

# Magnetic Resonance Imaging of the Newborn Brain: Automatic Segmentation of Brain Images into 50 Anatomical Regions

Ioannis S. Gousias<sup>1,2\*</sup>, Alexander Hammers<sup>1,3</sup>, Serena J. Counsell<sup>1,4</sup>, Latha Srinivasan<sup>1</sup>, Mary A. Rutherford<sup>1,4</sup>, Rolf A. Heckemann<sup>1,3</sup>, Jo V. Hajnal<sup>1,4</sup>, Daniel Rueckert<sup>2</sup>, A. David Edwards<sup>1,4</sup>

**1** Faculty of Medicine, Imperial College London, and Medical Research Council Clinical Sciences Centre, Hammersmith Hospital, London, United Kingdom, **2** Department of Computing, Imperial College London, South Kensington, London, United Kingdom, **3** The Neurodis Foundation (Fondation Neurodis), CERMEP – Imagerie du Vivant, Lyon, France, **4** Centre for the Developing Brain, King's College London, London, United Kingdom

## Abstract

We studied methods for the automatic segmentation of neonatal and developing brain images into 50 anatomical regions, utilizing a new set of manually segmented magnetic resonance (MR) images from 5 term-born and 15 preterm infants imaged at term corrected age called ALBERTs. Two methods were compared: individual registrations with label propagation and fusion; and template based registration with propagation of a maximum probability neonatal ALBERT (MPNA). In both cases we evaluated the performance of different neonatal atlases and MPNA, and the approaches were compared with the manual segmentations by means of the Dice overlap coefficient. Dice values, averaged across regions, were  $0.81 \pm 0.02$  using label propagation and fusion for the preterm population, and  $0.81 \pm 0.02$  using the single registration of a MPNA for the term population. Segmentations of 36 further unsegmented target images of developing brains yielded visibly high-quality results. This registration approach allows the rapid construction of automatically labeled age-specific brain atlases for neonates and the developing brain.

**Citation:** Gousias IS, Hammers A, Counsell SJ, Srinivasan L, Rutherford MA, et al. (2013) Magnetic Resonance Imaging of the Newborn Brain: Automatic Segmentation of Brain Images into 50 Anatomical Regions. PLoS ONE 8(4): e59990. doi:10.1371/journal.pone.0059990

**Editor:** Antoni Rodriguez-Fornells, University of Barcelona, Spain

**Received:** June 19, 2012; **Accepted:** February 22, 2013; **Published:** April 2, 2013

**Copyright:** © 2013 Gousias et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** ISG was supported by the United Kingdom Engineering and Physical Sciences Research Council, Action Medical Research United Kingdom and the Henry Smith Charity. He was also funded by a research scholarship from the "General Arnaoutis" Foundation. MR image data was collected with financial support from the Medical Research Council. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: ioannis.gousias04@imperial.ac.uk

## Introduction

Anatomical structures can be segmented in biological images by transfer of voxel labels from an analogous image previously segmented into anatomical regions, or atlas [1]. This requires an accurate alignment and correspondence of structurally equivalent regions between the atlas and the target image usually achieved using non-rigid registration [2,3,4,5,6]. Segmentation methods often first register the atlas to the target image and then segment the target image into anatomical structures based on transferred information [5,7,8,9,10,11,12], although registering multiple atlases to the same target with subsequent fusion of different segmentations will frequently improve the final segmentation result, compensating for nonsystematic errors in single registrations [13,14,15,16,17].

Magnetic Resonance (MR) images of the brains of newborn infants have been particularly difficult to segment [18] due to: the low tissue contrast; signal inhomogeneity; intersubject differences due to the rapid development of the brain, especially in white matter (WM) structures [19]. The development of automatic segmentation methods has been further hindered by the lack of gold standard data for comparison and validation [20].

In this paper we present two methods for the automatic segmentation of neonatal and developing brain MR images into

