (Vercauteren et al., 2009). The current SVFFD is composed with this SVF using the Baker-Campbell-Hausdorff (BCH) formula (c.f. Bossa et al., 2007; Schuh et al., 2014), i.e.,

$$t_i \frac{\partial \mathcal{D}}{\partial \boldsymbol{v}} = \boldsymbol{w}_i + \frac{1}{2}\left[t_i \boldsymbol{v}, \boldsymbol{w}_i\right] + \frac{1}{12}\left[t_i \boldsymbol{v}, \left[t_i \boldsymbol{v}, \boldsymbol{w}_i\right]\right] + \frac{1}{12}\left[\left[t_i \boldsymbol{v}, \boldsymbol{w}_i\right], \boldsymbol{w}_i\right] \tag{15}$$

where $t_0 = -0.5$, $t_1 = 0.5$, and $\boldsymbol{w}_i = \nabla_{\boldsymbol{\Phi}_i} \mathcal{D}$. This computation is similar to log-domain Demons (Vercauteren et al., 2008). By interpolating all vector fields (including the Lie bracket $[\cdot, \cdot]$) by cubic B-spline functions defined on the same lattice as $\boldsymbol{v}$, the approximated image dissimilarity gradient with respect to $\boldsymbol{c}$ is given directly by the spline coefficients of the right-hand side of equation (15). Following Lorenzi et al. (2013), we use the zeroth order approximation of the BCH formula. The coefficients of $\boldsymbol{w}_i$ are obtained via convolution of $\nabla_{\boldsymbol{\Phi}_i} \mathcal{D}$ with the cubic B-spline kernel (Modat et al., 2010).

2.6. Spatio-temporal atlas construction

Our proposed spatio-temporal atlas construction is illustrated in figure 3. In this schematic drawing, the horizontal axis corresponds to the temporal domain and the image planes orthogonal to it depict the 3D space of the atlas. Each image plane pictures an atlas template to which a given sample brain

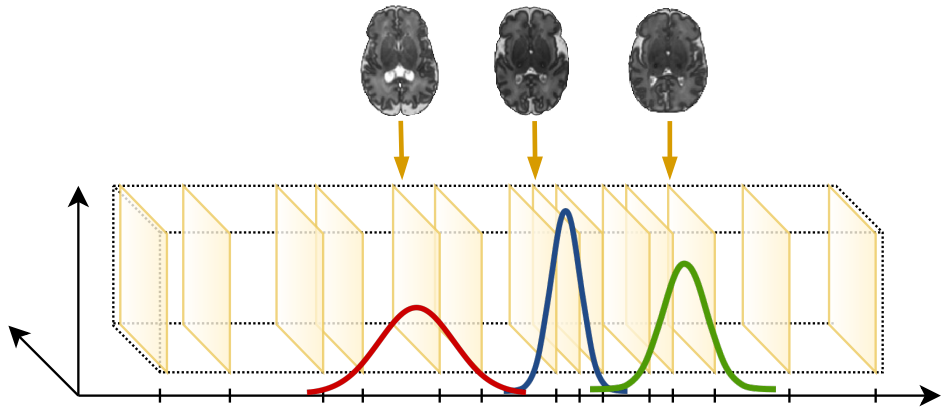


Figure 3: Illustration of atlas construction with time corresponding to the horizontal axis, and cross-sectional image planes depicting the spatial domain. Exemplary regression kernels are shown as overlaid curves. Arrows illustrate the subject-to-atlas mappings.

image acquired at time t_i is registered to, in order to infer the transformation which relates it to the atlas space. The temporal density of the observations is therefore illustrated by the spacing of the cross-sectional image planes along the time axis. The adaptation of the kernel width to this age distribution is shown by three exemplary Gaussian curves overlaid over the diagram.

Once the recursive mean intensity function in section 2.6.1 is defined, we modify this definition in section 2.6.2 to also incorporate an estimate of the longitudinal deformations of the atlas space from one time point to another. The steps which are iterated to construct a spatio-temporal atlas from n neonatal brain images of different ages are then summarised.

2.6.1. Recursive definition of spatio-temporal atlas

The mean intensity and brain shape template at iteration k is given by

$$\bar{I}_k(t) = \sum_{i=1}^{n} w_i(t) \left(\tilde{I}_i \circ \boldsymbol{T}_{k,i}^{-1}(t) \right) \tag{16}$$

where $\tilde{I}_i(\boldsymbol{x})$ is the i-th globally normalised image, and $\boldsymbol{T}_{k,i}$ is the current estimate of the diffeomorphism relating it to the spatio-temporal atlas space.

The weights used for temporal regression are normalised Gaussian kernels centred at time point t, with age-dependent standard deviation σ_t, i.e.,

$$w_i(t) = \frac{g_i(t)}{\sum_{j=1}^{n} g_j(t)} \tag{17}$$

where

$$g_i(t) = \frac{1}{\sigma_t\sqrt{2\pi}} \exp\left(\frac{-(t_i - t)^2}{2\sigma_t^2}\right) \tag{18}$$

These are truncated below a small threshold to limit their global extent. An adaptive kernel width has been shown by Serag et al. (2012a) to be advantageous over a fixed kernel width for the construction of a spatio-temporal atlas of the early developing brain. This is due to the age at scan distribution of most neonatal datasets, where neonates are ideally imaged close to birth and most are born around term between 37 to 40 weeks gestation. Adapting the kernel width to be narrower at term, combined with a data selection which eliminates sharp peaks in the age distribution, effectively smooths the age distribution and makes it more uniform. This results in temporally more consistent templates, given a more even distribution of anatomical samples. The weighted average given by equation (16) shifts the barycentre of deformed brain anatomies towards those samples that are closer in age to the atlas time point. Images of an earlier or later developmental phase have less or no influence. This is an important property given the large shape differences at different ages, and the aim to construct an atlas that resembles the mean brain shape at each stage of brain development during early life.

The subject-to-atlas deformations are modelled as the composition

$$\boldsymbol{T}_{k,i}(t) = \bar{\boldsymbol{\phi}}_k\,(t) \circ \boldsymbol{\phi}_{k,i} \tag{19}$$

where

$$\bar{\boldsymbol{\phi}}_k(t) = \exp\left(-\sum_{i=1}^{n} w_i(t) \log \boldsymbol{\phi}_{k,i}\right) \tag{20}$$

is the weighted Log-Euclidean mean (Arsigny et al., 2006) of inverse subject-to-atlas deformations obtained through SVFFD registration of $\tilde{I}_i$ to the age-matched average image of the previous iteration, i.e., $\bar{I}_{k-1}(t_i)$. The logarithmic maps are hereby given directly by the computed SVFs. The cross-sectional diffeomorphism mapping $\tilde{I}_i$ into the atlas space at t_i is given by

$$\boldsymbol{\phi}_{k,i} = \exp\left(\boldsymbol{v}_{k,i}\right) \tag{21}$$

where

$$\boldsymbol{v}_{k,i} = \arg\min_{\boldsymbol{v}} \mathcal{E}(\tilde{I}_i, \bar{I}_{k-1}(t_i); \boldsymbol{v}) \tag{22}$$

is the (local) minimum of the energy given by equation (6), and $\boldsymbol{v}_{0,i} = \mathbf{0}$.

2.6.2. Joint estimation of mean shape and growth

In the previous section, the transformations relating each brain image to different time points were considered to have the same anatomical co-domain. This neglects the sequence of residual mean deformations $\bar{\boldsymbol{\phi}}_k$ that deform the atlas space at each time point differently for the mean shape to flow towards the weighted barycentre of the nearby observations at the respective age. As illustrated in figure 4, we can track the sequence of space deformations applied at each time t, and use these to derive a longitudinal coordinate map,

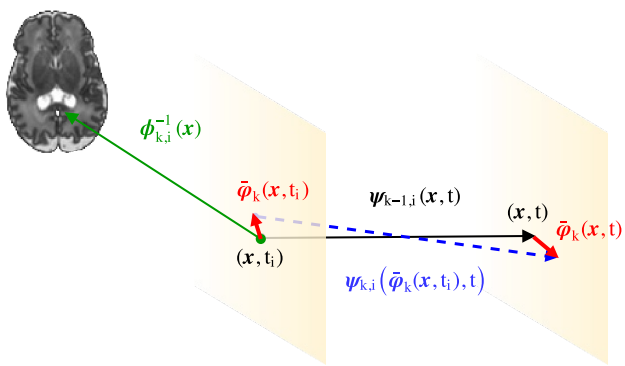


Figure 4: Recursive estimation of mean growth via composition of space deformations. The cross-sectional SVFFD $\boldsymbol{\phi}_{k,i}$ (green arrow), computed by non-rigid registration of the i-th image to the atlas template at age t_i (left plane), is composed with the current longitudinal deformation $\boldsymbol{\psi}_{k-1,i}$ (black arrow) to map the image to time t (right plane). This longitudinal estimate of mean growth evolves together with the spatial anatomical coordinates at the two time points which are deformed at iteration k by the age-specific mean deformations $\bar{\boldsymbol{\varphi}}_k$ (red arrows), resulting in the new estimate $\boldsymbol{\psi}_{k,i}$ (blue arrow).

which relates an observed age t_i to any other time point in order to correct for this anatomical mismatch in the co-domain, i.e.,

$$\begin{aligned} \boldsymbol{\psi}_{k,i}(t) &= \bar{\boldsymbol{\varphi}}_k(t) \circ \boldsymbol{\psi}_{k-1,i}(t) \circ \bar{\boldsymbol{\varphi}}_k^{-1}(t_i) \\ &= \left(\prod_{s=1}^{k} \bar{\boldsymbol{\varphi}}_s(t) \right) \left(\prod_{s=1}^{k} \bar{\boldsymbol{\varphi}}_s(t_i) \right)^{-1} \end{aligned} \tag{23}$$

where $\prod$ denotes a sequence of function compositions, and $\bar{\boldsymbol{\varphi}}$ are the adjusted Log-Euclidean mean transformations defined below. The non-recursive definition follows from the assumption that the initial longitudinal deformation between time t_i and any other time point is the identity transformation, i.e., $\boldsymbol{\psi}_{0,i}(t) = \mathrm{Id}$. Because each transformation is the exponential map of a SVF, all compositions can be approximated using the BCH formula, such that the composite transformation is another SVFFD.

The composition of the subject-to-atlas transformation given by (21), which is obtained by registering the i-th image to the template $\bar{I}_{k-1}(t_i)$, with the longitudinal deformation (23), yields an age-specific deformation, i.e.,

$$\boldsymbol{\varphi}_{k,i}(t) = \boldsymbol{\psi}_{k-1,i}(t) \circ \boldsymbol{\phi}_{k,i} \tag{24}$$

For a set of discrete time points, these transformations may be found by separate registrations of the i-th image to the different atlas templates as in Gholipour et al. (2017). However, obtaining it instead via composition with an estimate of the longitudinal atlas deformation, explicitly enforces temporal consistency and reduces the number of required registrations to n.

