

## Unbiased average age-appropriate atlases for pediatric studies

Vladimir Fonov<sup>a,\*</sup>, Alan C. Evans<sup>a</sup>, Kelly Botteron<sup>b</sup>, C. Robert Almlı́<sup>c,d</sup>,  
Robert C. McKinstry<sup>e,f</sup>, D. Louis Collins<sup>a</sup>  
and the Brain Development Cooperative Group<sup>1</sup>

<sup>a</sup> McConnell Brain Imaging Centre, Montreal Neurological Institute, Montreal, QC, Canada

<sup>b</sup> Department of Psychiatry, Washington University, St. Louis, MO, USA

<sup>c</sup> Developmental Neuropsychobiology Laboratory, Department of Neurology, Programs in Neuroscience, Occupational Therapy, Washington University School of Medicine, St. Louis, MO, USA

<sup>d</sup> Developmental Neuropsychobiology Laboratory, Department of Psychology, Programs in Neuroscience, Occupational Therapy, Washington University School of Medicine, St. Louis, MO, USA

<sup>e</sup> Mallinckrodt Institute of Radiology, Washington University Medical Center, St. Louis, MO, USA

<sup>f</sup> St. Louis Children's Hospital, Washington University Medical Center, St. Louis, MO, USA

### ARTICLE INFO

#### Article history:

Received 22 December 2009

Revised 1 July 2010

Accepted 16 July 2010

Available online 23 July 2010

#### Keywords:

Atlas template

Registration

Pediatric image analysis

### ABSTRACT

Spatial normalization, registration, and segmentation techniques for Magnetic Resonance Imaging (MRI) often use a target or template volume to facilitate processing, take advantage of prior information, and define a common coordinate system for analysis. In the neuroimaging literature, the MNI305 Talairach-like coordinate system is often used as a standard template. However, when studying pediatric populations, variation from the adult brain makes the MNI305 suboptimal for processing brain images of children. Morphological changes occurring during development render the use of age-appropriate templates desirable to reduce potential errors and minimize bias during processing of pediatric data. This paper presents the methods used to create unbiased, age-appropriate MRI atlas templates for pediatric studies that represent the average anatomy for the age range of 4.5–18.5 years, while maintaining a high level of anatomical detail and contrast. The creation of anatomical T1-weighted, T2-weighted, and proton density-weighted templates for specific developmentally important age-ranges, used data derived from the largest epidemiological, representative (healthy and normal) sample of the U.S. population, where each subject was carefully screened for medical and psychiatric factors and characterized using established neuropsychological and behavioral assessments. Use of these age-specific templates was evaluated by computing average tissue maps for gray matter, white matter, and cerebrospinal fluid for each specific age range, and by conducting an exemplar voxel-wise deformation-based morphometry study using 66 young (4.5–6.9 years) participants to demonstrate the benefits of using the age-appropriate templates. The public availability of these atlases/templates will facilitate analysis of pediatric MRI data and enable comparison of results between studies in a common standardized space specific to pediatric research.

© 2010 Elsevier Inc. All rights reserved.

### Introduction

Magnetic resonance imaging (MRI) has emerged as the premier modality of noninvasive imaging of normal structural and metabolic development of the brain in both infants and children. With the advent of modern MRI methods in the last 20 years, multiple groups have reported age-related changes in gray matter (GM) and white matter (WM) volumes, extent of myelination, and subcortical structures (Jernigan and Tallal, 1990; Jernigan et al., 1991; Filipek et al., 1994; Pfefferbaum et al., 1994; Blatter et al., 1995; Caviness et al., 1996, 1999; Giedd et al., 1996a,b, 1999; Reiss et al., 1996; Lange et al., 1997; Kennedy et al., 1998, 2003; Paus et al., 1999; Sowell et al., 1999,

2002, 2003, 2004a,b; Courchesne et al., 2000; Bartzokis et al., 2001; Blanton et al., 2001, 2004; De Bellis et al., 2001; Durston et al., 2001; Mazziotta et al., 2001a,b; Gogtay et al., 2002, 2004). However, significant variability has generally been seen in the volumetric and metabolic data across populations and between genders, complicated by reports of differences in regionally specific changes within individual brain growth trajectories (Giedd et al., 1996a, 1999; Gogtay et al., 2004). Furthermore, because most prior studies have limited number of subjects and included analysis of T1-weighted (T1w) data only, previous findings have not been easily extrapolated among studies, between specific age-groups, or to the general pediatric population.

To address these issues, the National Institutes of Health (NIH) MRI Study of Normal Brain Development has developed a large, combined cross-sectional and longitudinal, population-based study design to generate a meaningful normative database of T1-weighted (T1w),

\* Corresponding author.

E-mail address: [vladimir.fonov@mcgill.ca](mailto:vladimir.fonov@mcgill.ca) (V. Fonov).

<sup>1</sup> See Appendix A.

T2w and proton density weighted (PDw) structural images that will be useful in the study of both normal brain development, and childhood neurological and neuropsychiatric diseases (Evans and B.D.C. Group, 2006; Almlí et al., 2007). Previous reports (Evans and B.D.C. Group, 2006; Almlí et al., 2007; Waber et al., 2007) have detailed the study's design, imaging protocols and analysis, and behavioral/cognitive testing methods. This report describes the creation and usefulness of age-appropriate atlases based on the Objective 1 data (i.e., subjects aged 4.5–18.5 years) from the MRI Study of Normal Brain Development.

In addition, data characterizing cognitive and behavioral constructs for all infants, children, and adolescents in the study were acquired along with structural imaging data to enable examination and characterization of correlations between structure and function associated with ongoing developmental processes. Our hope is that the construction of a population-based, representative database of MRI structural and metabolic data correlated with validated cognitive/behavioral measurements will improve our ability to detect and interpret differences in brain development that correspond to pediatric psychiatric and neurological disorders.

Many automated techniques for registration, tissue classification, and statistical analysis use a template brain (Mazziotta et al., 2001a,b), including mni\_autoreg (Collins et al., 1994), SPM (Ashburner and Friston, 1997), and FSL (Smith et al., 2004). However, such techniques are not ideal for pediatric analysis because the templates were created by averaging MRI data from young adults. Since the developing brain is not simply a smaller version of an adult brain, the use of adult templates may introduce a bias in analysis. For example, Muzik et al. (2000) showed that, when using an adult template with SPM96, the registration of pediatric data was more variable than that of adult data. In addition, Wilke et al. (2002a,b) found that the analysis of pediatric data depended greatly on processing techniques and spatial normalization methods. In electroencephalography source analysis, Hoeksma et al. (2005) found differences between pediatric and adult data, and demonstrated that an adult target was less adequate for pediatric data. Machielsen et al. (2007) also found standard registration methods using the MNI (Montreal Neurological Institute) template to be less robust with pediatric data.

These types of problems indicate a need for developmental age specific brain templates. To achieve this age specificity, some studies have used data from a single subject for the template. For example, Jelacic et al. (2006) built an interactive Web-based atlas for subjects under 4 years of age that facilitates the comparison of a given subject with standard datasets from a database. Shan et al. (2006) built a digital pediatric brain structure atlas from T1w MRI scans from a single 9-year-old subject. However, the main problem with using single subject templates is that, despite being a typical healthy individual, the chosen subject may represent an extreme tail of the normal distribution for some brain regions. Moreover, a single subject template cannot represent the anatomical variability in the population. The solution to these problems is to build atlases from multiple subjects. In the pediatric literature, Joshi et al. (2004) used unbiased diffeomorphic atlas construction techniques to build a template of eight 2-year-old subjects. Kazemi et al. (2007) developed a neonatal atlas for spatial normalization of whole brain MRI, based on data from seven subjects. Bhatia et al. (2007) used an expectation-maximization framework to build an MRI atlas for 1- and 2-year-olds. However, these atlases either were created from a small number of subjects or cover a very narrow age range. More recently, Wilke et al. (2008) created a "Template-O-Matic" toolbox for creating population-specific templates based on the unsupervised tissue segmentation and linear coregistration of individual pediatric scans with regression on independent variables such as age and gender. Although this enables a user to generate an appropriate intensity average template volume for a particular study, anatomical details may be blurred in regions of high variability such as the cortex because only linear registration is

used. Therefore, in this paper we create a series of age-specific, nonlinearly registered pediatric templates from 324 subjects within the age range of 4.5 to 18.5 years that include T1w, T2w, and PDw averages as well as average tissue maps for GM, WM, and cerebrospinal fluid (CSF). Because the atlas-building process uses nonlinear registration, these templates have the advantage of being age-specific while retaining significant anatomical detail.

Many groups have investigated techniques for creating an anatomical average from a group of subjects such that the result is representative of the population. In some of the first work published on this topic, Guimond et al. (1998, 2000) developed methods of building a template atlas with both average intensity and average shape. These methods begin by selecting or creating an initial template, which may be a single subject or a linear average like the MNI305 volume used in mni\_autoreg, SPM, or FSL. Each subject in the group is then nonlinearly registered to the template, and the estimated transformation is used to resample the subject's MRI in the template space. A voxel-by-voxel average is computed across all subjects to produce the average-intensity image, and to warp this image to have an average shape, all nonlinear transformations are averaged together. The inverse of the average nonlinear transformation is then applied to resample the average-intensity image, resulting in a template with both an average unbiased shape and average intensity. To account for imperfections in the nonlinear registration procedure, multiple iterations are performed, each time using the new template as the registration target, until the difference between two successive templates is smaller than some threshold.

This procedure has been used as a general strategy in many subsequent papers that addressed different issues in the template-building process, such as the selection of the first template, data used to build the template, similarity function used to drive the registration, type of nonlinear transformation modeled, and method used for averaging. For example, Shattuck et al. (2008) used the nonlinear registration methods of AIR (Woods et al., 1998), FSL (Smith et al., 2004), and SPM to create average targets from 40 healthy normal controls. Wang et al. (2005) evaluated different template construction strategies for atlas-based segmentation and found that an intensity-average template based on nonlinear coregistration was best for the segmentation of 49 brain regions. Joshi and Miller (2000) and Joshi and Miller (2000), Joshi et al. (2004) used diffeomorphic registration to build unbiased average templates, a technique later modified by Lorenzen et al. (2005) to create an unbiased atlas as a Fréchet mean estimation process. Bhatia et al. (2007) interleaved tissue classification and nonlinear registration of the tissue probability maps to build an average three-dimensional (3D) MRI template.

To facilitate the processing of pediatric imaging data, we have produced a number of age-appropriate, representative, average brain templates using nonlinear deformation to standard coordinates. The construction of a registration target that is both age-appropriate and representative will allow meaningful correlation of anatomical changes and development. Furthermore, nonlinear deformation methods were used for their superior spatial detail and ability to register anatomies from different subjects and across different ages.

Here, we present the procedure used to create unbiased atlas templates that include a series of symmetric and asymmetric atlases. We created and compared atlases from two databases of MR images covering the age range of 4.5 to 43.5 years: (1) a collection of 324 pediatric (4.5–18.5 years) MRI scans from the NIH-funded MRI Study of Normal Brain Development (hereafter, NIHPD, for NIH pediatric database) (Evans and B.D.C. Group, 2006) and (2) an MRI database of young adult brains, using data from 152 subjects (aged 18.5–43.5 years) acquired at the Montreal Neurological Institute (MNI) as part of the International Consortium for Brain Mapping (known as the ICBM database) (Mazziotta et al., 1995). These data were used to create templates with the following characteristics: (1) average (over the population analyzed) normalized intensity; (2) average shape;

(3) (optionally) left–right symmetry; (4) high contrast-to-noise ratio; (5) high level of anatomical structural detail (as seen in the individual images); and (6) compatibility with new ICBM 152 space that is compatible with the older MNI305 stereotaxic space (Janke et al., 2006).

The main contributions of this paper concern the templates that are created and made available to the scientific community. To our knowledge, this is the only dataset containing (1) an epidemiologically ascertained sample of children aged 4.5 to 18.5 years old, representative of the U.S. population with respect to income (as a proxy for socioeconomic status) and race/ethnicity, (2) where each child has been carefully screened with respect to medical and psychiatric factors (including family history), and (3) has been very well characterized using a variety of standardized interviews, rating scales and cognitive tests (Evans and B.D.C. Group, 2006). These factors ensure that the templates will be useful as normative models.

In addition to the T1w templates for the NIHPD and the ICBM database, the following templates were also created: T2w and PDw templates, average brain masks and probabilistic atlases of GM, WM, and CSF maps. Finally, to demonstrate the usefulness of the pediatric templates, the bias of using a population-specific template is shown by comparing the results obtained using the NIHPD 4.5–18.5 templates and the new ICBM 152 template using deformation-based morphometry analysis (Chung et al., 2001).