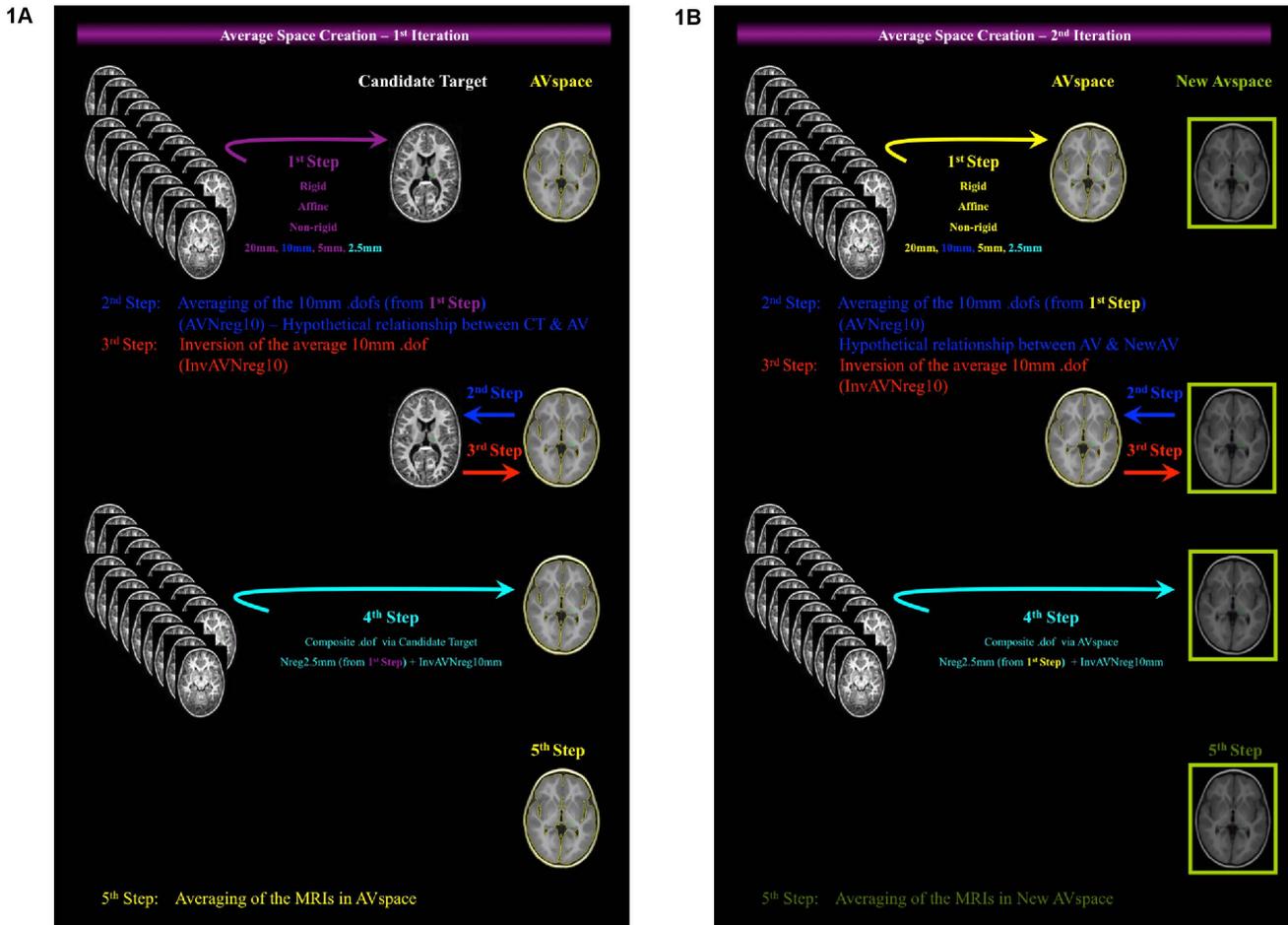
50 regions of interest (ROI) utilizing a new set of manually defined neonatal atlases called ALBERTs [21]. The first approach is based on fusion of anatomical prior information from various neonatal atlases. The second approach is based on propagation of labels from a maximum probability neonatal ALBERT (MPNA). For both methods we evaluated the performance of different atlases and MPNAs and compared the results to the gold-standard manual segmentations.

## Materials and Methods

### Image Acquisition

MR images were acquired using a 3.0 Tesla Philips Achieva scanner (Philips Medical Systems, Best, The Netherlands). The technical characteristics of the scans, as well as detailed demographics, can be found in our previous work [21]. T1-weighted magnetization prepared rapid-acquisition gradient echo volumes in the sagittal plane were acquired with an echo time of 4.6 ms and repetition time 17 ms; 124–150 sagittal slices of 1.6 mm thickness were acquired with a 210 mm field of view, a flip angle of 30°, and a 256×256 matrix, resulting in voxel sizes of  $0.82 \times 0.82 \times 1.6 \text{ mm}^3$ . We used the following data sets:

- 15 preterm neonates scanned at term (eight female), with a median gestational age at birth of 29 weeks (range 26–35