The residual atlas deformation is now given by the Log-Euclidean mean of the age-adjusted atlas-to-subject transformations. It is thus given by

$$\bar{\boldsymbol{\varphi}}_k(t) = \exp\left(-\sum_{i=1}^{n} w_i(t) \log \boldsymbol{\varphi}_{k,i}(t)\right) \tag{25}$$

which are further used to redefine the total age-dependent deformations, i.e.,

$$\begin{aligned} \boldsymbol{T}_{k,i}(t) &= \bar{\boldsymbol{\varphi}}_k(t) \circ \boldsymbol{\varphi}_{k,i}(t) \\ &= \bar{\boldsymbol{\varphi}}_k(t) \circ \boldsymbol{\psi}_{k-1,i}(t) \circ \boldsymbol{\phi}_{k,i} \end{aligned} \tag{26}$$

This concludes the definition of the coordinate maps involved in this new recursive spatio-temporal atlas construction based on the diffeomorphic SVFFD algorithm and the Log-Euclidean mean. As mentioned, all function compositions are efficiently approximated using the BCH formula, resulting in a compact representation of all coordinate maps. Moreover, while the mean intensity and shape template can be evaluated for any time t, during the atlas construction, only those images corresponding to the discrete time

points t_i are needed. The result of the atlas construction is not only the spatio-temporal atlas of mean shape and intensity, and the transformations mapping each anatomy into the atlas space, but also an estimate of the longitudinal deformations corresponding to mean growth. It should be noted, though, that these deformations primarily reflect coarse anatomical changes, which is in part due to the approximations and the remaining misalignment after deformable registration, but also because finer details are the result of the temporal kernel regression of intensities at each atlas coordinate. The final mean shape is a combination of both, the sequence of Log-Euclidean mean deformations, and the regression of the intensities.

In summary, the atlas construction iterates the following steps:

1. Generate $|\{t_i\}| \leq n$ template images $\bar{I}_{k-1}(t_i)$ given $\tilde{I}_i$ and $\boldsymbol{T}_{k-1,i}$ (16).

2. Compute $\boldsymbol{\phi}_{k,i} = \exp(\boldsymbol{v}_{k,i})$ from i-th image to template at t_i (21).

3. Compose maps $\boldsymbol{\psi}_{k-1,i}\, \forall t \in \{t_j | w_i(t_j) > 0\}$ given $\boldsymbol{\psi}_{k-2,i}$ and $\bar{\boldsymbol{\varphi}}_{k-1}$ (23).

4. Compose maps $\boldsymbol{\phi}_{k,i}$ with $\boldsymbol{\psi}_{k-1,i}$ $\forall t \in \{t_j | w_i(t_j) > 0\}$ (24).

5. Compute Log-Euclidean means $\bar{\boldsymbol{\varphi}}_k$ for each observed time t_i (26).

After about 5–10 iterations (or convergence based on residual deformations), template images of mean intensity and shape can be generated for a set of discrete time points together with the longitudinal deformations $\boldsymbol{\psi}$ between consecutive templates. The mean growth deformations can be used to initialise a subsequent longitudinal registration of the atlas time series, for example, using the TDFFD method (De Craene et al., 2012) with initial velocities given by the SVFs of the SVFFDs.

3. Results

3.1. Evaluation criteria

A number of measures were used to assess the quality of the pairwise registrations and the alignment of the brain anatomies in the atlas space. A more accurate alignment directly translates into sharper brain templates. Different subsets of criteria are used to select registration parameters, and to compare atlases built with different approaches. First, we define the per-voxel/-structure measures, before detailing the calculation of the summarising average measures.

3.1.1. Label entropy

A segmentation-based voxel-wise measure for evaluation of spatial normalisation procedures was proposed by Robbins et al. (2004). For this, the entropy of the labels assigned to a point in the standard space is evaluated. When all labels are in agreement, i.e., the atlas label is uniquely determined, the entropy is zero. It attains its maximum when all labels are equally likely, i.e., when each segmentation suggests a different label for a given atlas coordinate. The label entropy at atlas point $\boldsymbol{x}$ at time point t is given by

$$H_{\mathcal{L}}(\boldsymbol{x}, t) = -\sum_{c \in \mathcal{L}} \left(\frac{h_c(\boldsymbol{x}, t)}{n} \right) \log \left(\frac{h_c(\boldsymbol{x}, t)}{n} \right) \tag{27}$$

where $h_c(\boldsymbol{x}, t)$ is the number of times class label c is propagated from one of the brain segmentations by their respective image-to-atlas transformations using nearest neighbour interpolation, and $\mathcal{L}$ denotes the set of integral labels, i.e., either the set of Draw-EM tissue classes $\mathcal{L}_{\mathrm{T}} = \{1, \ldots, 9\}$ or the set of Draw-EM brain structure labels $\mathcal{L}_{\mathrm{S}} = \{1, \ldots, 87\}$. We refer to the entropy of

the tissue class labels as tissue (class) entropy, and the entropy of the whole brain structure labels as structure (label) entropy, respectively.

3.1.2. Intensity entropy

The entropy of intensity samples at atlas coordinate $(\boldsymbol{x}, t)$ quantifies the uncertainty of the mean intensity estimate at this point, i.e.,

$$H(\boldsymbol{x}, t) = -\sum_{j=1}^{b} \left(\frac{h_j(\boldsymbol{x}, t)}{n} \right) \log \left(\frac{h_j(\boldsymbol{x}, t)}{n} \right) \tag{28}$$

where a histogram $\boldsymbol{h}(\boldsymbol{x}, t) \in \mathbb{N}_0^b$ with b bins is used to estimate the probability distribution of the normalised intensities. The more diverse the intensities are, the more uncertain is the mean estimate.

3.1.3. Standard deviation

Another measure of voxel-wise intensity variation, also used by Gholipour et al. (2017) to quantify atlas sharpness, is the standard deviation, i.e.,

$$S(\boldsymbol{x}, t) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(\tilde{I}_i(\boldsymbol{x}, t) - \bar{I}(\boldsymbol{x}, t) \right)^2} \tag{29}$$

The lower this value, the lower is the standard error of the mean intensity.

3.1.4. Gradient magnitude

The magnitude of the gradient of mean intensities in the atlas space, i.e.,

$$G(\boldsymbol{x}, t) = \left\| \nabla \bar{I}(\boldsymbol{x}, t) \right\| \tag{30}$$

where ∇I is approximated using central differences, is proportional to the sharpness of image edges. When boundaries of structures with distinct appearance are aligned well, we observe a strong gradient at this boundary.

This criterion was also used by Gholipour et al. (2017). To specifically quantify the sharpness of the cortex in the mean images, we assess the gradient magnitude of the cGM probability map of the atlas.

3.1.5. Pairwise overlap

The overlap measure quantifies the accuracy of the pairwise registrations using the automatic whole brain segmentation. Given the number of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) label assignments for each structure and pair of segmentations of the i-th and j-th images propagated to the atlas at time point t_l, the mean Dice similarity coefficient (DSC) (Dice, 1945) of the c-th structure is given by

$$DSC_l(c) = \frac{1}{n^2} \sum_{i,j=1}^{n} \frac{2\, TP_{l,ij}(c)}{2\, TP_{l,ij}(c) + FP_{l,ij}(c) + FN_{l,ij}(c)} \tag{31}$$

3.1.6. Temporal consistency

To quantify the effect of the choice of temporal kernel width, we define a measure of temporal consistency (TC). The TC of structure c sampled at n_t ordered time points t_l (i.e., $t_l < t_{l+1}$) is the generalised DSC given by

$$TC(c) = \sum_{l=1}^{n_\mathrm{t}-1} \frac{2|\boldsymbol{p}_c(t_l) \cdot \boldsymbol{p}_c(t_{l+1})|}{\|\boldsymbol{p}_c(t_l)\|^2 + \|\boldsymbol{p}_c(t_{l+1})\|^2} \tag{32}$$

where the scalar product of two probabilistic segmentations is given by

$$\boldsymbol{p}_c(t_1) \cdot \boldsymbol{p}_c(t_2) = \int_\Omega p_c(\boldsymbol{x}, t_1) p_c(\boldsymbol{x}, t_2)\, d\boldsymbol{x} \tag{33}$$

and the squared ℓ_2-norm is $\|\boldsymbol{p}_c(t)\|^2 = \boldsymbol{p}_c(t) \cdot \boldsymbol{p}_c(t)$.

3.1.7. Average of spatio-temporal measures

The voxel-wise measures are averaged for each discrete atlas time point t_l using the probabilistic segmentation of the atlas into nine Draw-EM classes. Given a quality measure $Q(\boldsymbol{x}, t)$, the average for $c \in \mathcal{L}_{\mathrm{T}}$ and $t = t_l$ is

$$\mu_l[Q](c) = \frac{\int_\Omega p_c(\boldsymbol{x}, t_l) Q(\boldsymbol{x}, t_l)\, d\boldsymbol{x}}{\int_\Omega p_c(\boldsymbol{x}, t_l)\, d\boldsymbol{x}} \tag{34}$$

3.2. Selection of model parameters

3.2.1. Image dissimilarity measure

For the construction of a neonatal brain atlas using our previous pairwise construction techniques (Serag et al., 2012a; Schuh et al., 2014) and the new groupwise approach presented in section 2.6, we first compared registration results between all pairs of images for three subsets of ten brain images each (representing different phases of brain development) for two different commonly used dissimilarity measures, namely NMI and LNCC. In the work of Gholipour et al. (2014) on constructing a fetal brain atlas using SyGN (Avants et al., 2010), the use of LNCC resulted in slightly sharper templates. In our case, for both the FFD method (Rueckert et al., 1999) used in the original approach of Serag et al. (2012a) and our proposed SVFFD algorithm, mean overlap of brain structures was found to be consistently higher for NMI when compared to the results obtained with LNCC. A registration using NMI is also more efficient (on average less than 40% of CPU time). Given the better performance and lower computational cost of NMI, we used it for all pairwise neonatal brain image registrations. Note that NMI has also been used for the construction of previous spatio-temporal neonatal atlases (Kuklisova-Murgasova et al., 2011; Serag et al., 2012a; Schuh et al., 2014).

3.2.2. Control point spacing and regularisation weights

With the image dissimilarity measure fixed, we performed an exhaustive grid search over the regularisation weights and different control point spacings using the subset of images used to select the dissimilarity measure. For the reference method of Serag et al. (2012a) based on the displacement-based FFD model, a control point spacing of 2.5 mm resulted in the best compromise between overlap and amount of folding. In case of the diffeomorphic SVFFD model, a smaller control point spacing of 2 mm was selected.

A comparison of the evaluation measures for both models with the selected parameters, and evaluated for the three subsets of in total 30 neonatal subjects scanned at different PMAs, is given in figure 5. This first comparison of separate atlases constructed without temporal kernel regression for three time points using our previously proposed pairwise construction techniques (Serag et al., 2012a; Schuh et al., 2014) suggests that both registration methods achieve similar segmentation overlap using the selected parameters, but that the diffeomorphic SVFFD registration combined with the Log-Euclidean mean produces better atlas templates, i.e., suggesting an advantage of our inverse-consistent and diffeomorphic atlas construction presented in Schuh et al. (2014) over the original approach proposed by Serag et al. (2012a).