## Materials and methods

### Creation of an unbiased template

Over the last several years, several competing techniques have been developed for building population-specific templates. The rationale behind building a population-specific atlas is described in (Mazziotta et al., 2001a,b); the practical impact of such an atlas on the analysis of functional data is described in Good et al. (2001), and its impact on the analysis of pediatric data is given in (Wilke et al., 2002a,b, 2003; Kazemi et al., 2007). To reiterate the most important issues, an unbiased brain template is needed (1) to provide a registration target for automatic image processing techniques (e.g., those in Evans et al., 1993; Collins et al., 1994; Thompson and Toga, 2002); (2) to act as an atlas for volume estimation of brain regions (Amit et al., 1991; Christensen et al., 1994; Collins et al., 1999; Mazziotta et al., 2001a,b; Toga and Thompson, 2001; Thompson and Toga, 2002; Essen and David, 2002; Thompson and Toga, 2002; Toga and Thompson, 2007; Essen and David, 2005; Seghers et al., 2004; Grabner et al., 2006); and (3) to function as a reference for a particular population group in order to study intra- and inter-group variability or growth (Thompson and Toga, 2002; Gerig et al., 2006).

Recently, a number of algorithms have been published for constructing population-specific templates. The first approaches to building average templates were based on manual linear coregistration of individual scans into some kind of normalized reference space (e.g., Evans et al., 1993), a process later improved by using automatic tools (Collins et al., 1994) to register individual scans into the common space (Janke et al., 2006). Unfortunately, the variability of human brain anatomy leads to a limited resemblance between the average template and the real scans of individual subjects. Moreover, templates produced by linear registration were not very suitable for the automatic segmentation of brain substructures by deformable template algorithms (Carmichael et al., 2005). As described above, several methods were developed to produce a template more representative of the anatomy (Guimond et al., 1998, 2000, 2001; Mazziotta et al., 2001a,b; Bhatia et al., 2004; Joshi et al., 2004; Lorenzen et al., 2005; Essen and David, 2002; Wang et al., 2005). Generally, these methods may be classified into two types: (1) feature-matching algorithms that rely on matching homologous features of the individual scans and (2) intensity-matching algorithms that use

some generic cost function. The procedure described here belongs to the second type.

All templates described below use the original ICBM 152 (linear average) template as the initial reference target template volume for linear registration and intensity normalization. Our method is iterative, requiring  $N$  ( $2*N$  for the symmetric template) nonlinear registrations to be performed at each iteration step, where  $N$  is the number of subjects. We empirically show that the method converges to a stable solution after 20 iterations, thus requiring a total of  $20*N$  nonlinear registrations to be performed ( $40*N$  for the symmetric template).

### Nonlinear average

The work described here depends on the nonlinear registration engine of Automatic Nonlinear Image Matching and Anatomical Labeling (ANIMAL) (Louis Collins et al., 1995). While other non-linear registration techniques could have been chosen (Ardekani et al., 2005; Avants et al., 2006; Lorenzen and Joshi, 2003), we selected ANIMAL because we have extensive experience with the procedure and a recent study (Guizzard et al., 2009) has shown that when used with appropriate parameters, the results of the ANIMAL inter-subject registration procedure are comparable to ART (Ardekani et al., 2005) and SyN (Avants et al., 2008), the two top ranked registration procedures in the recent nonlinear registration evaluation paper of Klein et al. (2009). While in theory the linear elastic regularization used in ANIMAL does not guarantee the recovered transformation to be diffeomorphic, the set of registration parameters used here constrains the transformation to be smooth, bijective and invertible; characteristics needed for the atlas building procedure described below.

To estimate the required nonlinear transformation between a source and a target volume, the ANIMAL algorithm attempts to match hierarchically image gray-level intensity features in local neighborhoods arranged on a 3D grid by maximizing the cross-correlation of intensities between the source and target images. First, the deformations required to match blurred versions of the source and target data are estimated, producing a dense 3D deformation field, where a displacement vector is stored at each node of the field that best matches the local neighborhoods. Then, this deformation field is upsampled and used as input to the next iteration of the procedure, where the blurring is reduced and the estimation of the deformation field is refined. In this manner, large smooth deformations are recovered first, and finer, more local, deformations are recovered last. The schedule of grid step size, blurring, neighborhood size, and iterations is given in Table 1.

Our atlas generation technique is based on the work of (Guimond et al., 1998, 2001) and employs the principles of average model construction using elastic body deformations from (Miller et al., 1997). We use the *minimum deformation template* notation from the latter. Essentially, the problem can be formulated as follows: Given a set of  $n$  3D MRI volumes ( $I_1 \dots I_n$ ), our objective is to find a 3D template

**Table 1**

Nonlinear registration parameters. Step size is defined as the distance between control nodes for the free-form deformation recovered by ANIMAL. The blurring kernel is the size of the full-width-half-maximum of the Gaussian kernel used to blur the source and target data. The neighborhood size is the diameter of the local neighborhood used to estimate the local correlation that defined local similarity.

Iteration	Step size (mm)	Blurring kernel (mm)	Neighborhood size (mm)	Local iterations
1–4	32	16	96	20
5–8	16	8	48	20
9–12	8	4	24	20
13–16	4	2	12	10
17–20	2	1	6	10

$\Phi$ , which satisfies two constraints simultaneously, one for intensity, and one for the transformation. The first constraint is to minimize the mean squared intensity difference between the template  $\Phi$  and each subject  $I_i$ , transformed to match the template:

$$\Phi^* = \arg \min_{\Phi} \left[ \sum_{i=1}^n \int_{\text{volume}} (\Phi(v) - I_i(\Psi_{i,\Phi}(v)))^2 dv \right], \quad (1)$$

where  $v$  is a volume coordinate,  $\Psi_{i,\Phi}$  are the individual 3D mappings from the template  $\Phi$  to each subject volume  $I_i$ ,  $\Phi(v)$  is the template intensity at location  $v$ , and  $I_i(\Psi_{i,\Phi}(v))$  is the intensity in the subject's MRI after transformation by  $\Psi_{i,\Phi}$ . The transformation  $\Psi$  is constrained using simple elastic body model, such that for each subject  $i$ :

$$\Psi_{i,\Phi} = \arg \min_{\Psi} \int_{\text{volume}} (\Phi(v) - I_i(\Psi(v)))^2 dv \quad (2)$$

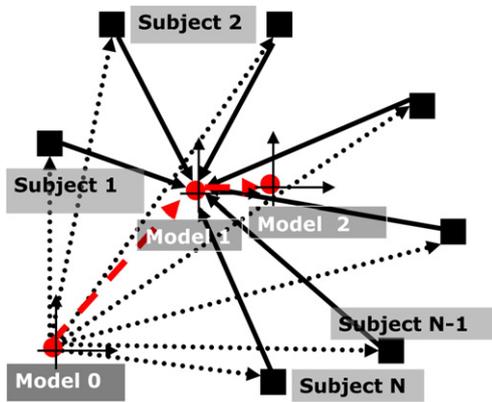
The second constraint is to minimize the magnitude of all deformations  $\Psi_{i,\Phi}$  required to map the template  $\Phi$  to each subject  $i$ :

$$\Phi^* = \arg \min_{\Phi} \left[ \sum_{i=1}^n \int_{\text{volume}} |\Psi_{i,\Phi}(v) - v|^2 dv \right], \quad (3)$$

In short, we are simultaneously minimizing Eqs. (1) and (3); Eq. (2) is minimized for each subject-template pair.

The transformation  $\Psi_{i,T}$  is represented with a dense deformation vector field  $h$ , such that  $\Psi(x) = x + h(x)$  and  $\Psi^{-1}(x) = x + h^{-1}(x)$  where  $h(x)$  may be defined on a discrete grid with a given distance (step size) between nodes, as in the ANIMAL algorithm (see Fig. 1). Following this formalism, we have developed an iterative algorithm minimizing the mean square differences in Eqs. (1) and (3). Each iteration of the algorithm interleaves minimization of both objective functions, first the mean square difference in terms of deformations (Eq. (3)) and second, the mean square difference in terms of intensity (Eq. (1)). To denote the mapping of each subject found at each consecutive step of the algorithm we use  $X_{i,k}$ , and the current approximation of the template  $T_k$ , when the algorithm converges  $X_{i,k} \rightarrow \Psi_{i,\Phi}, T_k \rightarrow \Phi^*$  producing the minimum deformation template  $\Phi^*$  and mapping  $\Psi_{i,\Phi}$ , from template to each subject. The algorithm is as follows:

1. Given  $T_k$  (the approximation template  $\Phi^*$  at iteration  $k$ ), for each scan  $I_i$ , calculate  $X_{i,k}$  (mappings from template to an individual scan  $i$ , on the iteration  $k$ ), using the  $Y_{i,k-1}^{-1}$  (inverse corrected mappings of the scan  $i$ , iteration  $k-1$ ) as a starting point. The identity transform for the first iteration and the linear ICBM 152 average was used as  $T_0$ .



**Fig. 1.** Schematic representation of the model building algorithm: dotted lines represent mapping of a voxel in the initial model (Model 0) to each subject, solid lines represent mapping of individual subjects into the next model (Model 1), dashed lines represent the voxel-wise residual error of the models at each iteration.

2. Calculate the residual error based on the average deformation  $X_{0,k}$  of the current template  $T_k$ :

$$X_{0,k} = x + \sum_n h_i(x) / n \quad (4)$$

3. Calculate corrected inverse mappings:  $Y_{i,k} = X_{i,k}^{-1} \cdot X_{0,k}$ , where “ $\cdot$ ” indicates composition of transformations. This step corresponds to the function minimization of Eq. (3) (i.e., deformation related), note that  $Y_{i,k}$  is defined in the space of each subject, and must be numerically inverted for use, hence the name *inverse mapping*.
4. Apply corrected inverse mappings to individual subjects and generate an average that will be used as a new template, thus minimizing Eq. (1) (i.e., intensity related):

$$T_{k+1}(x) = \sum_n I(Y_{i,k}(x)) / n. \quad (5)$$

5. Repeat from step 1 until convergence is reached.

In nonlinear registration, the process is repeated with diminishing step sizes in a hierarchical fashion. For the convergence condition, the root mean square (RMS) magnitude of the average residual deformation vector field generated in step 2 is computed, and the process is stopped once the difference between two subsequent steps falls below a certain threshold. In general, directly averaging deformation fields is not guaranteed to produce a diffeomorphic transformation. Some authors have suggested using a Log-Euclidean setting (Arsigny et al., 2006), however we do not use such a scheme. Our algorithm is similar to a numerical estimation technique, where the goal is to use a computationally simpler method that yields progressively smaller errors as the method converges. As such, the potential error incurred in this step becomes insignificant as the method converges. Our experiments showed that performing four iterations for a given step size was sufficient to achieve convergence at the given level of detail, down to a 2-mm step size. In contrast to the previously published method (Guimond et al., 1998, 2001), we always use the coordinate system of the current template to calculate nonlinear deformation fields  $X_i$ , thus ensuring that individual deformation vectors defined at each location have a common origin between different subjects. Moreover, information from the previous iteration is used to initialize the nonlinear registration at the next iteration, which is particularly important in terms of speed for the convergence of the iterative process.

### Symmetric model

As human brains have a certain degree of asymmetry (Toga and Thompson, 2003), the average template is expected to be asymmetric to reflect the average inequalities between the left and right hemispheres. However, in some studies, it may be desirable to treat both hemispheres equally. For example, when estimating left–right differences in a population, it is preferable not to use an asymmetric template, since it is difficult, if not impossible, to disambiguate the template's asymmetry from the population results. For example, detection of local volume differences with respect to the template should be equally sensitive on both sides of the brain.

To build a symmetric template, we introduce a transformation  $F$  that flips (or mirrors) a scan  $I$  in the  $x$  direction, around the midline. The flipped scan is denoted as  $F(I)$ . Also we denote the transformation that maps the template  $\Phi$  to the flipped scan as  $\Psi^f$ . From a mathematical point of view, we would like  $\Phi$  to have the following property: for each scan  $I$ , and corresponding template mapping  $\Psi$ :

$$\Psi(I) = F(\Psi^f(F^{-1}(I))) \quad (6)$$

i.e., registering the flipped image and then flipping the result should be the same as registering the unflipped image. (Note that the flipping operator has the property that  $F = F^{-1}$ .)

To achieve this, we have added another step into the non-linear registration portion of the iterative algorithm described above. If the registration procedure was perfect, we would only need to complete one registration, and then flip the result to build the symmetric model. However, to address imperfections in the ANIMAL registration procedure, we perform two non-linear registrations for each subject: one with the original image and with the flipped image. We do not treat the two registrations independently; instead we ensure that non-linear mappings calculated for the pair satisfy Eq. (6):

1. Given  $T_k$  (the approximation of template  $\phi^*$  at iteration  $k$ ), for each scan  $I_i$ , calculate  $X_{i,k}$  (mappings from template to an individual scan  $i$ , on the iteration  $k$ ), using the  $Y_{i,k-1}^{-1}$  (inverse corrected mappings from step 3 above) from the previous iteration as a starting point (identity is used for the first iteration). Calculate a mapping between the template  $T_k$  and the flipped version of the scan  $F(I_i)$ :  $X'_{i,k}$ , using the flipped version of  $Y_{i,k-1}^{-1}$  as a starting point. Then, calculate the average between  $X_{i,k}$  and  $F \cdot X'_{i,k} \cdot F^{-1}$ , producing  $X_{i,k}$  and  $X'_{i,k} = F \cdot X_{i,k} \cdot F^{-1}$ . From this point, transformations  $X_{i,k}$  and  $X'_{i,k}$  are treated independently, and the rest of the algorithm continues, averaging transformations as if twice as many subjects were used.
2. The new template calculated at the end of the iteration is replaced with

$$T'_{k+1} = \text{average}(T_{k+1}, F(T_{k+1}))$$

The resulting averages are always symmetric by this construction.