**Figure 1. Illustration of the algorithm for the creation of an average space/template, emphasizing schematically on the different spaces involved in the transformations.** **1a** First Iteration: 1st Step: Non-rigid registration of cohort to the candidate target; 2nd Step: Averaging of the nonrigid 10 mm transformation (blue arrow); 3rd Step: Inversion of the average nonrigid 10 mm transformation (red arrow); 4th Step: Composite transformation (nonrigid 2.5 mm+inverted average 10 mm); 5th Step: Averaging of the MRIs in average (AV) space. **1b.** Second Iteration: 1st Step: Non-rigid registration of cohort to the new candidate (AV space); 2nd Step: Averaging of the nonrigid 10 mm transformation (blue arrow); 3rd Step: Inversion of the average nonrigid 10 mm transformation (red arrow); 4th Step: Composite transformation (nonrigid 2.5 mm+inverted average 10 mm); 5th Step: Averaging of the MRIs in New AV space. doi:10.1371/journal.pone.0059990.g001

weeks) and a median gestational age at the time of the scan of the subjects of 40 (37–43) weeks, manually segmented [21]. Two were twins with a gestational age of 29 weeks, scanned at 40 weeks.

- Five term control neonates (two female), with a median age at scan of 41 (39–45) weeks, manually segmented [21].
- 36 preterm neonates scanned at birth that had not been manually segmented (sixteen female), with a median gestational age at birth of 29 (24–36) weeks.

Approval for scanning the subjects had been obtained from the Hammersmith Hospital Research Ethics Committee, and written informed consent obtained prior to scanning. Post-processing of anonymised scan data that had been acquired for clinical purposes did not require individual consent from the individuals who had been scanned.

**MR Image Pre-processing**

T1-weighted 3D image volumes were obtained in DICOM format and converted to the NIFTI format using the UCLA

Laboratory of Neuro-Imaging’s Debabeler Software ([www.loni.ucla.edu/Software/Debabeler](http://www.loni.ucla.edu/Software/Debabeler)). The image matrix was reduced in superior, anterior, posterior, and lateral directions to contain five empty slices ( $5 \times 0.82 = 4.1$  mm) after the last slice containing skin. Inferiorly, five empty slices ( $5 \times 1.6 = 8$  mm) were added after the last slice where the posterior floor of the skull was visible. The reduction in matrix size simplified the subsequent bias correction step in that inferior extracranial signal would not have to be

**Table 1. Neonatal Templates based on different candidate target and different prior MRIs.**

Template	Candidate Target	Fused information
Template_01	Term-born	All 19 remaining MRIs.
Template_02	Term-born	All 15 preterm MRIs.
Template_03	Preterm	All 14 remaining preterm MRIs.
Template_04	Term-born	All 4 remaining term-born MRIs.

doi:10.1371/journal.pone.0059990.t001

considered. The padding around the skull was maintained because our previous work had shown the skull to be an essential landmark for successful registration in young children [16,17]. T1-weighted image volumes were corrected for non-uniformity using the FAST Software from the FMRIB Software Library (FSL version 4, [22]). T1-weighted images were resampled, creating isotropic voxels of  $0.82 \times 0.82 \times 0.82 \text{ mm}^3$  using windowed sinc interpolation, to make them compatible with Analyze AVW 8.1 software. Re-orientation of the sagittal T1 volumes was performed with the horizontal line defined by the anterior and posterior commissures (AC-PC orientation) and the sagittal planes parallel to the midline [23]. We reduced the number of interpolation steps during reorientation by coregistration of the native images onto the re-orientated versions of themselves. This was performed using a method based on normalised mutual information and 7<sup>th</sup> degree B-spline interpolation. Coregistration was performed using SPM2 (Statistical Parametric Mapping, Wellcome Trust Centre for Neuroimaging, UCL, London) [24,25] under Matlab version 6.5 (Mathworks Inc, Sherborn, MA, USA). A 16-bit voxel depth was maintained throughout the process.

**Delineation, Manual Segmentation, and Nomenclature**

The MR images had been manually segmented into 50 ROIs each, covering the whole brain, using newly created protocols established according to previously described principles [26,27,28]. Each voxel belongs to one ROI only, and their ensemble thus constitutes a brain atlas: a label-based encephalic ROI template (ALBERT), as described in detail in Gousias et al., 2012 [21]. The 60-pages illustrated Appendix of the aforementioned companion paper consists of the protocols for all the regions.

In the remainder, the brain atlases are designated as ALBERTs; ALBERTs having been registered onto other individual brain MRIs via their underlying MRI and MRI-to-MRI registrations as “transformed ALBERTs”; average greyscale MRIs as “Templates”; fused atlases in a template space (i.e. in an average space as opposed to an individual space) as maximum probability neonatal ALBERT (MPNA).