Having determined the registration parameters for constructing a spatio-temporal neonatal atlas using direct registrations between all pairs of images, we performed a second parameter search for our new groupwise approach based on previously determined parameter ranges. This was done, because at each iteration the target image is the current average brain image, which has a different appearance than an actual brain scan. For this groupwise

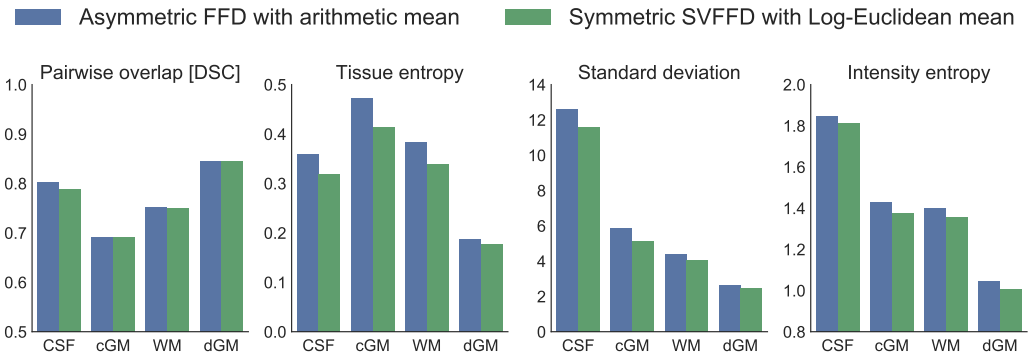


Figure 5: Comparison of evaluation measures for different pairwise atlas constructions based on a subset of in total 30 neonatal subjects used for parameter selection. This comparison suggests that, although overlap between image pairs is similar, the template created with the SVFFD model and Log-Euclidean mean is slightly sharper.

parameter search, we constructed a single time point atlas at 40 weeks PMA from a subset of 64 brain images of term-born neonates. The SVFFD parameters chosen for the pairwise atlas construction were found to be suitable also for the groupwise method. Due to the iterative refinement, the approach demonstrated to be more robust towards variations in parameters. However, the initial target images after affine alignment contain noticeable artefacts due to a significant amount of residual anatomical misalignment. We therefore constrain the deformations at early iterations more strongly using a sparser control point placement, and progressively half the control point spacing of the SVFFD after a few iterations. For the first two iterations, a control point spacing of 8 mm at the highest image resolution level is used. At iterations three to five, the control point spacing is 4 mm, whereas it is 2 mm for the final iterations. We found that the quality of the atlas did not substantially improve after 8 iterations, as can be seen in figure 6.

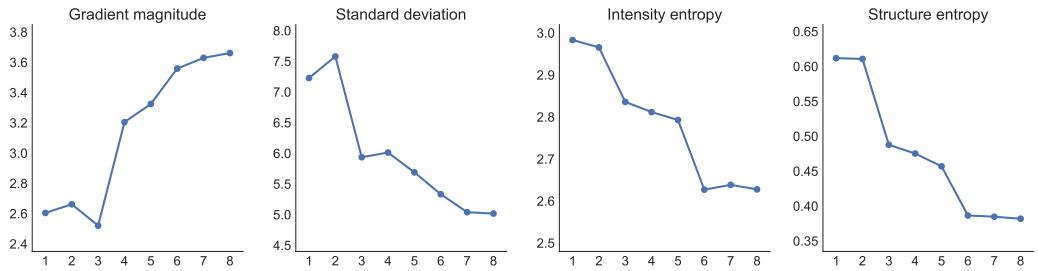


Figure 6: Convergence of atlas quality measures in case of groupwise construction. The horizontal axis corresponds to number of completed iterations, k. The jumps from iteration $k = 2$ to $k = 3$, and from $k = 5$ to $k = 6$ are the result of halving the minimum control point spacing of the SVFFD at the highest image resolution level between iterations.

3.2.3. Standard deviation of temporal regression kernel

For the construction of mean images at weekly intervals of the age range from 36 to 44 weeks PMA, for which a sufficient number of brain scans is available in our dataset, we use an algorithm similar to Serag et al. (2012a) to determine an age-specific standard deviation, σ_t, of the temporal regression kernel for each age $t_l \in \{36, 37, 38, 39, 40, 41, 42, 43, 44\}$. Our kernel width selection differs from the algorithm proposed by Serag et al. (2012a) in that no stopping criterion based on the difference in brain volumes is considered, because our goal is a uniform distribution of brain samples at the equally spaced discrete time points. We first specify a constant kernel width, σ, and determine the median number, m, of images with non-zero kernel weight for each discrete atlas age t_l. We then iteratively increase or decrease each kernel width σ_{t_l} by a small value until the number of images is equal to m.

Using this algorithm, we determined adaptive kernel widths for different target σ values, and constructed a separate spatio-temporal atlas for each corresponding set of kernel widths listed in table 2, and visualised in figure 7,

using our previous approach presented in Schuh et al. (2014). The TC of tissue classes for each constructed atlas is shown in figure 8. As expected, it is observed that TC improves with increasing kernel widths. When $\sigma \to \infty$, the atlas becomes constant in time, and thus has maximum TC, but fails to model temporal shape differences. Reducing σ results in more distinct and sharper mean brain shapes for each time point, with fewer brain images contributing to each template image. We therefore also compared the average pairwise overlap of spatially normalised Draw-EM structures and the mean voxel-wise standard deviation of T2w image intensity samples in figure 9. We expect an atlas with a consistent quality across time to achieve a similar DSC value for each time point, without sacrificing a substantial reduction of overall quality. We should, however, also consider the higher anatomical detail at later time points due to increased cortical folding.

A target σ value below half a week results in insufficient overlap to ensure temporal smoothness of the mean brain images at weekly intervals. A target kernel width of 0.75 to 1 week produces temporally smooth atlases with noticeable temporal shape differences. After visual inspection, we found that a good compromise for our dataset is obtained for target $\sigma = 1$. The kernel width σ_t for any time point is obtained by linear interpolation, i.e.,

$$\sigma_t = (1 - \alpha)\, \sigma_{\lfloor t \rfloor} + \alpha\, \sigma_{\lfloor t \rfloor + 1} \quad \text{where} \quad \alpha = \lfloor t \rfloor - t \tag{35}$$

3.3. Comparison of atlas construction methods

With the selected parameters, we constructed a spatio-temporal neonatal atlas from 36 to 44 weeks PMA using the discussed construction methods, and evaluated these at weekly intervals. We compared the atlases constructed

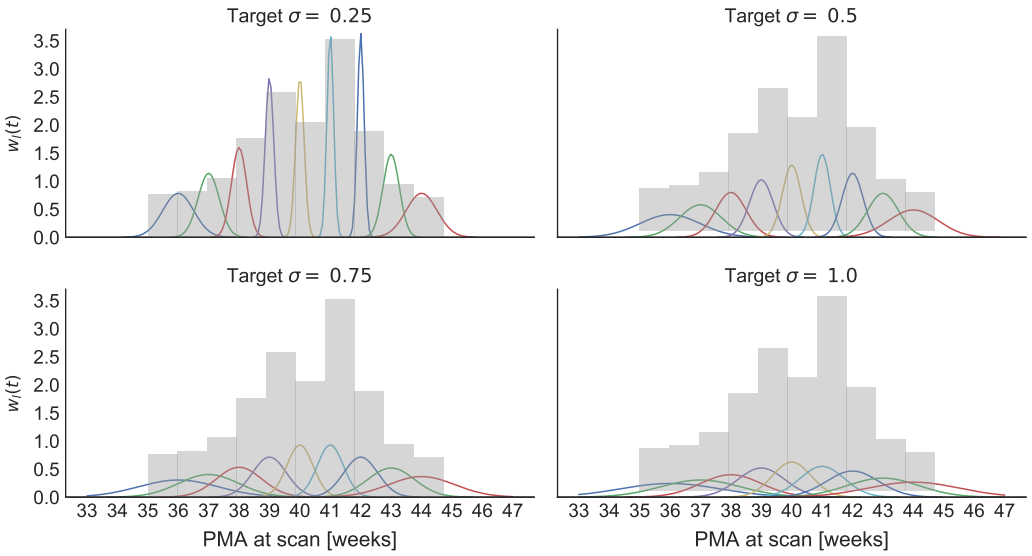


Figure 7: Temporal regression kernels for different ages and target σ values. The respective standard deviations of the shown regression kernels centred at the discrete time points t_l are listed in table 2. The curves are overlayed on the PMA at scan histogram.

$\boldsymbol{\sigma}$	0.25	0.50	0.75	1.00
$\boldsymbol{m}$	44	101	143	196
$\boldsymbol{\sigma_{36}}$	0.51	0.99	1.31	1.64
$\boldsymbol{\sigma_{37}}$	0.35	0.69	0.99	1.31
$\boldsymbol{\sigma_{38}}$	0.25	0.50	0.75	1.00
$\boldsymbol{\sigma_{39}}$	0.14	0.39	0.56	0.77
$\boldsymbol{\sigma_{40}}$	0.14	0.31	0.43	0.64
$\boldsymbol{\sigma_{41}}$	0.11	0.27	0.43	0.73
$\boldsymbol{\sigma_{42}}$	0.11	0.35	0.56	0.86
$\boldsymbol{\sigma_{43}}$	0.27	0.51	0.77	1.18
$\boldsymbol{\sigma_{44}}$	0.51	0.82	1.09	1.50

Table 2: Adaptive kernel widths.

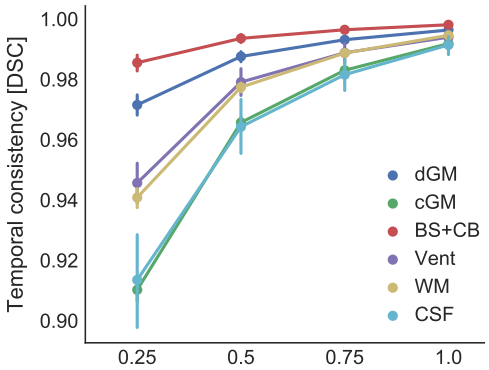


Figure 8: TC for different target σ values.

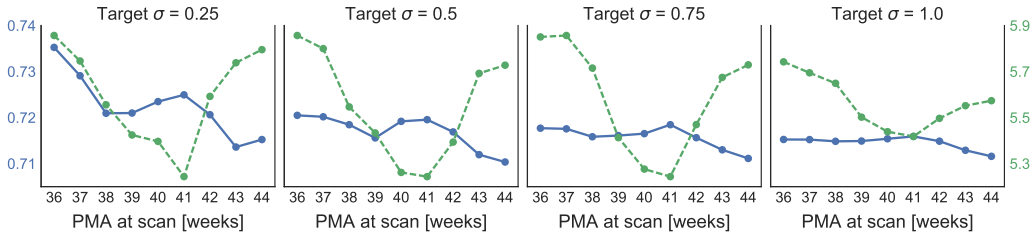


Figure 9: Quality measures for discrete time points of atlas constructed with approach proposed in Schuh et al. (2014) and different target σ values. The left vertical axis corresponds to the blue curve of the pairwise overlap (DSC), whereas the right axis corresponds to the dotted green curve depicting the mean voxel-wise standard deviation of T2w intensities.

with our proposed approaches based on the SVFFD algorithm and Log-Euclidean mean deformations to the original method of Serag et al. (2012a), which uses the classic FFD algorithm and arithmetic mean of displacements. Our previous method proposed in Schuh et al. (2014) is closely related to the approach of Serag et al. (2012a). Both derive the transformations relating each observation to the atlas space from transformations between each unique pair of brain images. This construction technique is thus referred to as "pairwise" method, while the new groupwise construction presented here is referred to as "groupwise" approach. A key difference between our "pairwise" approach and the original "reference" method of Serag et al. (2012a) is, that we utilise topology preserving deformation models, and the Log-Euclidean mean instead of the arithmetic mean of the pairwise mappings.