The final 2 mm symmetric and asymmetric templates for the entire NIHPD group were used as a starting point to generate the corresponding templates for the remaining age sub-ranges using the procedures described above.

**Subjects**

*NIH pediatric database*

In the course of the NIH-funded MRI Study of Normal Brain Development (see Evans and B.D.C. Group, 2006) for a description of the study and details of the MRI acquisition), MRI data was collected from 433 children aged 4.5–18.5 years (see Fig. 2 for a histogram of age distribution). In the project, T1w, T2w, and PDw data were obtained from six sites across the United States. The T1w data were acquired on either Siemens or General Electric (GE) 1.5 T scanners with a 3D RF-spoiled gradient echo acquisition with a repetition time

(TR) = 22–25 ms, echo time (TE) = 10–11 ms, flip angle 30°, refocusing pulse of 180°, sagittal acquisition with a field of view (FOV) of 256 mm SI and 204 mm anterior-posterior (AP). The slice thickness was 1.0 mm for Siemens and 1.1–1.5 mm for GE. The 2D T2w/PDw dual contrast fast spin echo sequence was acquired in the axial direction with TR = 3500 ms, TE1 = 15–17 ms, TE2 = 115–119 ms, FOV of 256 mm AP and 224 mm left-right (LR) with a 2 mm slice thickness. The ethics committees of the respective scanning sites approved the study, and informed consent for all subjects was obtained from the children’s parents or children of adult age (subjects older than 18 years). Although the MRI data contained both primary and fallback acquisitions, we used only the primary acquisition data because of its higher resolution and contrast. Quality control of the data was applied to eliminate scans that did not adhere to protocol or that suffered from severe motion artifacts. In the end, data from 324 subjects passed quality control and were used in the processing described below.

All NIHPD subjects were divided into the following age groups: (a) 4.5–18.5 years (all 324 subjects); (b) 4.5–8.5 years (82 subjects); (c) 7.0–11.0 years (112 subjects); (e) 7.5–13.5 years (162 subjects); (f) 10.0–14.0 years (105 subjects); (g) 13.0–18.5 years (108 subjects). These specific age group atlases were selected in an attempt to capture potentially critical aspects of brain development as they may be related to pubertal status. In our samples, puberty ranges from roughly 9–10 years through 16–17 years of age (based on the assessment by Petersen et al., 1988). Thus, the 4.5–8.5, 7.0–11.0, 7.5–13.5, 10.0–14.0, and 13.0–18.5 atlases would represent pre-puberty, pre- to early puberty, pre- to mid-puberty, early to advanced puberty, and mid-puberty through post-puberty, respectively. The selection of these ages was reinforced by graphic data presented by Waber et al. (2007), which consistently showed changes in the performance trajectories for most neuropsychological assessments between the ages of 9–10 years through 14–15 years. This selection also ensured each group contains a large number of subjects. Finally, because the age ranges are overlapping, the data from some subjects were used to generate several templates. Note that this should not cause any bias, as the templates are to be used independently.

*ICBM database*

Within the ICBM project, MRI data from 152 young normal adults (18.5–43.5 years; see Fig. 2 for a histogram of age distribution) were acquired on a Philips 1.5T Gyroscan (Best, Netherlands) scanner at the Montreal Neurological Institute (Mazziotta et al., 1995). The T1w data

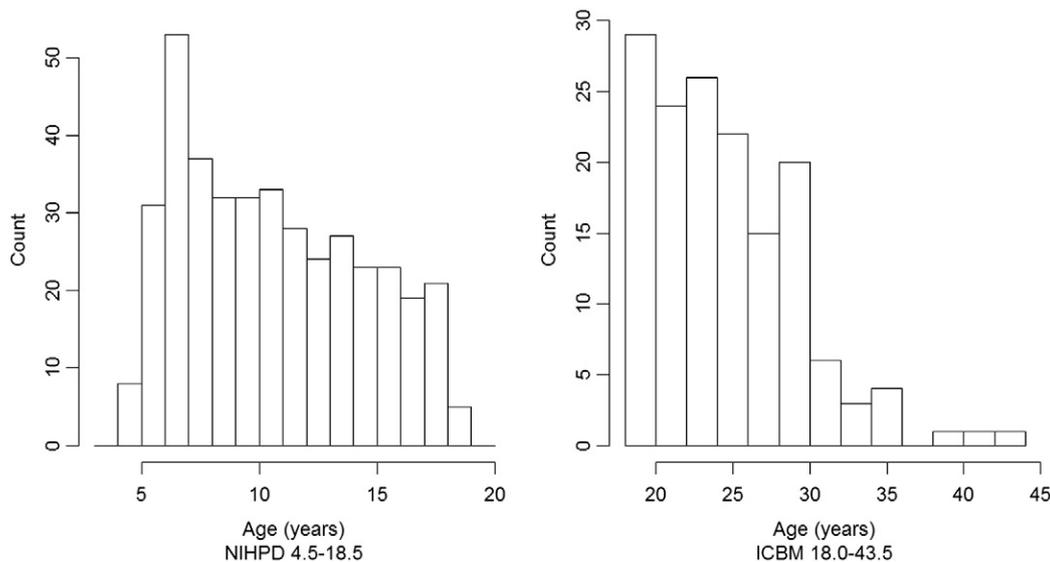


Fig. 2. NIHPD 4.5–18.5 age distribution (left) of the 324 subjects that passed QC and were included in template generation; ICBM 152 age distribution (right).

were acquired with a spoiled gradient echo sequence (sagittal acquisition, 140 contiguous 1-mm thick slices, TR = 18 ms, TE = 10 ms, flip angle 30°, rectangular FOV of 256 mm SI and 204 mm AP). The T2w/PDw data were acquired as a dual contrast fast spin echo sequence acquired in the axial direction with TR = 3300 ms, TE<sub>1</sub> = 34 ms, TE<sub>2</sub> = 120 ms, FOV of 256 mm AP and 224 mm LR, with a 2 mm slice thickness. The Ethics Committee of the Montreal Neurological Institute approved the study, and informed consent was obtained from all participants.

#### Image processing tools

The following data preprocessing steps were applied to all MRI scans prior to building the atlas: (1) N3 non-uniformity correction (Sled et al., 1998); (2) linear normalization of each scan's intensity to be in the same range as the ICBM 152 template by a single linear histogram scaling (Nyul and Udupa, 1999); (3) automatic linear (nine parameters) registration to the ICBM 152 stereotaxic space using mritotal from the MINC mni\_autoreg software package (Collins et al., 1994); and (4) brain mask creation using BET from the FSL package (Smith, 2002). Only the voxels within the brain volume after linear mapping into stereotaxic space were used for the nonlinear registration procedure described below.

For the actual implementation, we used programs from the MINC image processing framework, namely, minctracc for linear and nonlinear registration, xfmavg and xfminvert for operations on transformation maps, and mincaverage to calculate intensity averages, all of which are publicly available (packages.bic.mni.mcgill.ca). To resample the individual images, we used the B-spline algorithm from ITK based on (Thevenaz et al., 2000) (publicly available from [www.itk.org](http://www.itk.org)). All models were generated on a cluster consisting of 16 dual Pentium-III 1.4 GHz machines running Ubuntu Linux 8.04, using Sun Grid Engine 6.1 to distribute computations among the machines. The total time required to build an average template for 324 subjects was 90 hours, not counting the preprocessing.

## Results

#### Algorithm behavior

Average asymmetric and symmetric templates were generated for all subjects in the NIHPD group (4.5–18.5 years). Fig. 3 shows qualitatively the progression of the average asymmetric template and its standard deviation map at different iterations for a given step size. In the figure, the anatomical detail, in particular near the cortex, becomes increasingly better defined and the voxel-wise intensity variability is reduced with successive iterations.

To quantitatively track the convergence of the model, Fig. 4 shows the voxel-wise RMS magnitude of the residual error at each iteration for the asymmetric (black squares) and symmetric (red circles) fitting processes for all NIHPD subjects (4.5–18.5 years). Both curves show similar behavior with respect to the step size and number of iterations, although the displacements are understandably slightly larger for the symmetric model. Another measure of the goodness of fit is the change in the voxel-wise intensity standard deviation, calculated during the averaging of 324 individual warped scans (see Fig. 5). Note how the values of the residual error decrease for a given scale value and then increase at the next scale, before decreasing once again. These jumps are due to the decreases in scale (finer resolution), where more differences are recovered between subjects. If all scans are perfectly normalized, this graph should asymptotically reach the noise level of the acquisitions. The behavior is similar for the creation of the symmetric and asymmetric templates.

#### Average anatomy templates

The algorithm was applied to each of the age subgroups of the NIHPD and to the subjects in the ICBM database. Fig. 6 shows the final average asymmetric T1w templates for the six NIHPD age ranges and the ICBM young adult population: In each case, the templates provide significant anatomical detail in the central region, cerebellum, brainstem, and cortex, even though a large number of subjects were averaged for each template (e.g., 82 [4.5–8.5 years], 112 [7–11 years], and 152 subjects for the ICBM young adult average). See the T1w pediatric templates in Fig. 7 for better detail.

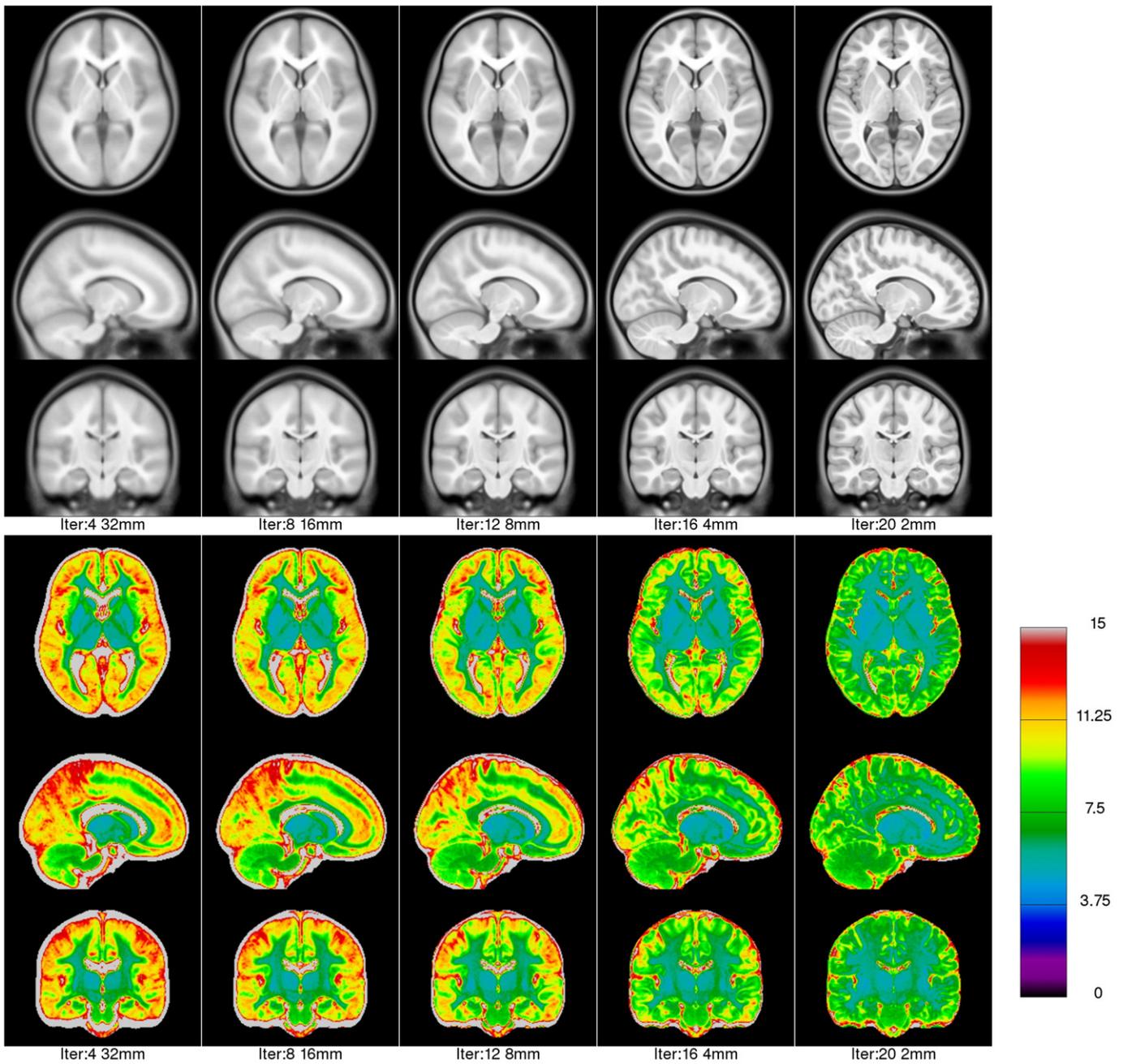
For each age-range dataset from the NIHPD and for the subjects in the ICBM database, templates of T2w and PDw modalities were generated (see Fig. 8). In addition, tissue probability maps were created using a genetic tissue classification algorithm on T1w images (Tohka et al., 2007), followed by a partial volume effect estimation of the tissue probability maps using all three modalities (T1w, T2w, PDw) (Tohka et al., 2004). For each subject, the individual T2w, PDw, and tissue probability maps were warped using the deformation field obtained during the creation of the T1w model, and averaged together to create the average T2w, PDw (c.f. Fig. 8), GM, WM, and CSF tissue probability maps, respectively. Fig. 9 shows the tissue probability maps for the full age range of the NIHPD and for all subjects in the ICBM database. Fig. 10 shows the detailed GM, WM, and CSF templates for the six age-specific NIHPD pediatric templates and the ICBM young adult template. Fig. 11 identifies some anatomical differences between the NIHPD 4.5–8.5 and ICBM 18.5–43.5 templates. The tissue probability maps, brain masks, and the average T1w, T2w, and PDw templates are publicly available in both MINC and NIFTI formats (<http://www.bic.mni.mcgill.ca/ServicesAtlases/NIHPD-obj1>).