**Automatic Segmentation via Multiple ALBERTs**

**Registration.** After pre-processing, every neonatal subject was paired with every other neonatal subject for image registration, resulting in 380 (20×19) image pairs. All pairs were aligned using 3D voxel-based registration in three steps using IRTK Software (available via <http://wwwhomes.doc.ic.ac.uk/~dr/software/>): rigid, affine and non-rigid registration. Parameter settings were tuned to the specific challenges posed by images of

neonates [17]. Blurring of both target and source images during the subsequent affine registration improved results. Furthermore, we increased the resolution levels from one to three, increased the number of iterations from 100 to 200, decreased the length of steps from 3.2 to 2 and used the correlation coefficient as the similarity measure in this step. For the final non-rigid registration, iterations were increased from 10 to 100 compared with adult-to-adult registration. The non-rigid step was based on the manipulation of a free-form deformation represented by displacements on a grid of control points blended using cubic B-splines [29] (available via <http://wwwhomes.doc.ic.ac.uk/~dr/software/>) and maximizing normalised mutual information (NMI; [30]). The registration was refined in a multi-resolution fashion by stepwise reduction of the control point spacing from 20 mm to 10 mm, 5 mm and finally 2.5 mm. Registration pairs were processed in parallel on a cluster of approximately 400 Linux PCs, controlled by Condor software (Version 6.7.13, <http://www.cs.wisc.edu/condor/>).

**Label propagation.** The output of the registration of an image pair is a transformation that maps the neonatal source image to the neonatal target image. These transformations were then applied to the ALBERTs using nearest-neighbour interpolation, resulting in 19 individualized propagated transformed ALBERTs for each of the 20 target brains.

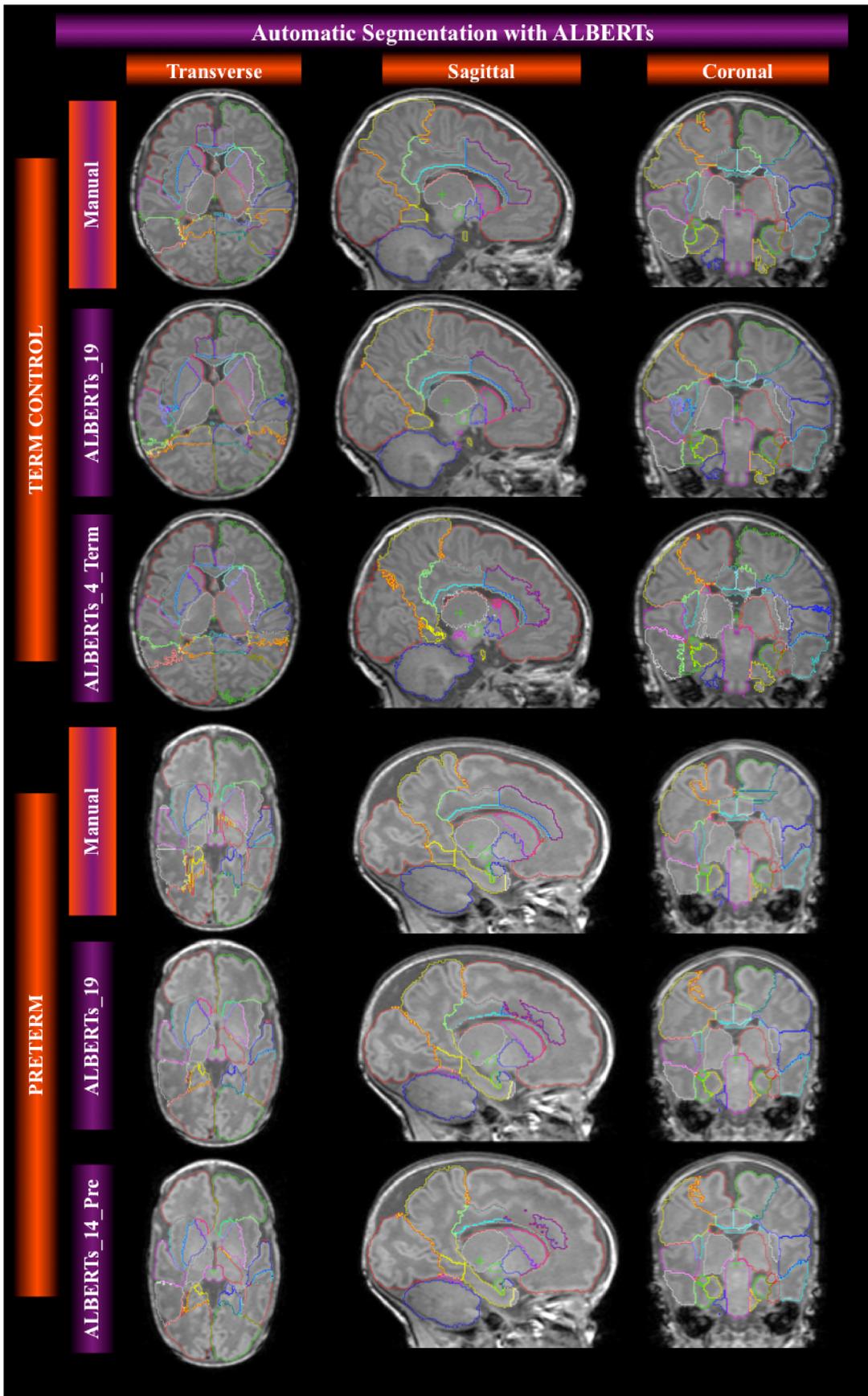
**Decision fusion.** Each resulting transformed ALBERT assigns a structure label to every voxel in the corresponding MR image volume. To combine the information from multiple individual label sets into a single segmentation, we applied vote-rule decision fusion. The consensus class of each voxel was defined as the modal value of the distribution of the individual label assignments [31]. This approach yielded good results in our previous studies [14,16,17,26]). In the case of non-unique modes, one of the modal values was assigned at random. Even vs. odd numbers of individual label sets resulted in twice the number of equivocal voxels, but in absolute terms, the fraction was very small (less than 1% of the total number of voxels) [16]. Three versions were created.