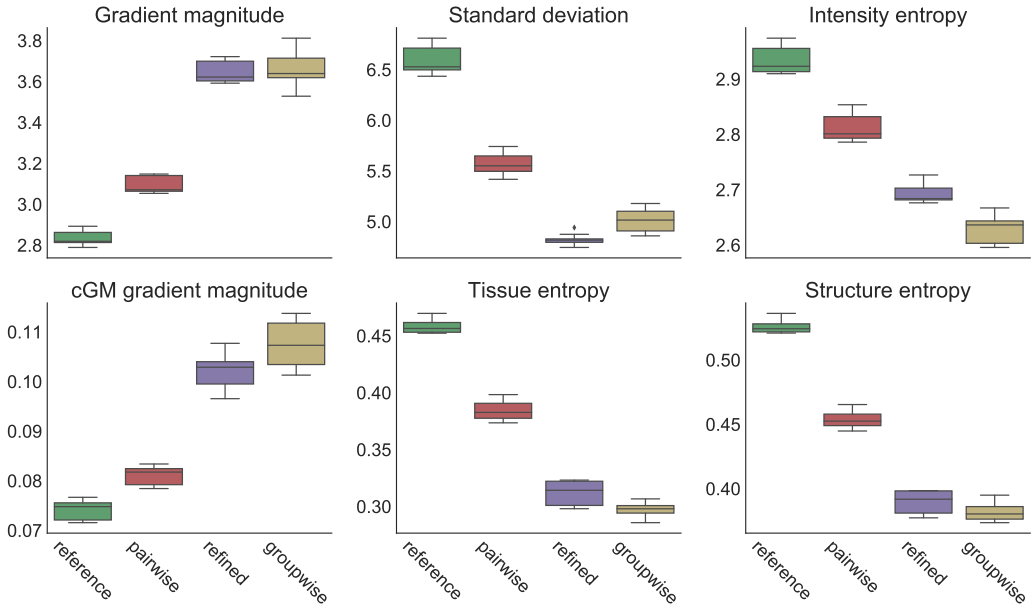
To reduce the residual anatomical misalignment in the atlas space further, we additionally performed a separate iterative refinement of the template brain images at discrete time points constructed with the "pairwise" approach. For this, the mean diffeomorphisms relating each subject to the

Method	Summary
reference	Approach of Serag et al. (2012a) using arithmetic mean of FFDs.
pairwise	Log-Euclidean mean of pairwise SVFFDs (Schuh et al., 2014).
refined	The "pairwise" atlas with 10 iterations of deformation refinement.
groupwise	Proposed temporally consistent groupwise atlas construction.

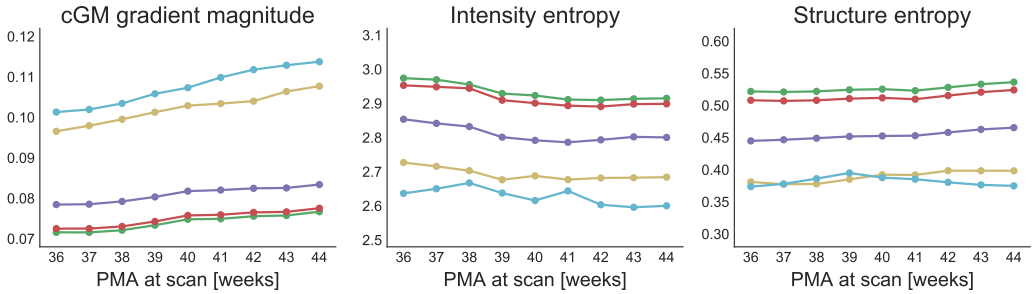
Table 3: Summary of compared spatio-temporal atlas construction methods.

atlas space at a specific time point are refined by registering each brain image to the respective mean intensity image. The new subject-to-atlas mapping is then given by the composition of this new transformation with the Log-Euclidean mean, i.e., similar to equation (19). Given these refined subject-to-atlas transformations, a new template image is computed, and the process is repeated. This iterative approach is comparable to the spatio-temporal fetal atlas construction of Gholipour et al. (2017), but using our SVFFD algorithm and the Log-Euclidean mean instead of SyGN. We compare this "refined" atlas to our proposed "groupwise" approach. A summary of the compared construction methods is given in table 3.

The mean of the voxel-wise measures averaged separately for each time point for each method over the entire brain masks is reported in table 4. Box plots of these measures visualising also the temporal variance of the average measures are shown in figure 10a, whereas plots of individual average measures over time are shown in figure 10b. The latter visualisation highlights the temporal smoothness even for average measures with a larger interquartile range, and explains this larger variance of average measures by the difference between early and late atlas time points.



(a) Box plots of mean values for different time points.



(b) Mean values over PMA at scan.

Figure 10: Mean quality measures of constructed spatio-temporal atlases.

	reference	pairwise	refined	groupwise
Gradient magnitude	2.837	3.096	3.648	**3.661**
Standard deviation	6.591	5.563	**4.821**	5.017
Intensity entropy	2.933	2.811	2.693	**2.628**
Tissue entropy	0.459	0.384	0.311	**0.297**
Structure entropy	0.526	0.454	0.389	**0.382**
cGM gradient magnitude	0.074	0.081	0.102	**0.108**

Table 4: Mean quality measures of constructed spatio-temporal atlases. For each evaluation criterion, the result for the atlas which is most favourable is highlighted in bold.

As expected, a number of the evaluation measures are strongly correlated. This is especially the case for the voxel-wise label entropy of tissue classes and individual brain structures, because both are based on the automatic Draw-EM segmentations. We therefore focus on the structure entropy as it is based on a more fine granular parcellation. In regards to the intensity based evaluation criteria, the measures of intensity variance, namely standard deviation and entropy, demonstrate a similar relative performance of the construction methods. A comparison of these voxel-wise measures averaged separately for each tissue class in figure 11 shows that the apparent worse performance of the "groupwise" method when compared to the iteratively "refined" pairwise atlas is caused by a higher standard deviation of intensities for cerobrospinal fluid (CSF). This is explained by the fact that this tissue class is at the boundary of the brain mask at which the NMI similarity criterion may not be evaluated during the registration as a result of only partial overlap of the foreground regions. An improved handling of partial boundary overlap during the registration would likely improve the within CSF measures. Of higher interest is however the atlas quality for brain tissues. The average

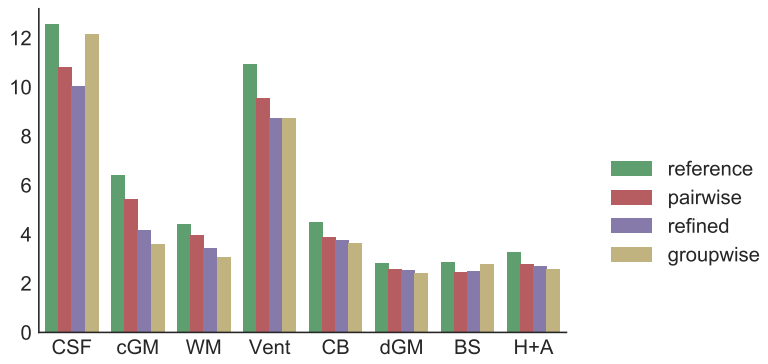


Figure 11: Intensity standard deviation averaged separately for CSF, cGM, WM, ventricles (Vent), cerebellum (CB), dGM, brainstem (BS), hippocampi and amygdalae (H+A).

measures for cGM, WM, and deep grey matter (dGM) for each constructed spatio-temporal atlas are compared in figure 12. It can be observed that all methods perform similarly well for deep brain structures, as these require less local deformation to be aligned, but that our proposed methods improve the quality of the atlas noticeably within WM and especially cGM regions. Our pairwise atlas construction using the Log-Euclidean mean of inverse-consistent diffeomorphic mappings is more consistent near the cortex than the approach proposed by Serag et al. (2012a). This is in part because we are able to more locally align cortical features with a lower control point spacing of the SVFFD, which does not introduce folding. In contrast, the FFD model requires stronger regularisation in order to preserve the topology of the anatomies. Moreover, the iterative approaches clearly outperform the pairwise methods due to the iterative compensation of residual anatomical misalignment. Our proposed "groupwise" spatio-temporal atlas construction performs slightly better than the individual iterative refinement of atlas templates at discrete time points.

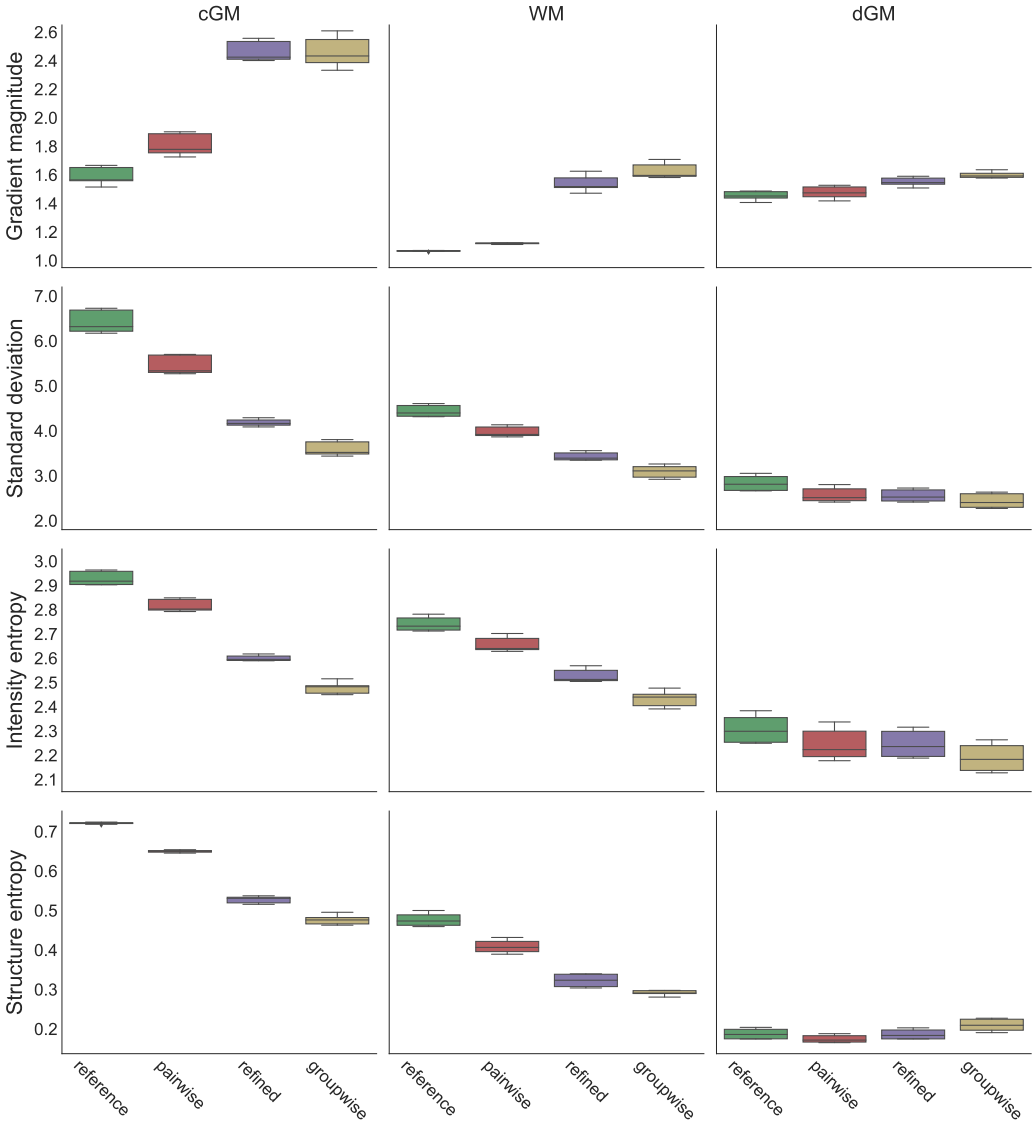


Figure 12: Atlas quality measures averaged separately for cGM, WM, and dGM.