Subtle morphological differences between each of these templates (Figs. 6–11) correspond to the maturation of the cerebral anatomy. For example, with all templates normalized to the same overall brain size, with age the corpus callosum thins, flattens slightly and lengthens slightly in the AP direction (Figs. 11 and 12). In addition, the lateral ventricles increase in size and the sulcal spaces widen in adulthood. In the frontal lobe, the ratio of WM to GM appears to increase with age. Further, the basal ganglia and thalamus appear wider and longer with increasing age, the pons enlarges with age and the posterior part of the brain (cerebellum and occipital pole) appears to shift in the superior direction, with the cerebellum widening with age (see Fig. 12).

#### Deformation-based morphometry example study

As an example of the potential effect the choice of template can have on analysis, a DBM study of the youngest subjects from the NIHPD was completed using four different target templates: the 7.0–11.0 years, the 10.0–14.0 years, and the 13.0–18.5 years NIHPD atlas templates as well as the ICBM young adult atlas template. The test set included subjects in the age range of 4.5–6.9 years that passed MRI quality control with the primary acquisition sequence and thus were comparable to the atlas templates ( $n = 66$  subjects). The objective here was not to complete a full DBM study, but rather to quantify the differences (or potential bias) that choice of template might have on eventual analysis. The templates are compared in a pair-wise fashion such that one template (7.0–11.0 years) is close to the appropriate age of the subjects and the second template is selected from the remaining three that are further away in age. These results clearly show that the bias (or difference) between templates increases as the difference in average age between templates is increased.

Each of the T1w MRI volumes in the test set was processed four times according to the standard data preprocessing steps (as described above), each time using one of the four aforementioned templates. After preprocessing, the nonlinear registration algorithm ANIMAL was used to estimate the mapping between each template



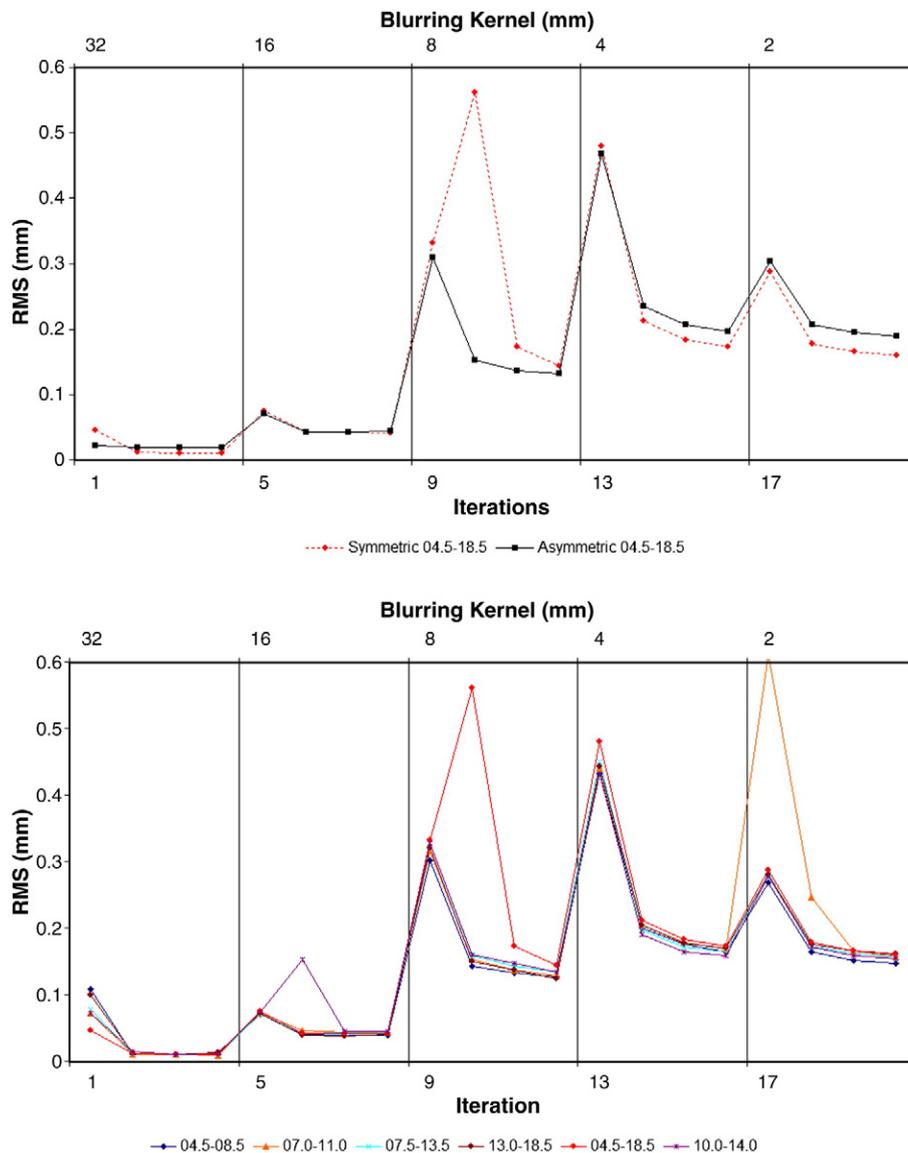
**Fig. 3.** Average asymmetric template (4.5–18.5 years old) generated at each level of fitting. The grey scale images show the intensity average anatomy, while the rainbow colour scale shows the intensity standard deviation for selected iterations in the hierarchical fitting process. One can see that as fitting progresses, anatomical features become less blurred and the intensity variability is reduced. The intensity range of the average data sets runs from 0 to 100.

and the preprocessed, brain-masked, linearly transformed data for each of the test subjects. This procedure yielded  $66 \times 4$  deformation fields.

The Jacobian determinant  $J$  was estimated for each node in each deformation field. The log Jacobian was computed, producing four fields of the local volume difference,  $\log J_{NIHPD\ 7.0-11.0}$ ,  $\log J_{NIHPD\ 10.0-14.0}$ ,  $\log J_{NIHPD\ 13.0-18.5}$ , and  $\log J_{ICBM}$ , for each subject of the test set. As the log Jacobian maps are defined in the space of the templates, for comparison, they need to be transformed into a common space. All NIHPD log Jacobian maps were transformed through the nonlinear deformation, by which each NIHPD template was mapped to the space of the ICBM template for analysis. A voxel-wise, pair-wise Student's  $t$ -test was then performed on the absolute difference from 0.0 between the resampled  $\log J_{NIHPD\ 7.0-11.0}$  and the  $\log J_{NIHPD\ 10.0-14.0}$ ,  $\log J_{NIHPD\ 13.0-18.5}$  and the  $\log J_{ICBM}$

templates. To account for the multiple-comparisons we have used False Discovery Rate (FDR) of 5% to calculate threshold for statistically significant differences (Genovese et al., 2002).

Since the test subjects are not drawn from the same age range as the target templates, the average log Jacobian map is not expected to be null. This is indeed the case, and for each template, the average magnitude of the deformation bias increases with age. The results of the Student's  $t$ -test shown in Fig. 13 demonstrate regions where this bias is significantly (corrected for multiple comparisons,  $FDR = 5\%$ ) different between pairs of templates. When the age difference between templates is small, for example when  $\log J_{NIHPD\ 7.0-11.0}$  and  $\log J_{NIHPD\ 10.0-14.0}$  are compared, the potentially biased regions are quite small and focused near the center of the brain. However, as the age between the templates increases, the size of the significantly



**Fig. 4.** RMS magnitude of the residual error vector field for each iteration (i.e., the bias in the average deformation for the current template), x axis shows the step-size in mm. On the top image, the symmetric (red circles) and asymmetric (black squares) NIHDP 4.5–18.5 models are compared. On the bottom, the different NIHDP age sub-ranges are plotted for the asymmetric atlas creation. One can see that at each iteration for each step size, the average RMS residual error magnitude is reduced, indicating that the optimization procedure is reaching a minima.

different regions increases as well. When an adult template is used to analyze a pediatric dataset in the 4.5–6.9 years range, there is a systematic bias in the estimation of tissue growth or shrinkage in the central regions of the brain, particularly around the ventricles. This result is not surprising, given the different appearance of the ventricles and the corpus callosum in these templates (see Figs. 6 and 11).

## Discussion

### On the method

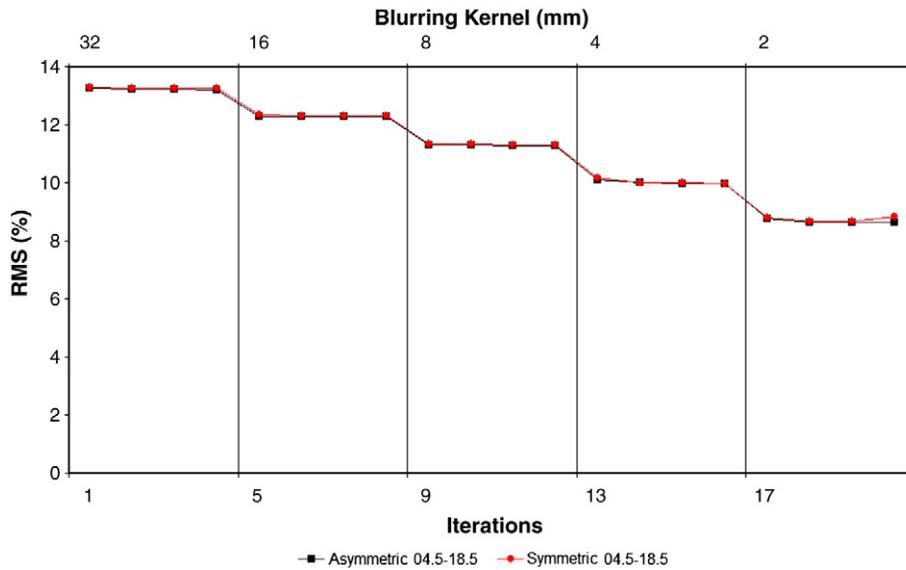
We have developed and characterized a method of creating unbiased symmetric and asymmetric templates of MRI data from large ensembles of subjects. Our method uses iterative refinement with successively finer scales of nonlinear registration to yield templates with a high degree of anatomical detail, even at the cortex. For this paper, we created unbiased symmetric and asymmetric templates of pediatric data for six (overlapping) age ranges, using MRI data available to qualified researchers from the NIH MRI Study of Normal

Brain Development. For comparison, we built a young adult template from MRI data from 152 young adults who had participated in the ICBM project (Mazziotta et al., 1995). In each case, the templates include nonlinear averages of T1w, T2w, and PDw images, average brain masks, and average GM, WM, and CSF maps. These atlases are publicly available from <http://www.bic.mni.mcgill.ca/ServicesAtlases>, where they can be viewed and downloaded.

Results of the iterative averaging procedure demonstrate that it is possible to generate average maps of anatomy from large numbers of subjects and retain detail not only for the central region of the brain, but also at the cortex (see Figs. 6–10). Figs. 3–5 show that the iterative process behaves well and converges with a small number of iterations.