First, we created fused atlases for all 20 subjects based on fusion of all remaining 19 transformed ALBERTs (ALBERTs<sub>19</sub>). Secondly, for the 15 preterms, we also created fused atlases based on only the remaining 14 preterm transformed ALBERTs (ALBERTs<sub>14\_Pre</sub>); and finally for the term population (n = 5), we also created fused atlases based on only the remaining four term transformed ALBERTs (ALBERTs<sub>4\_Term</sub>).

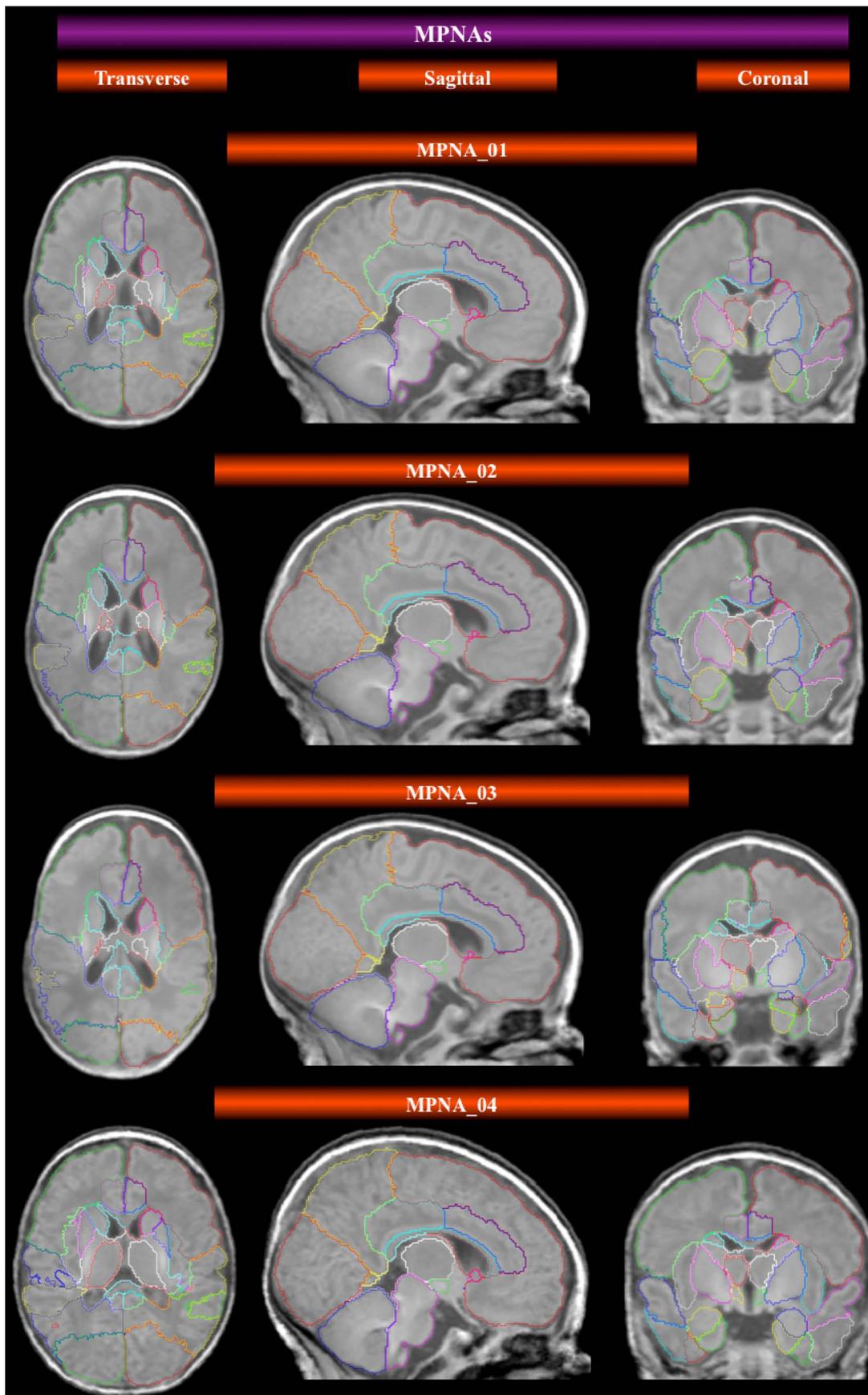
**Table 2.** MPNAs based on different templates and different ALBERTs.