3.4. Visual assessment of neonatal atlases

Having observed improvements in the defined quantitative measures, we further assess the quality of the constructed atlases visually. Templates for 40 weeks PMA constructed with the reference method of Serag et al. (2012a), our derived pairwise approach using the SVFFD model and Log-Euclidean mean, the template of the pairwise method after 10 iterations of individual subject-to-atlas deformation refinement, and the template generated by the proposed groupwise approach after 8 iterations are compared in figure 13a. Most noticeable are the sharper cortical details achieved by the iterative approaches. Differences in the atlases obtained with either of the two pairwise or iterative techniques, respectively, are more subtle, and less clear by inspecting a single cross-section. Slightly clearer cortical boundaries may be observed in the atlas created with our pairwise method when compared to the reference approach, in particular at the frontal lobe. These differences are more apparent when examining coloured renders of the voxel-wise intensity and tissue class entropy measures in figure 13b. Our pairwise method achieves more anatomical consistency in the alignment of the individual brain scans in the atlas space, but both methods produce atlases of high quality. An advantage of our approach is a consistent modelling of mean shape by topology-preserving deformations.

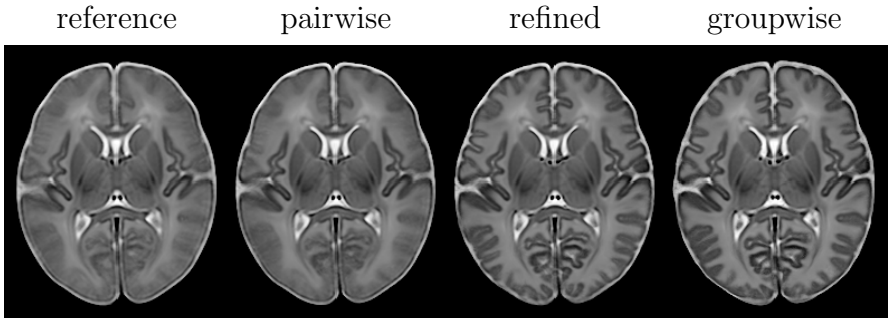
The improved temporal consistency of the atlas constructed with our groupwise approach, as opposed to an individual iterative refinement of each time point, is highlighted in the detail views of smaller ROIs in figure 14. When brain scans are registered individually to each template, developing cortical structures may be missing at certain time points or some folding

patterns recede instead of deepen further, which is not expected for normal development. This is not observed in the temporal sequences generated by our groupwise method, which consistently maps each brain scan to each individual time point considering the longitudinal inter-atlas deformations.

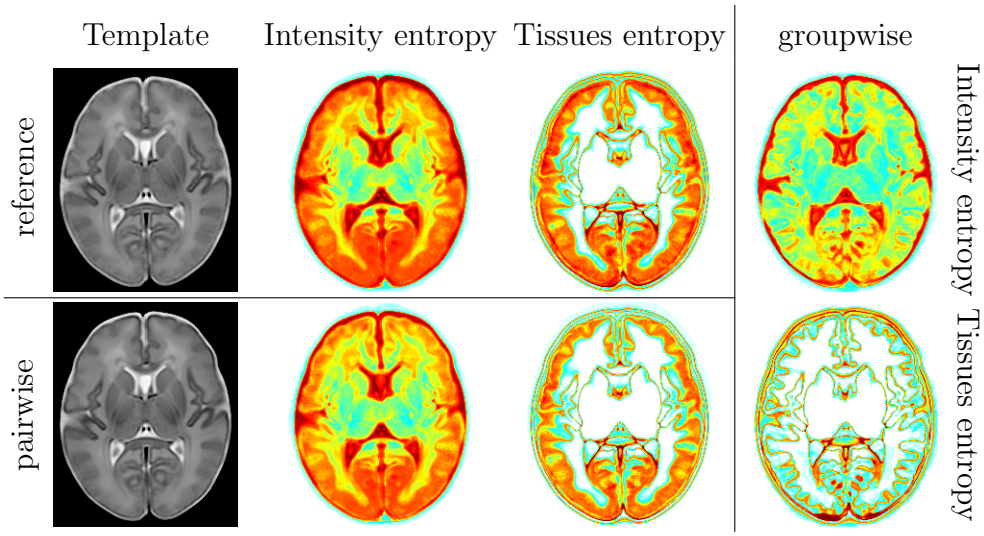
T1w and T2w mean shape and intensity images at weekly intervals for the age range from 36 to 44 weeks PMA generated with the proposed groupwise technique are displayed in figure 15, along with the spatial tissue probability map for cGM. The clear markedness of cortical folds is reflected in the sharpness of the spatial cGM probability distribution.

3.5. Qualitative comparison to previous atlases

Thus far, we compared atlases constructed with our methods based on the SVFFD algorithm to the atlas built according to the approach of Serag et al. (2012a) from the same set of images, and an improved registration using the Medical Image Registration ToolKit (MIRTK). In figure 16, we give a comparison to the spatio-temporal neonatal brain atlases pre-dating our work, that were made publicly available by the original authors of these methods. The significant improvement in image quality of the new atlases is clearly noticeable. One reason for this is that the brain images acquired by dHCP are of remarkably higher quality than previous datasets of preterm-born neonates that were available for the construction of these previous atlases. Another reason is that at each time point of our new atlases, 196 anatomies are being averaged, whereas the mean images created by Serag et al. (2012a) are the average of 15-19 individual brain scans. This explains the lower cortical detail of our pairwise atlas. High quality and marked cortical folds are present in the atlas generated by our groupwise method.

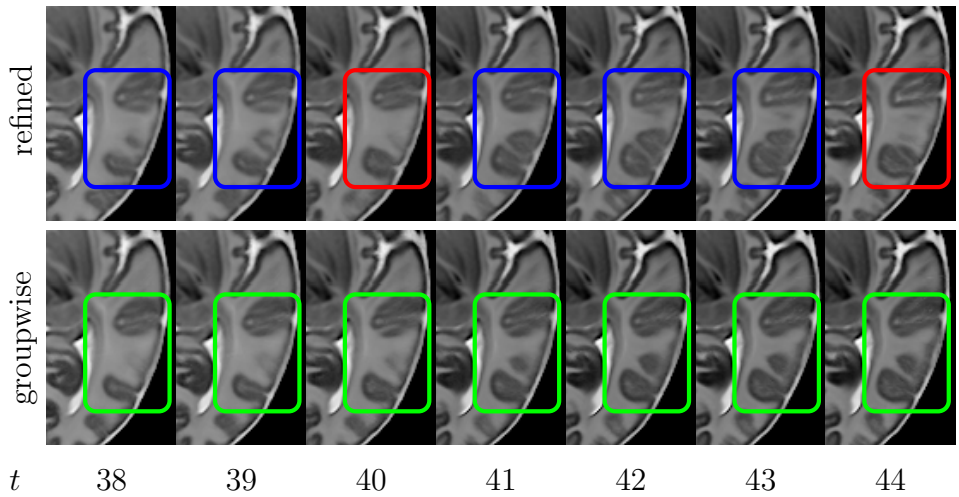


(a) Axial mid-sections of mean intensity and shape templates at 40 weeks PMA.

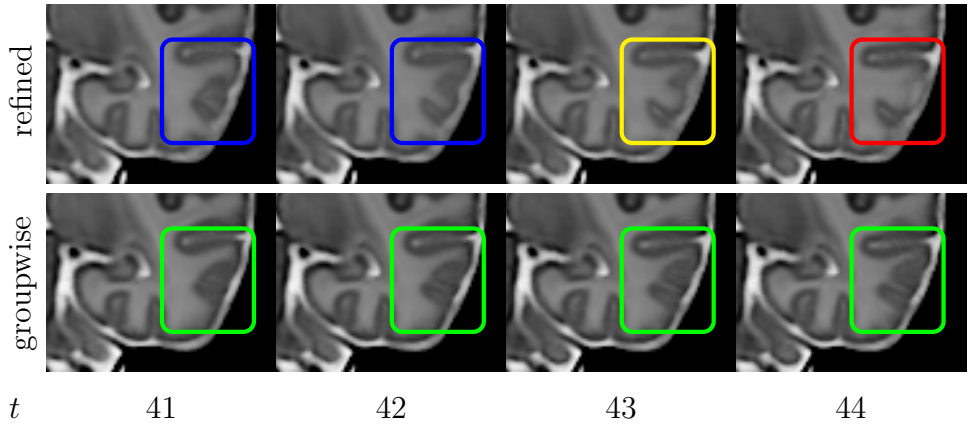


(b) Template at 36 weeks constructed with the pairwise approach (top left) using the arithmetic mean of FFDs, and (bottom left) our method using the SVFFD algorithm and Log-Euclidean mean. Subtle differences in mean intensity at the frontal cortex can be observed; (middle left) colour mappings of voxel-wise intensity and (middle right) tissue entropy reveal further cortical differences, where white corresponds to zero entropy, and red indicates higher entropy. The entropy measurements for the new groupwise approach are presented on the right.

Figure 13: Visual comparison of templates constructed with different approaches.



(a) Cortical structures develop in the left temporal lobe, but vanish at weeks 40 and 44 for the individual refinement of time points (top, red boxes). The sequence of our groupwise method shows a consistent formation of a sulcus.



(b) In the sequence obtained by individual refinement of each time point, a forming sulcus vanishes between week 43 (yellow box) and 44 (red box). Our groupwise approach generates a temporal sequence consistent with increasing cortical folding.

Figure 14: Detail views of axial cross-sections of atlases constructed (top) with the pairwise approach followed by iterative refinement, and (bottom) our groupwise method.