### On the atlases

The templates were created for specific age ranges of subjects, selected from an epidemiological sample of normal healthy children 4.5–18.5 years old, that are representative of the U.S. population and have been carefully screened for medical and psychiatric factors and have been characterized using a series of standardized rating scales,



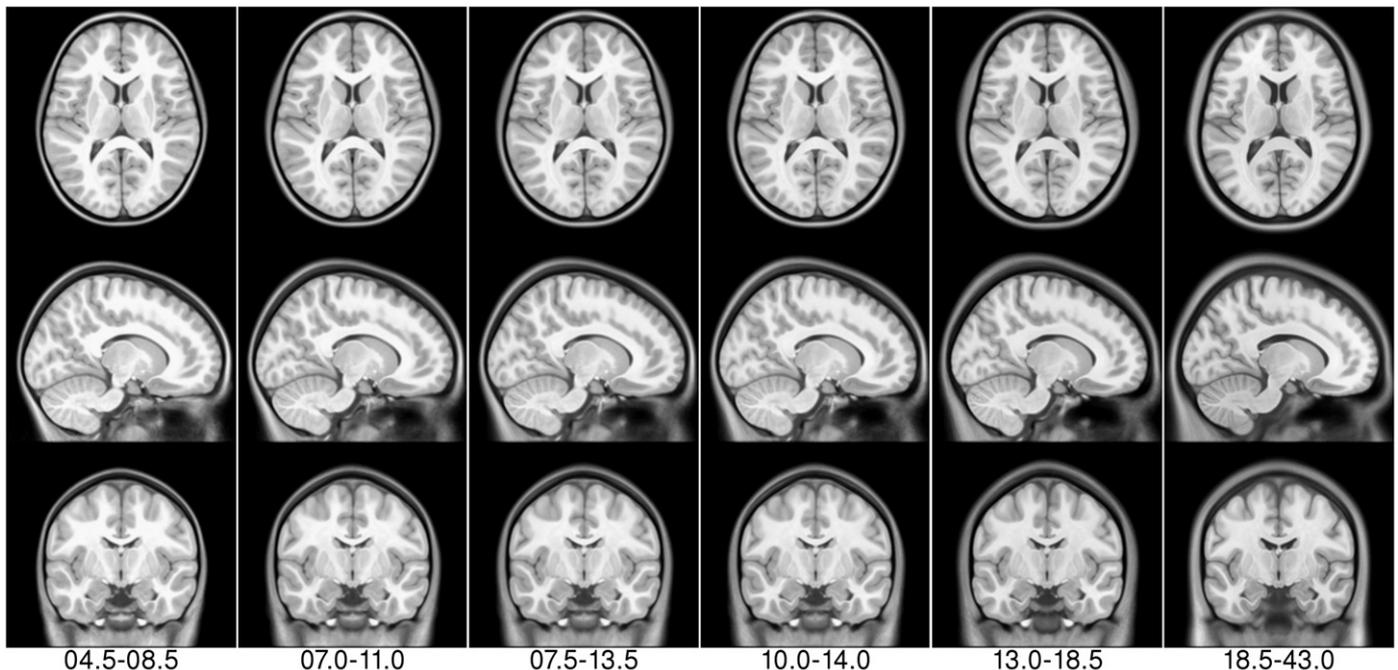
**Fig. 5.** RMS of intensity standard deviation (SD) between individual scans at each iteration for the creation of the NIHPD 4.5–18.5 years old atlas, x axis shows the step-size in mm. As the procedure advances, the RMS intensity SD between iterations decreases progressively for creation of both symmetric (red circles) and asymmetric (black squares) models.

cognitive tests and interviews. The use of such a cohort makes these templates practically useful for both clinical and more basic research in pediatric studies.

The generated symmetric and asymmetric templates should enable better unbiased analyses of pediatric data, with each type of template appropriate for certain types of analysis. For example, a symmetric template is better suited to analyze left–right differences in a particular population, whereas the asymmetric templates should be used as registration targets for all other studies where left–right comparison is not the major goal. In addition, one only has to manually segment one side of the brain when building a symmetric segmentation atlas.

Not surprisingly, our DBM study demonstrated that different templates give rise to different results; therefore, using an adult template

for pediatric data will yield different results than an age-appropriate template. Furthermore, comparisons between templates showed that this variation increases as the average age between templates increases. Experiments with the test set demonstrated that using an atlas close to the appropriate age yields fewer regions of potential bias than using an adult atlas. Indeed, Fig. 13 shows large regions where the deformation field is different from 1.0, indicating regions where the atlas is, on average, either larger or smaller than the corresponding regions of the internal test subjects. Since the experiments presented here only show that a difference exists, it is not possible to judge which template is preferable; selection of the *best* template will be task-specific. However, one might assume that a more accurate template (in terms of average morphometry) is better.



**Fig. 6.** NIHPD asymmetric templates (first six columns) + ICBM asymmetric template (rightmost column) for the T1w modality.

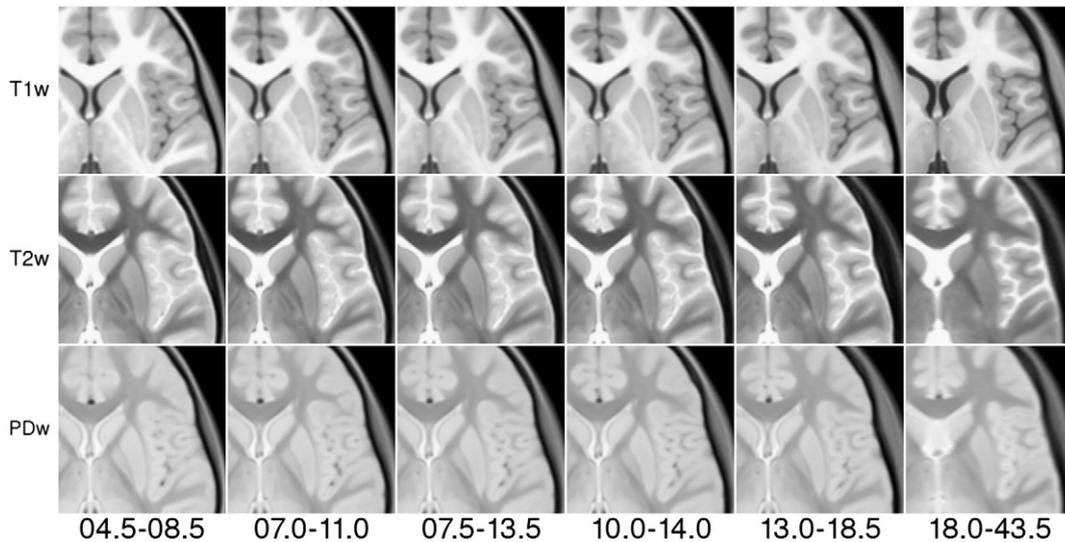


Fig. 7. Close up of the T1w, T2w and PDw (from top to bottom) atlas data to show cortical detail.

#### Comparison to other atlas building strategies

Our atlas-building strategy bears some similarities to previous iterative methods (Guimond et al., 1998, 2000; Joshi et al., 2004; Bhatia et al., 2007), but with some important differences: For example, in Guimond et al. (1998, 2000), the subject was registered to a template, and the deformations were averaged, inverted, and then applied to the average resampled data to remove bias. Here, we compute the deformations from the template to each subject, after linear registration in stereotaxic space, to allow estimation of the nonlinear deformation in the template space and justify vector averaging. Grabner et al. (2006) extended the work of Guimond et al. to include steps to build a symmetric template similar to those we use here. However, in contrast to both Guimond and Grabner, who used tri-linear interpolation to resample the MRI data, we use spline interpolation to yield slightly better results (Thévenaz et al., 2000). Guimond and Grabner also start

from scratch at each iteration; that is, at iteration  $n$ , they recompute the registration steps  $0, 1, \dots, n$ , where iteration  $0$  is a linear transformation to the target, whereas we use the transformation computed at iteration  $n-1$  as the starting point for iteration  $n$ , which helps maintain the stability of the process. Moreover, unlike Joshi et al. (2004), who used a large deformation diffeomorphic fluid approach that integrates streamlines (i.e., velocity field integration) into the deformation averaging approach, our work (and that of Guimond and Grabner) uses a linear elastic model, enabling a simpler averaging of vectors to estimate the mean deformation field. Finally, Bhatia et al. (2007) alternated group-wise combined segmentation and B-spline registration of the tissue classes in a global optimization procedure to form the templates. By contrast, our technique fits T1 intensities directly using a local optimization registration procedure.

Whereas we average data from all subjects within a group, the Template-O-Matic method (Wilke et al., 2008) uses statistical analysis

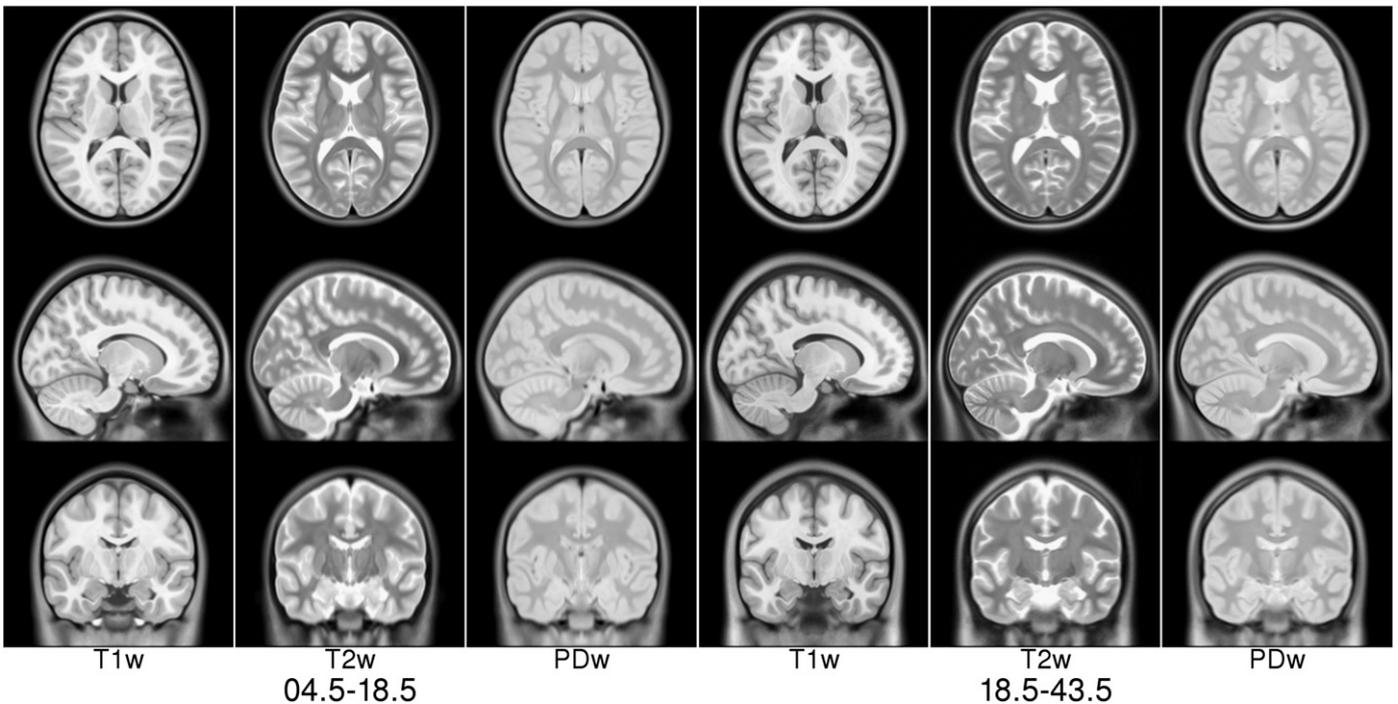
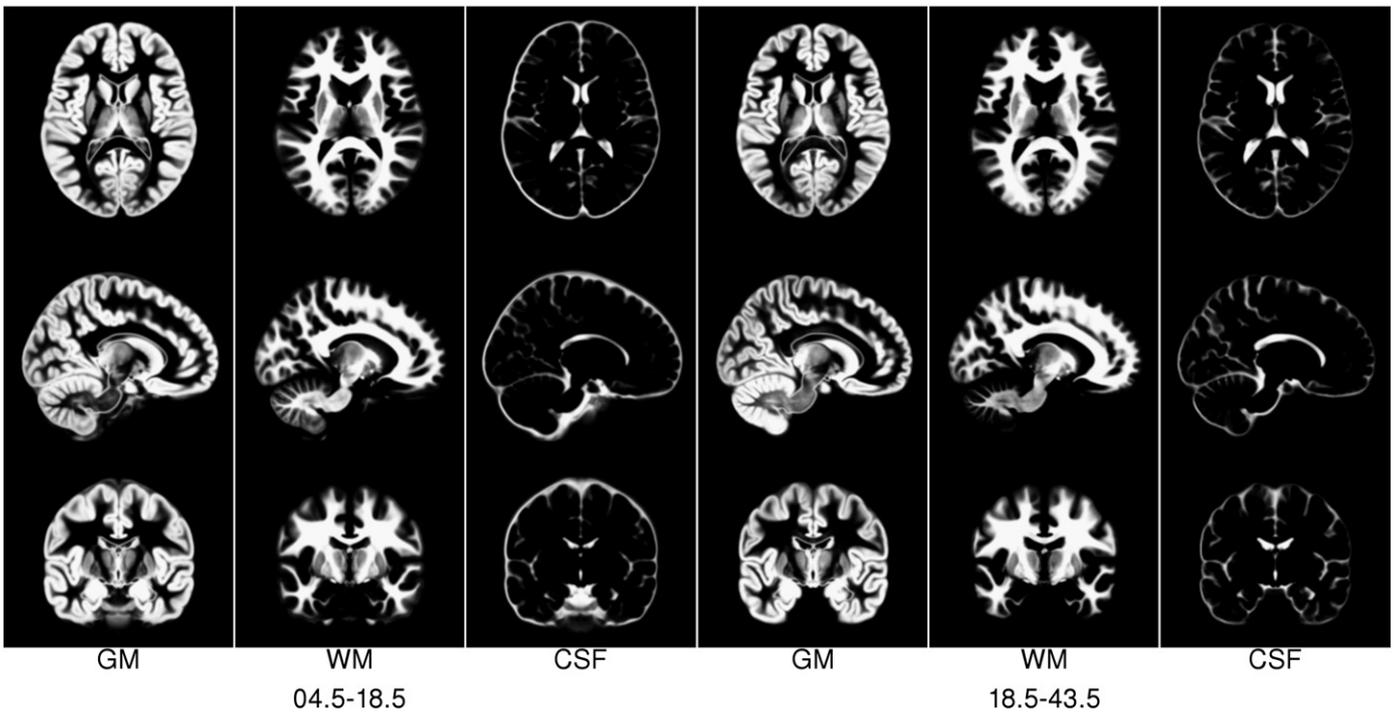