MPNA	Template	Fused information in a leave-one-out fashion
MPNA_01	Template 01	19 remaining ALBERTs
MPNA_02	Template 02	19 remaining ALBERTs
MPNA_02_Preterms	Template 02	14 remaining preterm ALBERTs
MPNA_03	Template 03	19 remaining ALBERTs
MPNA_03_Preterms	Template 03	14 remaining preterm ALBERTs
MPNA_04	Template 04	19 remaining ALBERTs
MPNA_04_Preterms	Template 04	14 remaining preterm ALBERTs
MPNA_04_Terms	Template 04	4 remaining term ALBERTs

doi:10.1371/journal.pone.0059990.t002



**Figure 2. Automatic segmentation using different groups of ALBERTs and decision fusion.** Comparison with manual gold standard ALBERT. doi:10.1371/journal.pone.0059990.g002



**Figure 3. Average space templates and corresponding MPNAs.** All MPNAs shown here are derived from fusion of all remaining 19 transformed ALBERTs. Only the template creation differs in terms of the initial candidate target (see Figure 1): term-born for MPNA\_01 MPNA\_02 and MPNA\_04, preterm for 03; and in terms of the MRIs averaged to create the template space: all remaining 19 for MPNA\_01; all 15 preterms for MPNA\_02; all remaining 14 preterms for MPNA\_03, and all remaining 4 terms for MPNA\_04.  
doi:10.1371/journal.pone.0059990.g003

### Automatic Segmentation via Probabilistic Templates and MPNAs

**Neonatal template creation.** One of the term-born controls was selected as the candidate target. All remaining 19 data sets were registered to the candidate target with rigid, affine, and non-rigid registration starting with 20 mm spacing down to 2.5 mm as described above. The process is illustrated in Figure 1. The 10 mm non-rigid transformations were averaged and the average transformation was inverted, on the assumption that this average transformation maps the hypothetical average space we want to create to the candidate target space. Combining each 2.5 mm non-rigid transformation with the inverse average 10 mm non-rigid transformation we transferred each image, through the candidate target space, to the average space. In the second iteration (Figure 1b) we used the mean intensity image in average space as the new candidate target, in order to reduce possible bias arising from the choice of the first candidate target. We registered the 19 data sets to the new candidate average and repeated the steps twice. After the second iteration, we obtained the new average space. Similar approach has been used for the creation of pediatric templates [32].

In order to assess the influence of the bias resulting from the choice of the candidate target on subsequent MPNA registrations we created four different average spaces. For three of these, the visually most representative and symmetrical of the term controls

was selected as a candidate target and used to transform 1) the whole cohort, 2) the preterm data sets only, 3) the term data sets only. Finally, we selected the visually most representative and symmetrical preterm data set and transformed only the preterm data sets to this candidate target (Table 1).

After the creation of the average spaces for the cohort of neonates, each data set was registered to each average space using the same parameter settings as previously and the segmentations were fused to create eight different MPNAs (Table 2), each corresponding to one of the four template spaces (Table 1). This is similar to the creation of a maximum probability atlas for the pediatric population [32].

### Validation

Validation of all automatic segmentations was achieved via overlap measurements, expressed as a Dice index (twice the intersection divided by the union; [33,34]) between the automatically created segmentations and the corresponding manually created ALBERT, which served as the gold standard. Automatic segmentations were based on individual pairwise registrations and subsequent label fusion of:

- all 19 remaining manually created ALBERTs - ALBERTs\_19,
- 14 remaining preterms - ALBERTs\_14\_Pre,
- four remaining terms - ALBERTs\_4\_Term.