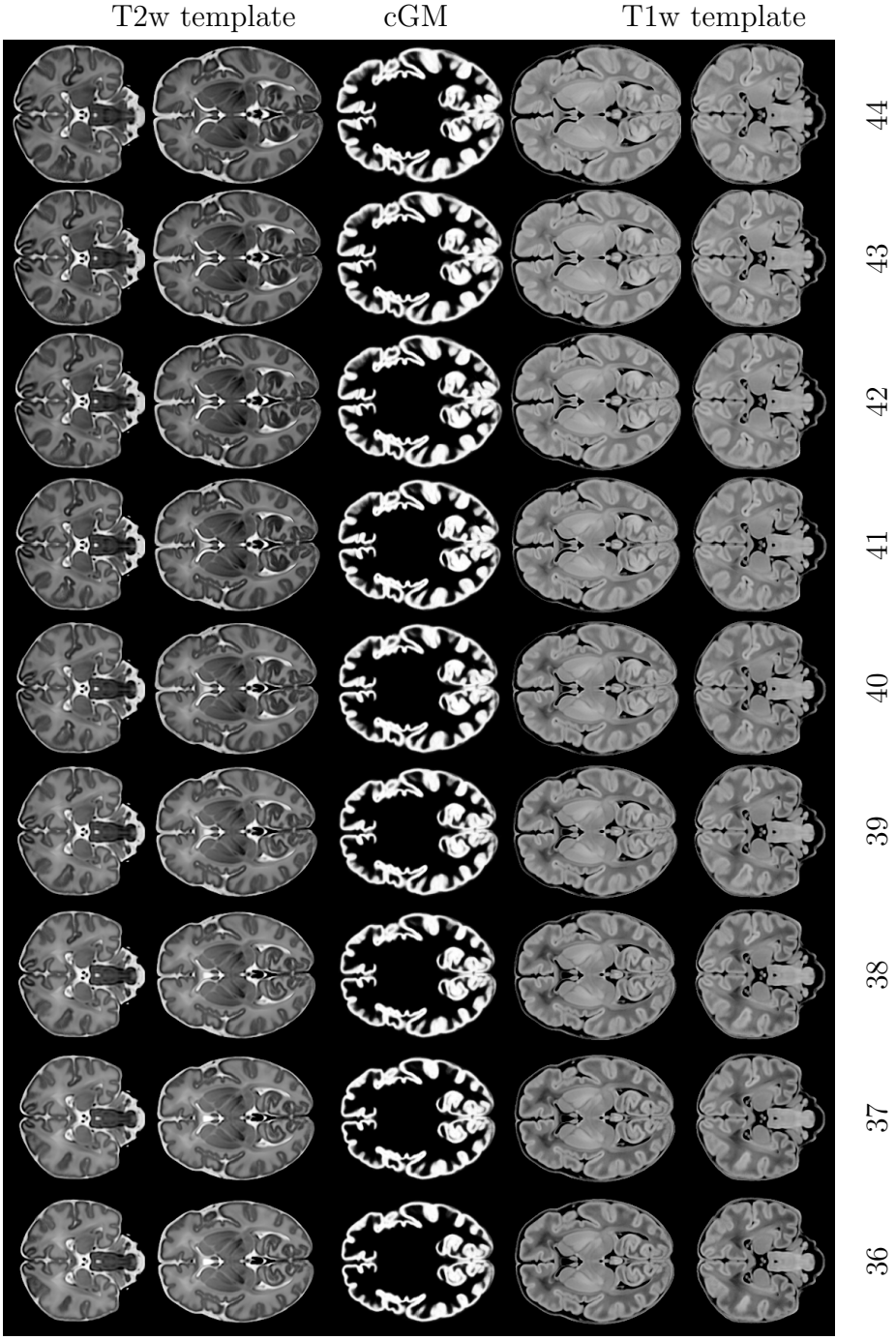
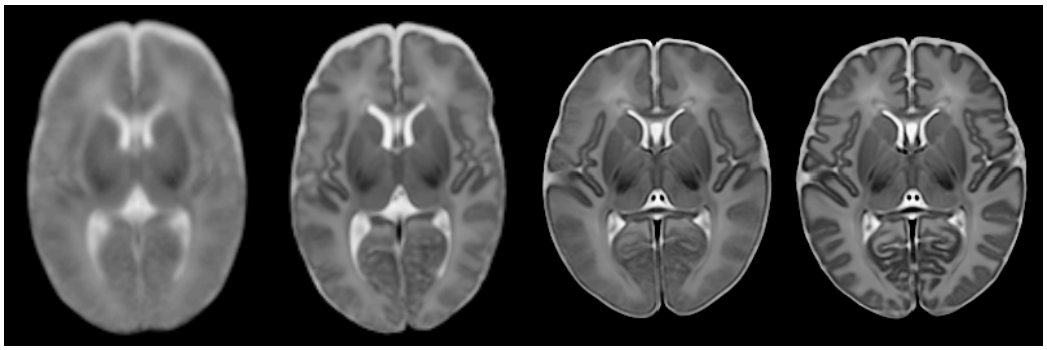


Figure 15: Spatio-temporal atlas constructed with the proposed groupwise approach.



(a) (b) (c) (d)

Figure 16: Qualitative comparison of mean T2w image at 40 weeks of (a) atlas created by Kuklisova-Murgasova et al. (2011) from 142 preterm-born neonates using Log-Euclidean mean of affine transformations with constant kernel width of 2 weeks, (b) atlas created by Serag et al. (2012a) from 204 preterm-born neonates using arithmetic mean of pairwise FFDs with adaptive kernel widths, (c) our pairwise atlas constructed from 275 term-born neonates using the Log-Euclidean mean of SVFFDs also with adaptive kernel widths, and (d) the atlas constructed from the same brain images using our groupwise approach.

4. Discussion

In this work, we have proposed a new iterative method for the construction of a spatio-temporal atlas of neonatal brain morphology with markedly increased anatomical detail when compared to neonatal brain atlases that were made publicly available before[2]. This new approach is compared to our previously proposed techniques based on mean deformations derived from weighted spatial mappings between all unique pairs of acquired brain images (Serag et al., 2012a; Schuh et al., 2014). Moreover, similar to the fetal atlas construction of Gholipour et al. (2017), which extended the single-time

[2] http://brain-development.org/brain-atlases/

point SyGN approach to a spatio-temporal model using the adaptive temporal kernel regression used by Serag et al. (2012a), we show that an individual refinement of discrete time points, by means of registering the brain images to each respective template brain image separately, may introduce implausible temporal inconsistencies. The atlas constructed with our groupwise technique does not suffer from these temporal inconsistencies.

For the comparison of the iterative approaches, we used our SVFFD algorithm instead of the greedy SyN method utilised by Gholipour et al. (2017) for several reasons. First, our main goal was to demonstrate the temporal inconsistencies that may be introduced by independent non-linear registrations to different time points due to the highly non-convex nature of the energy functions. Secondly, although SyN (Avants et al., 2008) has been demonstrated to be among the best performing deformable registration methods (Klein et al., 2009; Ou et al., 2014), the greedy optimisation can only provide us with the final displacement field and its inverse, instead of a (non-stationary) velocity field, unless the intermediate instantaneous velocity fields were to be recorded. A mean diffeomorphism corresponding to the average velocity fields can thus not be computed. The properties of a velocity-based parametrisation are hereby traded in for a more efficient registration, which does not require the costly space-time optimisation of LDDMM (Beg et al., 2005). A further drawback of a greedy optimisation, even when the intermediate update fields are preserved, is that it may lead to a suboptimal flow of diffeomorphisms in terms of the deviation from the shortest path connecting the identity mapping from the final deformation along the manifold of diffeomorphisms. When defining a mean deformation based on these velocity

fields, a deformation corresponding to a longer path contributes more to the mean shape deformation than a flow that follows the geodesic. Finally, although SyN is a greedy method, it has an average runtime of several hours[3]. For comparison, our SVFFD algorithm has an average multi-threaded runtime of 8–15 min when executed on a machine with 8 CPU cores, and the result is a SVF based on which a mean diffeomorphism can be computed.

It is worthwhile to note that in order to increase the sharpness of the constructed mean images, the Laplacian sharpening implemented in the Insight Segmentation and Registration Toolkit (ITK)[4] (Yoo, 2004) is applied. We noticed that this filter is used by the SyGN implementation in ANTs, on which the results presented by Gholipour et al. (2017) are based, and therefore adopted this post-processing step. Note also that the same edge sharpening has been done for all atlases compared in this work, except for the publicly available atlas templates shown in figures 16a and 16b.

5. Conclusions

We have proposed two approaches for the construction of a spatio-temporal atlas of the developing human brain based on the unbiased SVFFD algorithm. Our first approach, presented in Schuh et al. (2014), improved upon an earlier method developed by Serag et al. (2012a). Unlike the initial technique using the FFD algorithm (Rueckert et al., 1999) implemented in IRTK and arithmetic mean of displacements (Seghers et al., 2004), we utilised the SVFFD algorithm and Log-Euclidean mean of SVFs (Arsigny et al., 2006).

[3] An average runtime of ten hours is reported in Tustison and Avants (2013).
[4] https://itk.org/

This preserves the topology of the brain anatomies and ensures inverse consistency of the computed spatial mappings. Using MIRTK for both pairwise methods, we demonstrated slight improvements in cortical entropy measures using the SVFFD algorithm. A disadvantage of these approaches is, however, the quadratic computational complexity with respect to the number of brain images. A conceptual limitation is the reliance on accurate pairwise correspondences. We thus proposed a new groupwise method, which demonstrates better temporal consistency when compared to other spatio-temporal extensions of iterative methods (Gholipour et al., 2017), and results in average brain shape and intensity images with significantly more cortical details.

We conclude with a remark by Sled and Nossin-Manor (2013), that an inherent difficulty in brain atlas formation is the normal variability among individuals, and the uncertainty of the PMA. In time, these are mitigated by larger datasets acquired shortly after birth. The acquisition and public dissemination of more than a thousand high-quality brain MR images of the developing brain from 20 to 44 weeks PMA is the goal of the dHCP. The method presented in this work facilitates the construction of a spatio-temporal atlas of brain morphology from this and other large datasets.

An open source implementation of the SVFFD algorithm and our atlas construction methods are publicly available as part of MIRTK[5]. The atlases constructed from the dHCP cohort will be publicly available for download from `https://data.developingconnectome.org`.

[5] `https://biomedia.doc.ic.ac.uk/software/mirtk/`

Acknowledgements

This work has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) [grant number 319456]. We are grateful to the families who generously supported this trial. The work was supported by the NIHR Biomedical Research Centers at Guy's and St Thomas' NHS Trust.

Appendix A. Cubic B-spline parametrisation of SVF

The SVF, whose parameters are optimised by the SVFFD algorithm, is

$$\boldsymbol{v}(\boldsymbol{x}) = \sum_{l,m,n=0}^{3} w_l(\tilde{x}_1) w_m(\tilde{x}_2) w_n(\tilde{x}_3) \boldsymbol{c}_{\{(i_1+l)+N_1((i_2+m)+N_2\,(i_3+n)\}} \tag{A.1}$$

with coefficient vectors $\boldsymbol{c}_i \in \mathbb{R}^3$. The local coordinates of $\boldsymbol{x} \in \Omega \subset \mathbb{R}^3$ with respect to the control point lattice of size $N_1 N_2 N_3$ are given by a rotation $\boldsymbol{R} \in \mathbb{R}^{3\times 3}$, translation $\boldsymbol{d} \in \mathbb{R}^3$, and anisotropic scaling by factors corresponding to the uniform spacing δ_d of control points along the d-th axis of the lattice, i.e., $\hat{\boldsymbol{x}} = \mathrm{diag}(\delta_1, \delta_2, \delta_3)^{-1} \boldsymbol{R}\boldsymbol{x} + \boldsymbol{d}$. The coordinates of $\boldsymbol{x}$ relative to the edges of the cell of the lattice that it is contained in are then given by $\tilde{\boldsymbol{x}} = \hat{\boldsymbol{x}} - \lfloor \hat{\boldsymbol{x}} \rfloor$, and $\boldsymbol{i} = \lfloor \hat{\boldsymbol{x}} \rfloor - 1$ is the multi-index of the "leftmost" control point with non-zero weight. The weights $w_l(\tilde{x}_i)$ are defined in terms of the polynomials B_i of the cubic B-spline function with knots at -2, -1, 0, 1, and 2. These are derived using Cox-de Boor's recursion formula (de Boor, 2001), i.e.,

$$\begin{aligned} w_0(\tilde{x}_i) &= B_3(\tilde{x}_i + 1) & w_1(\tilde{x}_i) &= B_2(\tilde{x}_i + 0) \\ w_2(\tilde{x}_i) &= B_1(\tilde{x}_i - 1) & w_3(\tilde{x}_i) &= B_0(\tilde{x}_i - 2) \end{aligned} \tag{A.2}$$

where

$$
\begin{aligned}
B_0(r) &= (r+2)^3 \,/\, 6 \\
B_1(r) &= \left(-3r^3 - 6r^2 + 4\right) / \, 6 \\
B_2(r) &= \left(3r^3 - 6r^2 + 4\right) / \, 6 \\
B_3(r) &= (2-r)^3 \,/\, 6
\end{aligned}
\tag{A.3}
$$