Fig. 8. NIHDP 4.5–18.5 template (left) and ICBM 18.5–43.0 template (right), showing the T1w, T2w and PDw average templates for each group.

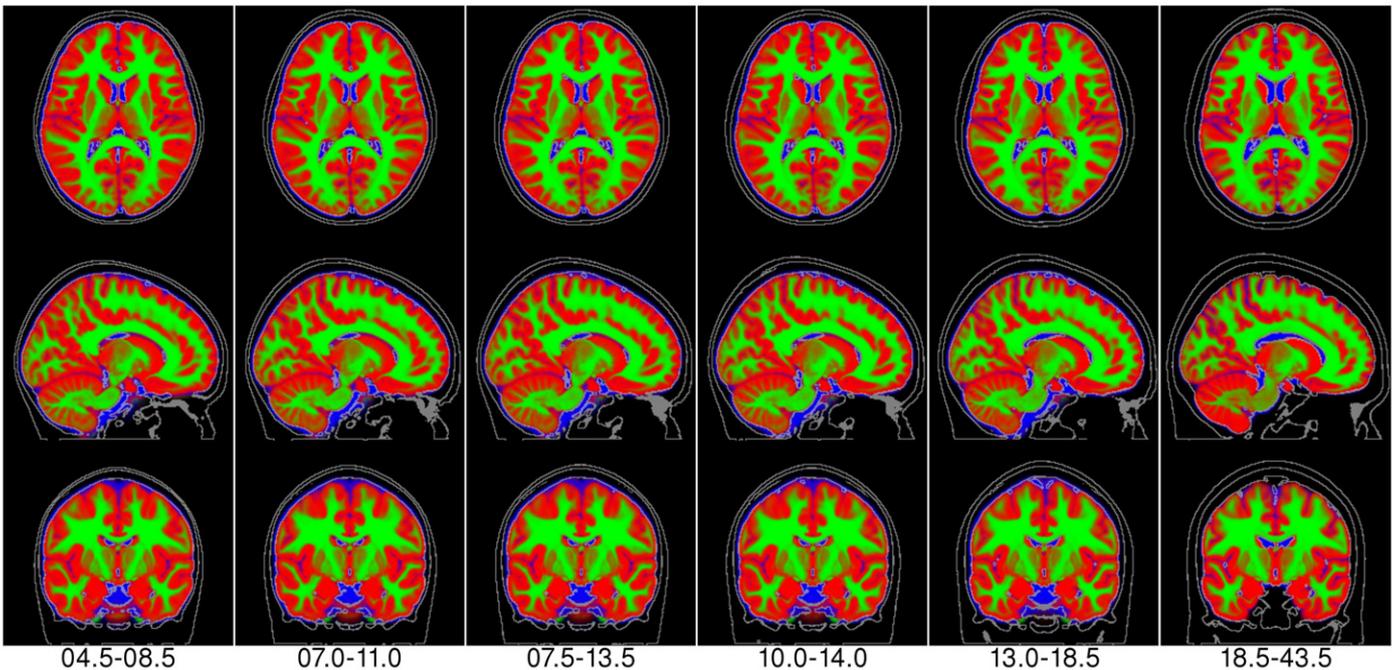


**Fig. 9.** Comparison of probabilistic atlas of the brain tissue types (GM, WM, CSF) for the NIHPD 4.5–18.5 atlas (leftmost 3 columns) and the ICBM 18.5–43.5 atlas (rightmost 3 columns). The brightest voxels indicate high probability of that tissue class. Note that the skin and skull outlines are overlaid on each subimage to facilitate comparisons.

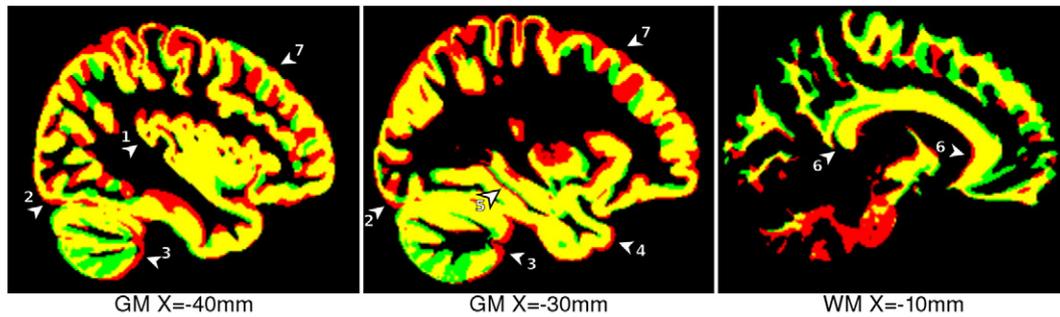
to compute weights of affine-registered GM and WM maps from each subject to generate customized tissue map templates that match a particular pediatric population under study. This differs from our method, where (1) age ranges are predefined according to hypotheses regarding aspects of brain development and (2) nonlinear iterative registration is used to align all datasets. The latter results in the much clearer and sharper average tissue templates seen in Fig. 9, compared to Fig. 3 of Wilke et al. (2008). Still, the statistical subject-weighting scheme deserves further investigation to determine, for instance, if it

can be combined with a nonlinear registration scheme similar to that described here.

A number of factors complicate the direct comparison of our template results with those published previously, including differences between the MRI data quality and number of subjects used to build the templates, the particular population studied, alignment method, registration strategy, scale of the deformation, and different metrics reported. With these caveats in mind, we compare our template results with those in the literature: Shan et al. (2006) created an atlas from the



**Fig. 10.** NIHPD templates (leftmost 6 columns) + ICBM template (rightmost column) of the combined tissue class atlas with red representing gray matter; green, white matter and blue color, CSF.



**Fig. 11.** Comparison between NIHDP 4.5–8.5 template (red) and ICBM 18.5–43.5 template (green), overlapping regions in yellow. The following anatomical differences are highlighted: (1) thicker insular cortex in pediatric atlas, (2) more posterior occipital pole in pediatric atlas, (3) different shape and GM/WM ratio in cerebellum, (4) more anterior temporal pole in pediatric atlas, (5) slightly different hippocampal shape, (6) flatter, thinner, longer corpus callosum in adult atlas, (7) thicker GM in pediatric atlas.

anatomy of a single 9-year-old subject. The atlases of Jelacic et al. (2006) allow the comparison of the anatomy of a given subject with those of other subjects, manually selected from a small group of standard normal scans. Kochunov et al. (2001), Park et al. (2005), and Wu et al. (2007) all described methods to select the *best* template from a collection of potential MRI scans. As a justification for using a single subject atlas, they cited the blurred appearance of older average templates such as the MNI305 or ICBM 152, which were created using only linear transformations. However, while a single subject template may be a good match globally for a specific subject under study, it is still possible that some local region of the template might represent an extreme of the normal distribution, which could potentially result in a biased analysis. Furthermore, when studying groups of subjects, it is necessary to align all subjects into a common coordinate space. With the *best* template strategy, this is impossible because there is a different *best* template for every subject being studied. By contrast, the atlases presented here represent the average anatomy of large groups of subjects; thus, they are less biased than atlases created from single subjects. In addition, the same template can be used as a common registration target for studies involving multiple subjects. Finally, the iterative nonlinear registration strategy used here results in templates with high anatomical detail throughout the brain, thus obviating the need to justify the use of a single subject atlas for registration.

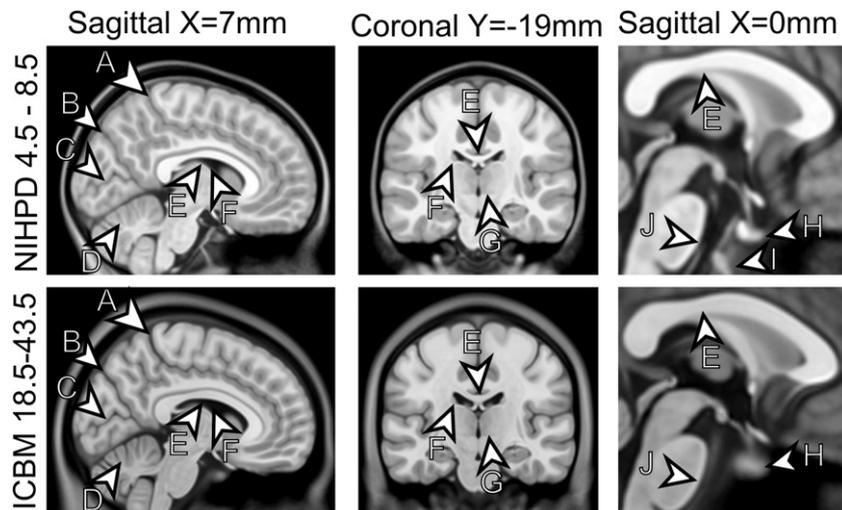
As far as multi-subject atlases, Joshi et al. (2004), Kazemi et al. (2007), and Bhatia et al. (2007) created atlases from eight, seven, and 22

subjects, respectively. Though an improvement on single subject templates, these atlases used substantially fewer subjects than those described here. Finally, these dedicated atlases represent the anatomy from a small, limited age range, whereas our atlases span ages from 4.5 to 43.5 years. Qualitatively, the atlas presented in Fig. 4 of Joshi et al. (2004) and that presented in Figs. 1 and 2 of Bhatia et al. (2007) appear to have slightly less detail in the cortex than the atlases presented here, perhaps due to the larger number and the (older) ages of subjects used to create our templates.

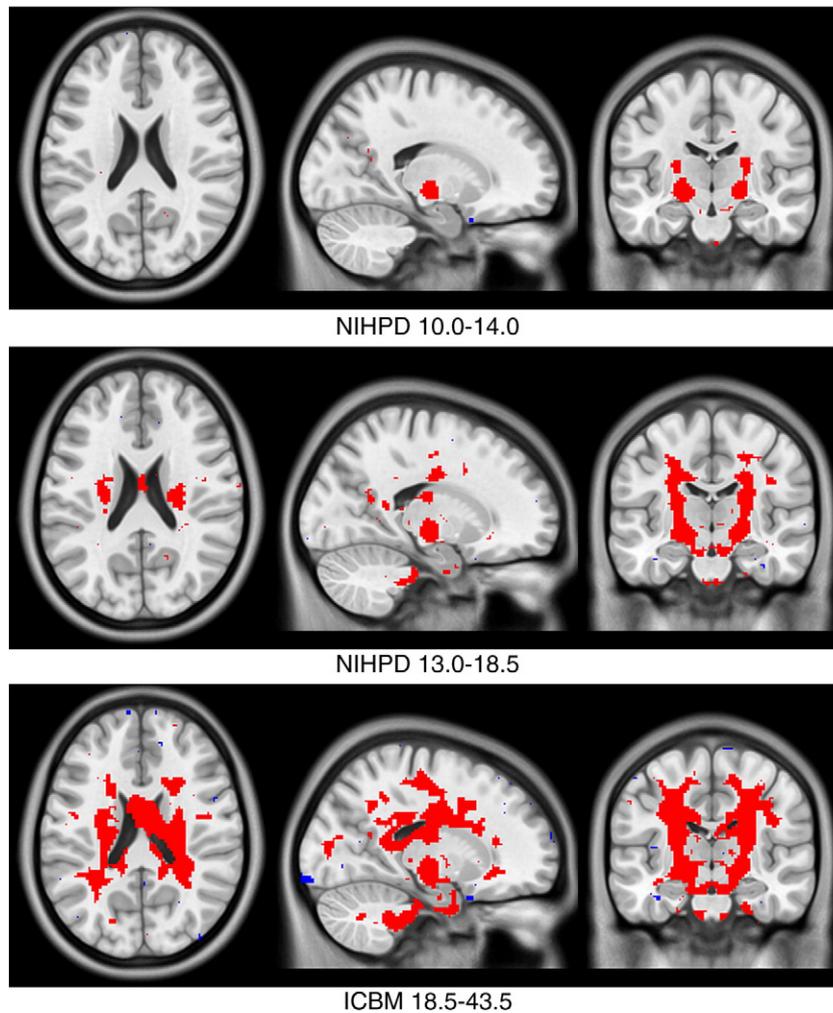
In conclusion, we have presented a method for unbiased atlas generation from large ensembles of MRI data. We have demonstrated that the iterative method converges and the resulting atlas templates maintain high anatomical detail throughout the brain. These publicly available templates are derived from a truly normal, well-characterized population and should facilitate spatial normalization and image analysis for better understanding of pediatric populations.