**Table 3. Validation results by means of SI measurements of the different templates, different ALBERTs of Optimum Segmentation and different fusion approaches.**

	Term-borns			Preterms		
	Average SI	Stdev	CV (%)	Average SI	Stdev	CV (%)
<b>MPNAs</b>						
MPNA_01	0.74	0.02	3	0.73	0.03	4
MPNA_02	0.73	0.02	3	0.72	0.02	3
MPNA_02_Preterms	0.70	0.02	3	0.72	0.02	3
MPNA_03	0.69	0.04	5	0.72	0.03	4
MPNA_03_Preterms	0.66	0.04	5	0.71	0.03	4
MPNA_04	0.75	0.02	3	<b>0.75<sup>††</sup></b>	0.03	4
MPNA_04_Preterms	0.72	0.02	3	0.74	0.03	4
MPNA_04_Terms	<b>0.81<sup>†</sup></b>	0.02	2	0.70	0.03	4
<b>Fusion of individual transformed ALBERTs</b>						
ALBERTs_19	<b>0.76*</b>	0.03	4	<b>0.81**</b>	0.02	3
ALBERTs_4	<b>0.79*</b>	0.02	3			
ALBERTs_14				<b>0.81**</b>	0.03	4
<b>Twin ALBERTs</b>						
ALBERTs_14				0.83	0.01	1
Twin ALBERT Single				0.80	0.002	0.32

SI: Similarity Index, same as Dice Index.

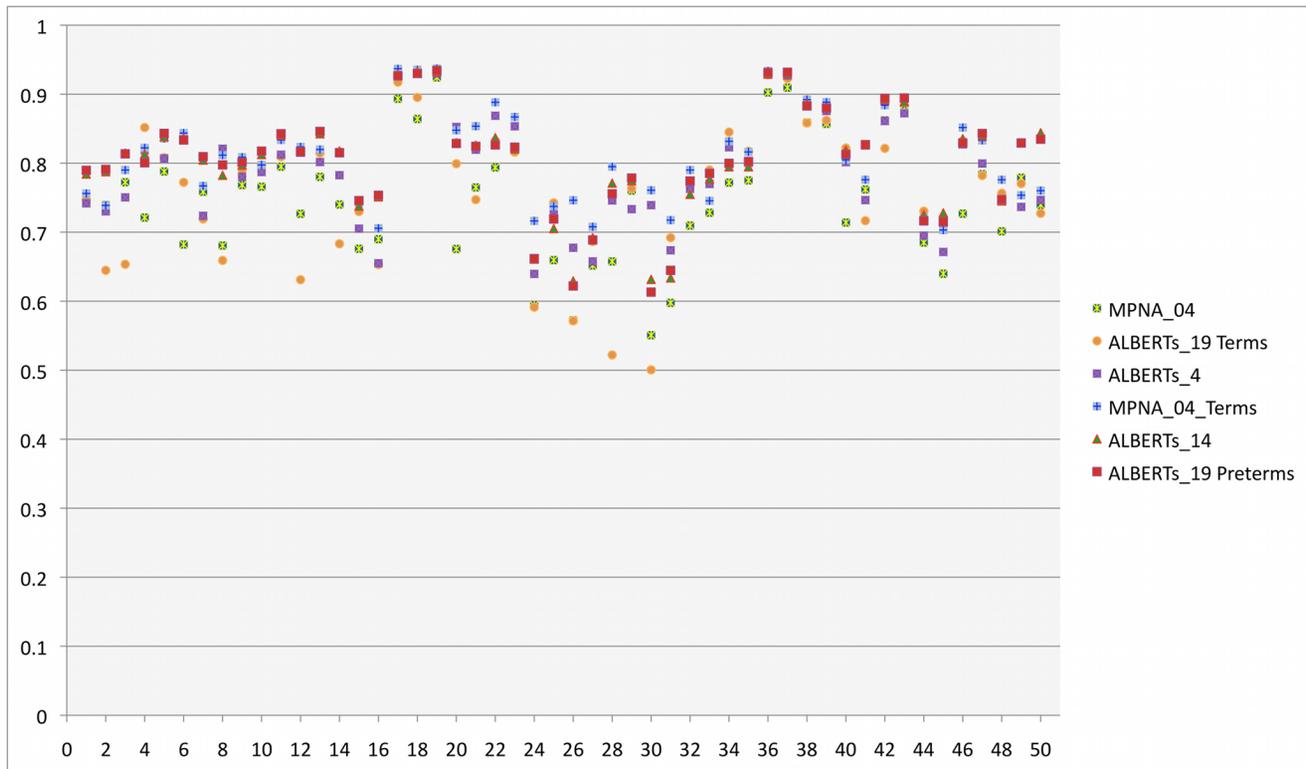
<sup>†</sup>Significant difference of MPNA\_04\_Terms with ALBERTs\_19 (two-tailed paired TTEST, p<0.001) and ALBERTs\_4 (two-tailed paired TTEST, p<0.05).