Appendix B. Iterative pairwise affine registration

The affine transformation $\boldsymbol{A}_{ij}^{(k)}$ is obtained by registering the j-th to the i-th image after application of the previous mean transformations, i.e., by registering the images $I_i \circ \left(\boldsymbol{A}_i^{(k-1)}\right)^{-1}$ and $I_j \circ \left(\boldsymbol{A}_j^{(k-1)}\right)^{-1}$. The blurring effect caused by interpolation, when resampling the images in the reference space, is avoided through composition with the transformations $\boldsymbol{B}_i \in \mathbb{R}^{4\times 4}$, which map each image sampling space to the common world space. Specifically, the map relating a sampling point $\hat{\boldsymbol{x}}_i \in \mathbb{R}^3$ in the space of the i-th image, with axes parallel to the sampling grid and scaled according to the image resolution, to its corresponding point $\hat{\boldsymbol{x}}_j$ in the j-th image is given by

$$
\begin{aligned}
\begin{pmatrix} \hat{\boldsymbol{x}}_j^T & 1 \end{pmatrix}^T &= \boldsymbol{B}_j^{-1} \left(\boldsymbol{A}_j^{(k-1)}\right)^{-1} \boldsymbol{A}_{ij}^{(k)} \boldsymbol{A}_i^{(k-1)} \boldsymbol{B}_i \begin{pmatrix} \hat{\boldsymbol{x}}_i^T & 1 \end{pmatrix}^T \\
&= \left(\boldsymbol{A}_j^{(k-1)} \boldsymbol{B}_j\right)^{-1} \boldsymbol{A}_{ij}^{(k)} \left(\boldsymbol{A}_i^{(k-1)} \boldsymbol{B}_i\right) \begin{pmatrix} \hat{\boldsymbol{x}}_i^T & 1 \end{pmatrix}^T \\
&= \left(\boldsymbol{C}_j^{(k-1)}\right)^{-1} \boldsymbol{A}_{ij}^{(k)} \boldsymbol{C}_i^{(k-1)} \begin{pmatrix} \hat{\boldsymbol{x}}_i^T & 1 \end{pmatrix}^T
\end{aligned}
\tag{B.1}
$$

Registering images with pre-transformations $\boldsymbol{A}_{\{i,j\}}^{(k-1)}$ is thus equivalent to registering images with the composite image to world maps denoted by $\boldsymbol{C}_{\{i,j\}}^{(k-1)}$. The registrations at the k-th iteration find the affine mappings $\boldsymbol{A}_{ij}^{(k)}$ that maximise the NMI of each image pair.

References

Alexa, M., 2002. Linear combination of transformations. In: Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques. San Antonio, Texas, USA, pp. 380–7.

Alexander, B., Murray, A. L., Loh, W. Y., Matthews, L. G., Adamson, C., Beare, R., Chen, J., Kelly, C. E., Rees, S., Warfield, S. K., Anderson, P. J., Doyle, L. W., Spittle, A. J., Cheong, J. L., Seal, M. L., Thompson, D. K., 2016. A new neonatal cortical and subcortical brain atlas: the Melbourne Children's Regional Infant Brain (M-CRIB) atlas. NeuroImage 147 (Supplement C), 841–51.

Arsigny, V., Commowick, O., Pennec, X., Ayache, N., 2006. A Log-Euclidean framework for statistics on diffeomorphisms. In: Larsen, R., Nielsen, M., Sporring, J. (Eds.), Proceedings of 9th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). Springer Berlin Heidelberg, pp. 924–31.

Arsigny, V., Fillard, P., Pennec, X., Ayache, N., 2007. Geometric means in a novel vector space structure on symmetric positivedefinite matrices. SIAM Journal on Matrix Analysis and Applications 29 (1), 328–47.

Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. NeuroImage 38 (1), 95–113.

Ashburner, J., Friston, K. J., 1999. Nonlinear spatial normalization using basis functions. Human Brain Mapping 7 (4), 254–66.

Ashburner, J., Friston, K. J., 2000. Voxel-based morphometry–the methods. NeuroImage 11 (6), 805–21.

Ashburner, J., Hutton, C., Frackowiak, R., Johnsrude, I., Price, C., Friston, K., 1998. Identifying global anatomical differences: Deformation-based morphometry. Human Brain Mapping 6 (5-6), 348–57.

Avants, B. B., Epstein, C. L., Grossman, M., Gee, J. C., 2008. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. Medical Image Analysis 12 (1), 26–41.

Avants, B. B., Yushkevich, P., Pluta, J., Minkoff, D., Korczykowski, M., Detre, J., Gee, J. C., 2010. The optimal template effect in hippocampus studies of diseased populations. NeuroImage 49 (3), 2457–66.

Beg, M. F., Khan, A., 2007. Symmetric data attachment terms for large deformation image registration. IEEE Transactions on Medical Imaging 26 (9), 1179–89.

Beg, M. F., Miller, M. I., Trouvé, A., Younes, L., 2005. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. International Journal of Computer Vision 61 (2), 139–57.

Benkarim, O. M., Sanroma, G., Zimmer, V. A., Muñoz-Moreno, E., Hahner, N., Eixarch, E., Camara, O., González Ballester, M. A., Piella, G., 2017. Toward the automatic quantification of in utero brain development in 3D structural MRI: A review. Human Brain Mapping 35 (4), 2772–87.

Bhatia, K., Hajnal, J., Puri, B., Edwards, A., Rueckert, D., 2004. Consistent groupwise non-rigid registration for atlas construction. In: Proceedings of IEEE International Symposium on Biomedical Imaging (ISBI). Vol. 1. pp. 908–11.

Bhatia, K. K., Aljabar, P., Boardman, J. P., Srinivasan, L., Murgasova, M., Counsell, S. J., Rutherford, M. a., Hajnal, J., Edwards, a. D., Rueckert, D., 2007. Groupwise combined segmentation and registration for atlas construction. In: Ayache, N., Ourselin, S., Maeder, A. (Eds.), Proceedings of 10th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). Springer Berlin Heidelberg, pp. 532–40.

Blencowe, H., Cousens, S., Oestergaard, M. Z., Chou, D., Moller, A. B., Narwal, R., Adler, A., Vera Garcia, C., Rohde, S., Say, L., Lawn, J. E., 2012. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: A systematic analysis and implications. The Lancet 379 (9832), 2162–2172.

Bossa, M., Hernandez, M., Olmos, S., 2007. Contributions to 3D diffeomorphic atlas estimation: application to brain images. In: Ayache, N., Ourselin, S., Maeder, A. (Eds.), Proceedings of 10th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). pp. 667–74.

Bossa, M., Zacur, E., Olmos, S., 2008. Algorithms for computing the group exponential of diffeomorphisms: Performance evaluation. In: Proceedings

of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. pp. 1–8.

Cachier, P., Rey, D., 2000. Symmetrization of the non-rigid registration problem using inversion-invariant energies: Application to multiple sclerosis. In: Delp, S. L., DiGoia, A. M., Jaramaz, B. (Eds.), Proceedings of 3rd International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). Springer Berlin Heidelberg, pp. 472–81.

Ceritoglu, C., Oishi, K., Li, X., Chou, M. C., Younes, L., Albert, M., Lyketsos, C., van Zijl, P. C. M., Miller, M. I., Mori, S., 2009. Multi-contrast large deformation diffeomorphic metric mapping for diffusion tensor imaging. NeuroImage 47 (2), 618–27.

Choi, Y., Lee, S., 2000. Injectivity conditions of 2D and 3D uniform cubic B-spline functions. Graphical Models 62 (6), 411–27.

Christensen, G. E., Rabbitt, R. D., Miller, M. I., 1996. Deformable templates using large deformation kinematics. IEEE Transactions on Image Processing 5 (10), 1435–47.

Cordero-Grande, L., Hughes, E. J., Hutter, J., Price, A. N., Hajnal, J. V., 2017. Three-dimensional motion corrected sensitivity encoding reconstruction for multi-shot multi-slice MRI: Application to neonatal brain imaging. Magnetic Resonance in Medicine 00.

Davis, B. C., Fletcher, P. T., Bullitt, E., Joshi, S., 2007. Population shape regression from random design data. In: Proceedings of 11th IEEE International Conference on Computer Vision (ICCV). pp. 1–7.

de Boor, C., 2001. A Practical Guide to Splines - Revised Edition. Springer-Verlag, New York.

De Craene, M., Piella, G., Camara, O., Duchateau, N., Silva, E., Doltra, A., D'hooge, J., Brugada, J., Sitges, M., Frangi, A. F., 2012. Temporal diffeomorphic free-form deformation: application to motion and strain estimation from 3d echocardiography. Medical Image Analysis 16 (2), 427–50.

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., Killiany, R. J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 31 (3), 968–80.

Dice, L. R., 1945. Measures of the amount of ecologic association between species. Ecology 26 (3), 297–302.

Ghadimi, S., Mohtasebi, M., Abrishami Moghaddam, H., Grebe, R., Gity, M., Wallois, F., 2017. A neonatal bimodal MR-CT head template. PLoS One 12 (1), e0166112.

Gholipour, A., Limperopoulos, C., Clancy, S., Clouchoux, C., Akhondi-asl, A., Estroff, J. A., Warfield, S. K., 2014. Construction of a deformable spatiotemporal MRI atlas of the fetal brain: Evaluation of similarity metrics and deformation models. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (Eds.), Proceedings of 17th International Conference

on Medical Image Computing and Computer-Assisted Intervention (MICCAI). Springer International Publishing, pp. 292–9.

Gholipour, A., Rollins, C. K., Velasco-Annis, C., Ouaalam, A., Akhondi-Asl, A., Afacan, O., Ortinau, C. M., Clancy, S., Limperopoulos, C., Yang, E., Estroff, J. A., Warfield, S. K., 2017. A normative spatiotemporal MRI atlas of the fetal brain for automatic segmentation and analysis of early brain growth. Scientific Reports 7 (1), 476.

Gousias, I. S., Edwards, a. D., Rutherford, M. a., Counsell, S. J., Hajnal, J. V., Rueckert, D., Hammers, A., 2012. Magnetic resonance imaging of the newborn brain: Manual segmentation of labelled atlases in term-born and preterm infants. NeuroImage 62 (3), 1499–509.

Grenander, U., Miller, M. I., 1998. Computational anatomy: An emerging discipline. Quarterly of Applied Mathematics 56 (4), 617–94.

Guimond, A., Meunier, J., Thirion, J., 1998. Automatic computation of average brain models. In: Wells, W. M., Colchester, A., Delp, S. (Eds.), Proceedings of 1st International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). pp. 631–40.

Hughes, E. J., Winchman, T., Padormo, F., Teixeira, R., Wurie, J., Sharma, M., Fox, M., Hutter, J., Cordero-Grande, L., Price, A. N., Allsop, J., Bueno-Conde, J., Tusor, N., Arichi, T., Edwards, A. D., Rutherford, M. A., Counsell, S. J., Hajnal, J. V., 2017. A dedicated neonatal brain imaging system. Magnetic Resonance in Medicine 78 (2), 794–804.

Hüppi, P. S., Schuknecht, B., Boesch, C., Bossi, E., Felblinger, J., Fusch, C., Herschkowitz, N., 1996. Structural and neurobehavioral delay in postnatal brain development of preterm infants. Pediatric Research 39 (5), 895–901.

Inder, T. E., Warfield, S. K., Wang, H., Hüppi, P. S., Volpe, J. J., feb 2005. Abnormal cerebral structure is present at term in premature infants. Pediatrics 115 (2), 286 LP – 294.

Joshi, S., Davis, B., Jomier, M., Gerig, G., 2004. Unbiased diffeomorphic atlas construction for computational anatomy. NeuroImage 23 Suppl 1, S151–60.