#### Disclaimer

The views herein do not necessarily represent the official views of the National Institute of Child Health and Human Development, the National Institute on Drug Abuse, the National Institute of Mental Health, the National Institute of Neurological Disorders and Stroke, the National Institutes of Health, the U.S. Department of Health and Human Services, or any other agency of the United States Government.



**Fig. 12.** Comparison between NIHDP 4.5–8.5 and ICBM 18.5–43.5 templates. When compared to the ICBM atlas, the NIHDP 4.5–8.5 atlas has thinner skull and scalp, narrower cortical sulci (A = Post Central Sulcus, B = Parieto-Occipital Sulcus, C = Calcarine Fissure), decreased separation of the cerebellar folia (D), thinner corpus callosum (E), smaller lateral ventricles (F), and thicker cortex overall. Internal architecture of the thalamus has a slightly different shape (G), Different shape of the pituitary gland (H), and the presence of the sphenoid-occipital synchondrosis (I), smaller pons (J).



**Fig. 13.** Regions of potential bias when using different atlases. Map of statistically significant differences in log Jacobians when mapping the NIHPD 4.5–6.9 age group to the NIHPD 7.0–11.0 (baseline for comparison) and the NIHPD 10.0–14.0 (top row), NIHPD 13.0–18.5 (middle row) and ICBM 18.5–43.5 (bottom row) templates, all presented in the space of the ICBM 18.5–45.0 template. Red color indicates regions where the selected templates produces significantly (5% False Discovery Rate (Genovese et al., 2002)) bigger log Jacobian determinant (i.e., a significant difference in local volume) compared to the NIHPD 7.0–11.0 template, and blue color indicates where the selected template yields a statistically significant smaller Jacobian determinant. One can see that the red regions are much larger than the blue regions, indicating potential bias non-age appropriate template for analysis of pediatric data in the 4.5–6.9 years range.

## Acknowledgments

This project has been funded in whole or in part with Federal funds from the National Institute of Child Health and Human Development, the National Institute of Drug Abuse, the National Institute of Mental Health, and the National Institute of Neurological Disorders and Stroke (Contract #s N01-HD02-3343, N01-MH9-0002, and N01-NS-9-2314, -2315, -2316, -2317, -2319 and -2320). Special thanks to the NIH contracting officers for their support. We also acknowledge the important contribution and remarkable spirit of John Haselgrove, Ph.D. (deceased).

## Appendix A. Brain Development Cooperative Group

The MRI Study of Normal Brain Development is a cooperative study performed by six pediatric study centers in collaboration with a Data Coordinating Center (DCC), a Clinical Coordinating Center (CCC), a Diffusion Tensor Processing Center (DPC), and staff of the National Institute of Child Health and Human Development (NICHD), the National Institute of Mental Health (NIMH), the National Institute for Drug Abuse (NIDA), and the National Institute for Neurological Diseases and Stroke (NINDS), Rockville, Maryland. Key personnel from the six pediatric study centers are as follows: Children's Hospital Medical

Center of Cincinnati, Principal Investigator William S. Ball, M.D., Investigators Anna Weber Byars, Ph.D., Mark Schapiro, M.D., Wendy Bommer, R.N., April Carr, B.S., April German, B.A., Scott Dunn, R.T.; Children's Hospital Boston, Principal Investigator Michael J. Rivkin, M.D., Investigators Deborah Waber, Ph.D., Robert Mulkern, Ph.D., Sridhar Vajapeyam, Ph.D., Abigail Chiverton, B.A., Peter Davis, B.S., Julie Koo, B.S., Jacki Marmor, M.A., Christine Mrakotsky, Ph.D., M.A., Richard Robertson, M.D., Gloria McAnulty, Ph.D.; University of Texas Health Science Center at Houston, Principal Investigators Michael E. Brandt, Ph.D., Jack M. Fletcher, Ph.D., Larry A. Kramer, M.D., Investigators Grace Yang, M.Ed., Cara McCormack, B.S., Kathleen M. Hebert, M.A., Hilda Volero, M.D.; Washington University in St. Louis, Principal Investigators Kelly Botteron, M.D., Robert C. McKinstry, M.D., Ph.D., Investigators William Warren, Tomoyuki Nishino, M.S., C. Robert Almlie, Ph.D., Richard Todd, Ph.D., M.D., John Constantino, M.D.; University of California Los Angeles, Principal Investigator James T. McCracken, M.D., Investigators Jennifer Levitt, M.D., Jeffrey Alger, Ph.D., Joseph O'Neil, Ph.D., Arthur Toga, Ph.D., Robert Asarnow, Ph.D., David Fadale, B.A., Laura Heinichen, B.A., Cedric Ireland B.A.; Children's Hospital of Philadelphia, Principal Investigators Dah-Jyuu Wang, Ph.D. and Edward Moss, Ph.D., Investigator Robert A. Zimmerman, M.D., and Research Staff Brooke Bintliff, B.S., Ruth Bradford, Janice Newman, M.B.A. The Principal Investigator of the data coordinating center at McGill University is Alan C. Evans, Ph.D., Investigators

Rozalia Arnaoutelis, B.S., G. Bruce Pike, Ph.D., D. Louis Collins, Ph.D., Gabriel Leonard, Ph.D., Tomas Paus, M.D., Alex Zijdenbos, Ph.D., and Research Staff Samir Das, B.S., Vladimir Fonov, Ph.D., Luke Fu, B.S., Jonathan Harlap, Ilana Leppert, B.E., Denise Milovan, M.A., Dario Vins, B.C., and at Georgetown University, Thomas Zeffiro, M.D., Ph.D. and John Van Meter, Ph.D. Investigators at the Neurostatistics Laboratory, Harvard University/McLean Hospital, Nicholas Lange, Sc.D. and Michael P. Froimowitz, M.S., work with data coordinating center staff and all other team members on biostatistical study design and data analyses. The Principal Investigator of the Clinical Coordinating Center at Washington University is Kelly Botteron, M.D., Investigators C. Robert Almli, Ph.D., Cheryl Rainey, B.S., Stan Henderson, M.S., Tomoyuki Nishino, M.S., William Warren, Jennifer L. Edwards, M.S.W., Diane Dubois, R.N., Karla Smith, Tish Singer and Aaron A. Wilber, M.S. The Principal Investigator of the Diffusion Tensor Processing Center at the National Institutes of Health is Carlo Pierpaoli, M.D., Ph.D., Investigators Peter J. Basser, Ph.D., Lin-Ching Chang, Sc.D., Chen Guan Koay, Ph.D. and Lindsay Walker, M.S. The Principal Collaborators at the National Institutes of Health are Lisa Freund, Ph.D. (NICHD), Judith Rumsey, Ph.D. (NIMH), Lauren Baskir, Ph.D. (NIMH), Laurence Stanford, Ph.D. (NIDA), Karen Sirocco, Ph.D. (NIDA) and from NINDS, Katrina Gwinn-Hardy, M.D. and Giovanna Spinella, M.D. The Principal Investigator of the Spectroscopy Processing Center at the University of California Los Angeles is James T. McCracken, M.D., Investigators Jeffrey R. Alger, Ph.D., Jennifer Levitt, M.D., Joseph O'Neill, Ph.D.

## References

- Almli, C.R., Rivkin, M.J., et al., 2007. The NIH MRI study of normal brain development (Objective-2): newborns, infants, toddlers, and preschoolers. *Neuroimage* 35 (1), 308–325.
- Amit, Y., Grenander, U., et al., 1991. Structural image restoration through deformable templates. *J. Am. Stat. Assoc.* 86 (414), 376–387.
- Ardekani, B.A., Guckemus, S., et al., 2005. Quantitative comparison of algorithms for inter-subject registration of 3D volumetric brain MRI scans. *J. Neurosci. Methods* 142 (1), 67–76.
- Arsigny, V., Commowick, O., et al., 2006. A Log-Euclidean framework for statistics on diffeomorphisms. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006*, pp. 924–931.
- Ashburner, J., Friston, K., 1997. Multimodal image coregistration and partitioning—a unified framework. *Neuroimage* 6 (3), 209–217.
- Avants, B., Grossman, M., et al., 2006. Symmetric diffeomorphic image registration: evaluating automated labeling of elderly and neurodegenerative cortex and frontal lobe. *Biomedical Image Registration*, pp. 50–57.
- Avants, B.B., Epstein, C.L., et al., 2008. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med. Image Anal.* 12 (1), 26–41.
- Bartzokis, G., Beckson, M., et al., 2001. Age-related changes in frontal and temporal lobe volumes in men: a magnetic resonance imaging study. *Arch. Gen. Psychiatry* 58 (5), 461–465.
- Bhatia, K.K., Hajnal, J.V., et al., 2004. Consistent groupwise non-rigid registration for atlas construction. *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on*.
- Bhatia, K., Aljabar, P., et al., 2007. Groupwise combined segmentation and registration for atlas construction. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2007*, pp. 532–540.
- Blanton, R.E., Levitt, J.G., et al., 2001. Mapping cortical asymmetry and complexity patterns in normal children. *Psychiatry Res.* 107 (1), 29–43.
- Blanton, R.E., Levitt, J.G., et al., 2004. Gender differences in the left inferior frontal gyrus in normal children. *Neuroimage* 22 (2), 626–636.
- Blatter, D.D., Bigler, E.D., et al., 1995. Quantitative volumetric analysis of brain MR: normative database spanning 5 decades of life. *AJNR Am. J. Neuroradiol.* 16 (2), 241–251.
- Carmichael, O.T., Aizenstein, H.A., et al., 2005. Atlas-based hippocampus segmentation in Alzheimer's disease and mild cognitive impairment. *Neuroimage* 27 (4), 979–990.
- Caviness Jr., V.S., Kennedy, D.N., et al., 1996. The human brain age 7–11 years: a volumetric analysis based on magnetic resonance images. *Cereb. Cortex* 6 (5), 726–736.
- Caviness Jr., V.S., Lange, N.T., et al., 1999. MRI-based brain volumetrics: emergence of a developmental brain science. *Brain Dev.* 21 (5), 289–295.
- Christensen, G.E., Rabbitt, R.D., et al., 1994. 3D brain mapping using a deformable neuroanatomy. *Phys. Med. Biol.* 39 (3), 609–618.
- Chung, M.K., Worsley, K.J., et al., 2001. A unified statistical approach to deformation-based morphometry. *Neuroimage* 14 (3), 595–606.
- Collins, D.L., Neelin, P., et al., 1994. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J. Comput. Assist. Tomogr.* 18 (2), 192–205.
- Collins, D.L., Zijdenbos, A.P., et al., 1999. ANIMAL+INSECT: improved cortical structure segmentation. *Information Processing in Medical Imaging: 16th International Conference, IPMI'99, Visegrád, Hungary, June/July 1999. Proceedings*: 210.
- Courchesne, E., Chisum, H.J., et al., 2000. Normal brain development and aging: quantitative analysis at in vivo MR imaging in healthy volunteers. *Radiology* 216 (3), 672–682.
- De Bellis, M.D., Keshavan, M.S., et al., 2001. Sex differences in brain maturation during childhood and adolescence. *Cereb. Cortex* 11 (6), 552–557.
- Durston, S., Hulshoff Pol, H.E., et al., 2001. Anatomical MRI of the developing human brain: what have we learned? *J. Am. Acad. Child Adolesc. Psychiatry* 40 (9), 1012–1020.
- Essen, V., David, C., 2002. Windows on the brain: the emerging role of atlases and databases in neuroscience. *Curr. Opin. Neurobiol.* 12 (5), 574–579.
- Essen, V., David, C., 2005. A Population-Average, Landmark- and Surface-based (PALS) atlas of human cerebral cortex. *Neuroimage* 28 (3), 635–662.
- Evans, A.C., B.D.C. Group, 2006. The NIH MRI study of normal brain development. *Neuroimage* 30 (1), 184–202.
- Evans, A.C., Collins, D.L., et al., 1993. 3D statistical neuroanatomical models from 305 MRI volumes. *Nuclear Science Symposium and Medical Imaging Conference, 1993., 1993 IEEE Conference Record*.
- Filipek, P.A., Richelme, C., et al., 1994. The young adult human brain: an MRI-based morphometric analysis. *Cereb. Cortex* 4 (4), 344–360.
- Genovese, C.R., Lazar, N.A., et al., 2002. Thresholding of Statistical Maps in Functional Neuroimaging Using the False Discovery Rate. *Neuroimage* 15 (4), 870–878.
- Gerig, G., Davis, B., et al., 2006. Computational anatomy to assess longitudinal trajectory of brain growth. *3D Data Processing, Visualization, and Transmission, Third International Symposium on*.
- Giedd, J.N., Rumsey, J.M., et al., 1996a. A quantitative MRI study of the corpus callosum in children and adolescents. *Brain Res. Dev. Brain Res.* 91 (2), 274–280.
- Giedd, J.N., Vaituzis, A.C., et al., 1996b. Quantitative MRI of the temporal lobe, amygdala, and hippocampus in normal human development: ages 4–18 years. *J. Comp. Neurol.* 366 (2), 223–230.
- Giedd, J.N., Blumenthal, J., et al., 1999. Brain development during childhood and adolescence: a longitudinal MRI study. *Nat. Neurosci.* 2 (10), 861–863.
- Gogtay, N., Giedd, J., et al., 2002. Brain development in healthy, hyperactive, and psychotic children. *Arch. Neurol.* 59 (8), 1244–1248.
- Gogtay, N., Giedd, J.N., et al., 2004. Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl. Acad. Sci. USA* 101 (21), 8174–8179.
- Good, C.D., Johnsrude, I.S., et al., 2001. A voxel-based morphometric study of ageing in 465 normal adult human brains. *Neuroimage* 14 (1 Pt 1), 21–36.
- Grabner, G., Janke, A.L., et al., 2006. Symmetric atlas and model based segmentation: an application to the hippocampus in older adults. *Med. Image Comput. Assist. Interv. Int. Conf. Med. Image Comput. Assist. Interv.* 9 (Pt 2), 58–66.
- Guimond, A., Meunier, J., et al., 1998. Automatic computation of average brain models. *Medical Image Computing and Computer-Assisted Intervention—MICCAI'98*, p. 631.
- Guimond, A., Meunier, J., et al., 2000. Average brain models: a convergence study. *Comput. Vis. Image Underst.* 77, 192–210.
- Guimond, A., Roche, A., et al., 2001. Three-dimensional multimodal brain warping using the demons algorithm and adaptive intensity corrections. *IEEE Trans. Med. Imaging* 20 (1), 58–69.
- Guizzard, N., Fonov, V., et al., 2009. Symmetric optimization scheme versus constrained symmetrization for non-linear registrations. *MICCAI workshop on "Medical Image Analysis on Multiple Sclerosis (validation and methodological issues)"*.
- Hoeksma, M.R., Kenemans, J.L., et al., 2005. Variability in spatial normalization of pediatric and adult brain images. *Clin. Neurophysiol.* 116 (5), 1188–1194.
- Janke, A., Evans, A., et al., 2006. MNI- and Talairach- space: everything you wanted to know but were afraid to ask. HBM, Florence, Italy.
- Jelacic, S., de Regt, D., et al., 2006. Interactive digital MR Atlas of the pediatric brain. *Radiographics* 26 (2), 497–501.
- Jernigan, T.L., Tallal, P., 1990. Late childhood changes in brain morphology observable with MRI. *Dev. Med. Child Neurol.* 32 (5), 379–385.
- Jernigan, T.L., Trauner, D.A., et al., 1991. Maturation of human cerebrum observed in vivo during adolescence. *Brain* 114 (Pt 5), 2037–2049.
- Joshi, S.C., Miller, M.L., 2000. Landmark matching via large deformation diffeomorphisms. *IEEE Trans. Image Process.* 9 (8), 1357–1370.
- Joshi, S., Davis, B., et al., 2004. Unbiased diffeomorphic atlas construction for computational anatomy. *Neuroimage* 23 (Suppl 1), S151–S160.
- Kazemi, K., Moghaddam, H.A., et al., 2007. A neonatal atlas template for spatial normalization of whole-brain magnetic resonance images of newborns: preliminary results. *Neuroimage* 37 (2), 463–473.
- Kennedy, D.N., Lange, N., et al., 1998. Gyri of the human neocortex: an MRI-based analysis of volume and variance. *Cereb. Cortex* 8 (4), 372–384.
- Kennedy, D.N., Haselgrove, C., et al., 2003. MRI-based morphometric of typical and atypical brain development. *Ment. Retard. Dev. Disabil. Res.* 9 (3), 155–160.
- Klein, S., Andersson, J., et al., 2009. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage* 46 (3), 786–802.
- Kochunov, P., Lancaster, J.L., et al., 2001. Regional spatial normalization: toward an optimal target. *J. Comput. Assist. Tomogr.* 25 (5), 805–816.
- Lange, N., Giedd, J.N., et al., 1997. Variability of human brain structure size: ages 4–20 years. *Psychiatry Res.* 74 (1), 1–12.
- Lorenzen, P. J. and S. C. Joshi (2003). High-dimensional multi-modal image registration. 2717: 234–243.
- Lorenzen, P., Davis, B., et al., 2005. Unbiased atlas formation via large deformations metric mapping. *Med. Image Comput. Assist. Interv. Int. Conf. Med. Image Comput. Assist. Interv.* 8 (Pt 2), 411–418.
- Louis Collins, D., Holmes, C.J., Peters, T.M., Evans, A.C., 1995. Automatic 3-D model-based neuroanatomical segmentation. *Human Brain Mapping* 3 (3), 190–208. <http://dx.doi.org/10.1002/hbm.460030304>.

- Machlisen, B., d'Agostino, E., et al., 2007. Linear normalization of MR brain images in pediatric patients with periventricular leukomalacia. *Neuroimage* 35 (2), 686–697.
- Mazziotta, J.C., Toga, A.W., et al., 1995. A probabilistic atlas of the human brain: theory and rationale for its development. The International Consortium for Brain Mapping (ICBM). *Neuroimage* 2 (2), 89–101.
- Mazziotta, J., Toga, A., et al., 2001a. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philos. Trans. R. Soc. B Biol. Sci.* 356 (1412), 1293–1322.
- Mazziotta, J., Toga, A., et al., 2001b. A four-dimensional probabilistic atlas of the human brain. *J. Am. Med. Inform. Assoc.* 8 (5), 401–430.
- Miller, M., Banerjee, A., et al., 1997. Statistical methods in computational anatomy. *Stat. Methods Med. Res.* 6 (3), 267–299.
- Muzik, O., Chugani, D.C., et al., 2000. Statistical parametric mapping: assessment of application in children. *Neuroimage* 538–549.
- Nyul, L.G., Udupa, J.K., 1999. On standardizing the MR image intensity scale. *Magn. Reson. Med.* 42 (6), 1072–1081.
- Park, H., Bland, P.H., et al., 2005. Least biased target selection in probabilistic atlas construction. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2005*, pp. 419–426.
- Paus, T., Zijdenbos, A., et al., 1999. Structural maturation of neural pathways in children and adolescents: in vivo study. *Science* 283 (5409), 1908–1911.
- Petersen, A.C., Crockett, L., et al., 1988. A self-report measure of pubertal status: reliability, validity, and initial norms. *J. Youth Adolesc.* 17 (2), 117–133.
- Pfefferbaum, A., Mathalon, D.H., et al., 1994. A quantitative magnetic resonance imaging study of changes in brain morphology from infancy to late adulthood. *Arch. Neurol.* 51 (9), 874–887.
- Reiss, A.L., Abrams, M.T., et al., 1996. Brain development, gender and IQ in children. A volumetric imaging study. *Brain* 119 (Pt 5), 1763–1774.
- Seghers, D., D'Agostino, E., et al., 2004. Construction of a brain template from MR images using state-of-the-art registration and segmentation techniques. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2004*, pp. 696–703.
- Shan, Z., Parra, C., et al., 2006. A digital pediatric brain structure atlas from T1-weighted MR images. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2006*, pp. 332–339.
- Shattuck, D.W., Mirza, M., et al., 2008. Construction of a 3D probabilistic atlas of human cortical structures. *Neuroimage* 39 (3), 1064–1080.
- Sled, J.G., Zijdenbos, A.P., et al., 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans. Med. Imaging* 17 (1), 87–97.
- Smith, S.M., 2002. Fast robust automated brain extraction. *Hum. Brain Mapp.* 17 (3), 143–155.
- Smith, S.M., Jenkinson, M., et al., 2004. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23, S208–S219.
- Sowell, E.R., Thompson, P.M., et al., 1999. Localizing age-related changes in brain structure between childhood and adolescence using statistical parametric mapping. *Neuroimage* 9 (6 Pt 1), 587–597.
- Sowell, E.R., Thompson, P.M., et al., 2002. Mapping sulcal pattern asymmetry and local cortical surface gray matter distribution in vivo: maturation in perisylvian cortices. *Cereb. Cortex* 12 (1), 17–26.
- Sowell, E.R., Peterson, B.S., et al., 2003. Mapping cortical change across the human life span. *Nat. Neurosci.* 6 (3), 309–315.
- Sowell, E.R., Thompson, P.M., et al., 2004a. Longitudinal mapping of cortical thickness and brain growth in normal children. *J. Neurosci.* 24 (38), 8223–8231.
- Sowell, E.R., Thompson, P.M., et al., 2004b. Mapping changes in the human cortex throughout the span of life. *Neuroscientist* 10 (4), 372–392.
- Thevenaz, P., Blu, T., et al., 2000. Interpolation revisited [medical images application]. *IEEE Trans. Med. Imaging* 19 (7), 739–758.
- Thévenaz, P., Blu, T., et al., 2000. Image Interpolation and Resampling. In: Bankman, I.N. (Ed.), *Handbook of Medical Imaging, Processing and Analysis*. Academic Press, San Diego, CA USA, pp. 393–420.
- Thompson, P.M., Toga, A.W., 2002. A framework for computational anatomy. *Comput. Vis. Sci.* 5 (1), 13–34.
- Toga, A.W., Thompson, P.M., 2001. Maps of the brain. *Anat. Rec.* 265 (2), 37–53.
- Toga, A.W., Thompson, P.M., 2003. Mapping brain asymmetry. *Nat. Rev. Neurosci.* 4 (1), 37–48.
- Toga, A.W., Thompson, P.M., 2007. What is where and why it is important. *Neuroimage* 37 (4), 1045–1049.
- Tohka, J., Zijdenbos, A., et al., 2004. Fast and robust parameter estimation for statistical partial volume models in brain MRI. *Neuroimage* 23 (1), 84–97.
- Tohka, J., Krestyannikov, E., et al., 2007. Genetic algorithms for finite mixture model based voxel classification in neuroimaging. *IEEE Trans. Med. Imaging* 26 (5), 696–711.
- Waber, D.P., De Moor, C., et al., 2007. The NIH MRI study of normal brain development: performance of a population based sample of healthy children aged 6 to 18 years on a neuropsychological battery. *J. Int. Neuropsychol. Soc.* 13 (05), 729–746.
- Wang, Q., Seghers, D., et al., 2005. Construction and Validation of Mean Shape Atlas Templates for Atlas-Based Brain Image Segmentation. *Information Processing in Medical Imaging*, pp. 689–700.
- Wilke, M., Schmithorst, V.J., et al., 2002a. Assessment of spatial normalization of whole-brain magnetic resonance images in children. *Hum. Brain Mapp.* 17 (1), 48–60.
- Wilke, M., Schmithorst, V.J., Holland, S.K., 2002b. Assessment of spatial normalization of whole-brain magnetic resonance images in children. *Hum. Brain Mapp.* 17 (1), 48–60.
- Wilke, M., Schmithorst, V.J., Holland, S.K., 2003. Normative pediatric brain data for spatial normalization and segmentation differs from standard adult data. *Magn. Reson. Med.* 50 (4), 749–757.
- Wilke, M., Holland, S.K., et al., 2008. Template-O-Matic: A toolbox for creating customized pediatric templates. *Neuroimage* 41 (3), 903–913.
- Woods, R.P., Grafton, S.T., et al., 1998. Automated image registration: I. General methods and intrasubject, intramodality validation. *J. Comput. Assist. Tomogr.* 22 (1), 139–152.
- Wu, M., Rosano, C., et al., 2007. Optimum template selection for atlas-based segmentation. *Neuroimage* 34 (4), 1612–1618.