Kapellou, O., Counsell, S. J., Kennea, N., Dyet, L., Saeed, N., Stark, J., Maalouf, E., Duggan, P., Ajayi-Obe, M., Hajnal, J., Allsop, J. M., Boardman, J., Rutherford, M. A., Cowan, F., Edwards, A. D., 2006. Abnormal cortical development after premature birth shown by altered allometric scaling of brain growth. PLoS Medicine 3 (8), 1382–1390.

Kazemi, K., Moghaddam, H. A., Grebe, R., Gondry-Jouet, C., Wallois, F., 2007. A neonatal atlas template for spatial normalization of whole-brain magnetic resonance images of newborns: Preliminary results. NeuroImage 37 (2), 463–73.

Klein, A., Andersson, J., Ardekani, B. a., Ashburner, J., Avants, B., Chiang, M.-C., Christensen, G. E., Collins, D. L., Gee, J., Hellier, P., Song, J. H., Jenkinson, M., Lepage, C., Rueckert, D., Thompson, P., Vercauteren, T., Woods, R. P., Mann, J. J., Parsey, R. V., 2009. Evaluation of 14 non-

linear deformation algorithms applied to human brain MRI registration. NeuroImage 46 (3), 786–802.

Kuklisova-Murgasova, M., Aljabar, P., Srinivasan, L., Counsell, S. J., Doria, V., Serag, A., Gousias, I. S., Boardman, J. P., Rutherford, M. A., Edwards, A. D., Hajnal, J. V., Rueckert, D., 2011. A dynamic 4D probabilistic atlas of the developing brain. NeuroImage 54 (4), 2750–63.

Kwon, S. H., Vasung, L., Ment, L. R., Huppi, P. S., 2014. The role of neuroimaging in predicting neurodevelopmental outcomes of preterm neonates. Clinics in Perinatology 41 (1), 257–283.

Lorenzen, P., Davis, B., Joshi, S., 2005. Unbiased atlas formation via large deformations metric mapping. In: Duncan, J. S., Gerig, G. (Eds.), Proceedings of 8th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). Vol. 8. Springer Berlin Heidelberg, pp. 411–8.

Lorenzen, P., Prastawa, M., Davis, B., Gerig, G., Bullitt, E., Joshi, S., 2006. Multi-modal image set registration and atlas formation. Medical Image Analysis 10 (3), 440–51.

Lorenzi, M., Ayache, N., Frisoni, G. B., Pennec, X., 2013. LCC-Demons: a robust and accurate symmetric diffeomorphic registration algorithm. NeuroImage 81, 470–83.

Makropoulos, A., Aljabar, P., Wright, R., Hüning, B., Merchant, N., Arichi, T., Tusor, N., Hajnal, J. V., Edwards, A. D., Counsell, S. J., Rueckert,

D., 2016. Regional growth and atlasing of the developing human brain. NeuroImage 125, 456–78.

Makropoulos, A., Gousias, I., Ledig, C., Aljabar, P., Serag, A., Hajnal, J., Edwards, A., Counsell, S., Rueckert, D., 2014. Automatic whole brain MRI segmentation of the developing neonatal brain. IEEE Transactions on Medical Imaging 33 (9), 1818–31.

Makropoulos, A., Robinson, E. C., Schuh, A., Wright, R., Fitzgibbon, S., Bozek, J., Counsell, S. J., Steinweg, J., Passerat-Palmbach, J., Lenz, G., Mortari, F., Tenev, T., Duff, E. P., Bastiani, M., Cordero-Grande, L., Hughes, E., Tusor, N., Tournier, J.-D., Hutter, J., Price, A. N., Murgasova, M., Kelly, C., Rutherford, M. A., Smith, S. M., Edwards, D. A., Hajnal, J. V., Jenkinson, M., Rueckert, D., 2017. The Developing Human Connectome Project: a minimal processing pipeline for neonatal cortical surface reconstruction. bioRxiv, 1–43.

Modat, M., Daga, P., Cardoso, M. J., Ourselin, S., Ridgway, G. R., Ashburner, J., 2012. Parametric non-rigid registration using a stationary velocity field. In: Proceedings of IEEE Workshop on Mathematical Methods in Biomedical Image Analysis. pp. 145–150.

Modat, M., Ridgway, G. R., Taylor, Z. A., Hawkes, D. J., Fox, N. C., Ourselin, S., 2009. A parallel-friendly normalised mutual information gradient for free-form registration. In: Proceedings of the Society of Photographic Instrumentation Engineers (SPIE), Medical Imaging: Image Processing. Vol. 7259.

Modat, M., Ridgway, G. R., Taylor, Z. A., Lehmann, M., Barnes, J., Hawkes, D. J., Fox, N. C., Ourselin, S., 2010. Fast free-form deformation using graphics processing units. Computer Methods and Programs in Biomedicine 98 (3), 278–84.

Molnár, Z., Rutherford, M., 2013. Brain maturation after preterm birth. Science Translational Medicine 5 (168), 168ps2.

Noblet, V., Heinrich, C., Heitz, F., Armspach, J. P., 2008. Symmetric non-rigid image registration: Application to average brain templates construction. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (Eds.), Proceedings of 11th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). Springer Berlin Heidelberg, pp. 897–904.

Noorizadeh, N., Kazemi, K., Grebe, R., Helfroush, M. S., Mahmoudzadeh, M., Golshaeyan, N., Wallois, F., 2013. Symmetric brain atlas template for newborns brain asymmetry studies, 6–9.

Oishi, K., Mori, S., Donohue, P. K., Ernst, T., Anderson, L., Buchthal, S., Faria, A., Jiang, H., Li, X., Miller, M. I., van Zijl, P. C. M., Chang, L., 2011. Multi-contrast human neonatal brain atlas: application to normal neonate development analysis. NeuroImage 56 (1), 8–20.

Ou, Y., Akbari, H., Bilello, M., Da, X., Davatzikos, C., 2014. Comparative evaluation of registration algorithms in different brain databases with varying difficulty: Results and insights. IEEE Transactions on Medical Imaging 33 (10), 2039–65.

Pai, A., Sommer, S., Sorensen, L., Darkner, S., Sporring, J., Nielsen, M., 2015. Kernel bundle diffeomorphic image registration using stationary velocity fields and wendland basis functions. IEEE Transactions on Medical Imaging 53 (6), 1369–80.

Pennec, X., Arsigny, V., 2013. Exponential barycenters of the canonical Cartan connection and invariant means on Lie groups. In: Nielsen, F., Bhatia, R. (Eds.), Matrix Information Geometry. Springer Berlin Heidelberg, pp. 123–66.

Robbins, S., Evans, A. C., Collins, D. L., Whitesides, S., 2004. Tuning and comparing spatial normalization methods. Medical Image Analysis 8 (3), 311–23.

Rueckert, D., Aljabar, P., Heckemann, R. A., Hajnal, J. V., Hammers, A., 2006. Diffeomorphic registration using B-splines. In: Larsen, R., Nielsen, M., Sporring, J. (Eds.), Proceedings of 9th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). pp. 702–9.

Rueckert, D., Sonoda, L. I., Hayes, C., Hill, D. L., Leach, M. O., Hawkes, D. J., 1999. Nonrigid registration using free-form deformations: application to breast MR images. IEEE Transactions on Medical Imaging 18 (8), 712–21.

Rutherford, M. A. (Ed.), 2002. MRI of the neonatal brain, 4th Edition. W.B. Saunders.
URL http://www.mrineonatalbrain.com/

Schuh, A., Murgasova, M., Makropoulos, A., Ledig, C., Counsell, S. J., Hajnal, J. V., Aljabar, P., Rueckert, D., 2014. Construction of a 4D brain atlas and growth model using diffeomorphic registration. In: STIA. Cambridge, MA.

Sederberg, T. W., Parry, S. R., 1986. Free-form deformation of solid geometric models. In: SIGGRAPH Proceedings of the 13th Annual Conference on Computer Graphics and Interactive Techniques. pp. 151–60.

Seghers, D., Agostino, E. D., Maes, F., Vandermeulen, D., 2004. Construction of a brain template from MR images using state-of-the-art registration and segmentation techniques. In: Barillot, C., Haynor, D. R., Hellier, P. (Eds.), Proceedings of 7th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). pp. 696–703.

Serag, A., Aljabar, P., Ball, G., Counsell, S. J., Boardman, J. P., Rutherford, M. A., Edwards, A. D., Hajnal, J. V., Rueckert, D., 2012a. Construction of a consistent high-definition spatio-temporal atlas of the developing brain using adaptive kernel regression. NeuroImage 59 (3), 2255–65.

Serag, A., Kyriakopoulou, V., Rutherford, M. A., Edwards, A. D., Hajnal, J. V., 2012b. A multi-channel 4D probabilistic atlas of the developing brain: Application to fetuses and neonates. Annals of the British Machine Vision Association (BMVA) 2012 (3), 1–14.

Shen, D., Davatzikos, C., 2002. HAMMER: hierarchical attribute matching mechanism for elastic registration. IEEE Transactions on Medical Imaging 21 (11), 1421–39.

Shi, F., Yap, P.-t., Wu, G., Jia, H., Gilmore, J. H., Lin, W., Shen, D., 2011. Infant brain atlases from neonates to 1- and 2-year-olds. PLoS One 6 (4), e18746.

Sled, J. G., Nossin-Manor, R., 2013. Quantitative MRI for studying neonatal brain development. Neuroradiology 55 (S2), 97–104.

Smith, S., 2002. Fast robust automated brain extraction. Human Brain Mapping 17 (3), 143–55.

Studholme, C., 2003. Simultaneous population based image alignment for template free spatial normalisation of brain anatomy. In: Gee, J. C., Maintz, J. B. A., Vannier, M. W. (Eds.), Proceedings of International Workshop on Biomedical Image Registration. Springer Berlin Heidelberg, pp. 81–90.

Studholme, C., Cardenas, V., 2004. A template free approach to volumetric spatial normalization of brain anatomy. Pattern Recognition Letters 25 (10), 1191–202.

Thompson, P. M., Toga, A. W., 2002. A framework for computational anatomy. Computing and Visualization in Science 5 (1), 13–34.

Tustison, N., Avants, B., Cook, P., Zheng, Y., Egan, A., Yushkevich, P., Gee, J., 2010. N4ITK: Improved N3 bias correction. IEEE Transactions on Medical Imaging 29 (6), 1310–20.

Tustison, N. J., Avants, B. B., 2013. Explicit B-spline regularization in diffeomorphic image registration. Frontiers in Neuroinformatics 7, 39.

Vercauteren, T., Pennec, X., Perchant, A., Ayache, N., 2008. Symmetric log-domain diffeomorphic registration: A demons-based approach. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (Eds.), Proceedings of 11th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). Springer Berlin Heidelberg, pp. 754–61.

Vercauteren, T., Pennec, X., Perchant, A., Ayache, N., 2009. Diffeomorphic demons: efficient non-parametric image registration. NeuroImage 45 (1), S61–72.

Woods, R. P., Grafton, S. T., Watson, J. D., Sicotte, N. L., Mazziotta, J. C., 1998. Automated image registration: II. Intersubject validation of linear and nonlinear models. Journal of Computer Assisted Tomography 22 (1), 153–65.

Wu, G., Wang, Q., Jia, H., Shen, D., 2012. Feature-based groupwise registration by hierarchical anatomical correspondence detection. Human Brain Mapping 33 (2), 253–71.

Yoo, T. S., 2004. Insight into Images: Principles and Practice for Segmentation, Registration, and Image Analysis. AK Peters Ltd.

Zöllei, L., 2006. A unified information theoretic framework for pair- and group-wise registration of medical images. Ph.D. thesis, Massachusetts Institute of Technology.