\*Not significant difference with each other.

<sup>††</sup>Significant difference of MPNA\_04 with ALBERTs\_19 and ALBERTs\_14 (two-tailed paired TTEST, p<1.0E-07).

\*\*Not significant difference with each other.

doi:10.1371/journal.pone.0059990.t003



**Figure 4. Validation of different approaches for automatic segmentation of the newborn brain.** Dice measurements for 50 ROIs, either fusing anatomical prior information from various combinations of ALBERTs or propagating labels of various MPNAs. Only the analytical results of the approaches that performed best are displayed. For translating the numbers into anatomical region names, see Table 4. doi:10.1371/journal.pone.0059990.g004

as well as back-registration of the eight MPNAs in the four average template spaces. We also tested the performance of twin pair ALBERT after single label propagation on the corresponding twin brain, compared to the fusion of non-corresponding priors.

We compared the performance of the best methods for each group using two-tailed paired TTEST, after Bonferroni correction for multiple comparisons, with regards to each ROI. For the preterm population we compared the results of ALBERTs\_19 with ALBERTs\_14 and MPNA\_04. For the term-borns we compared the results of ALBERTs\_19 with ALBERTs\_4 and MPNA\_04\_Terms. Besides, we used two-tailed paired TTEST to compare the overall performance of these methods for each group.

**Results**

A total of forty atlases for 3T MR data sets of neonates resulting from individual pairwise registration and label fusion were created automatically (Figure 2), consisting of 50 ROIs each. Each atlas is the result of label fusion of the remaining 19 ALBERTs (20 atlases), 14 ALBERTs in the cases of preterms (15 atlases), or 4 ALBERTs (5 atlases) in the cases of terms.

In the case of template-based segmentations resulting from single registrations of a template to a target, we used the four templates created using different candidate targets and fusing different cohorts (Figure 3) described in Table 1. The eight MPNAs created for the corresponding templates (Figure 3) have been described in Table 2. In total, this resulted in 160 (8x20) individualized segmentations via templates and MPNAs that were compared with their respective manual gold standard.

Validation was performed by means of Dice measurements. In Table 3 and Figure 4 we display the results of the validation of the different approaches for automatic segmentation when compared with manual gold standards. In Figure 4, we display some comparative Dice measurements for the approaches that performed best, either fusing ALBERTs or using MPNAs. The best methods for each group and the Dice indices for all 50 ROIs are displayed in Tables 4, 5, 6, 7, 8, 9.

We compared the performance of the best methods for each group using two-tailed paired TTEST, after Bonferroni correction for multiple comparisons, with regards to each ROI. For the preterm population we compared the results of ALBERTs\_19 with ALBERTs\_14 and MPNA\_04. For the term-borns we compared the results of ALBERTs\_19 with ALBERTs\_4 and MPNA\_04\_Terms. Two-tailed paired TTEST showed that the overall performance of ALBERTs\_19 and ALBERTs\_14 was significantly better than MPNA\_04 (Table 3). Also, for the term population, MPNA\_04\_Terms performed significantly better than ALBERTs\_19 and ALBERTs\_4 (Table 3).

For the preterms, in a regional level, ALBERTs\_19 performed similarly to ALBERTs\_14, without significant differences after Bonferroni correction for multiple comparisons (Tables 4–5). ALBERTs\_19 performed better than MPNA\_04 in all the regions in either one or both hemispheres, apart from the posterior part of the superior temporal gyrus and the posterior part of the cingulate gyrus (Tables 4, 6).

For the term-borns, in a regional level, ALBERTs\_19 performed similarly to ALBERTs\_4, without significant differences after Bonferroni correction for multiple comparisons (Tables 7–8). MPNA\_04\_Terms performed better than ALBERTs\_19 in the

**Table 4.** Dice statistics for 50 ROIs with fusion approach ALBERTs\_19 for preterms.

No	Name of Structure	Right Hemisphere					Left Hemisphere				
		Mean Dice	SD	CV (%)	Min	Max	Mean Dice	SD	CV (%)	Min	Max
<b>Temporal Lobe</b>											
1; 2	Hippocampus	0.79	0.08	10	0.53	0.85	0.79	0.06	7	0.65	0.86
3; 4	Amygdala	0.81	0.04	5	0.75	0.88	0.80	0.04	5	0.72	0.86
5; 6	Anterior temporal lobe, medial part	0.84	0.04	4	0.76	0.90	0.83	0.05	6	0.70	0.88
7; 8	Anterior temporal lobe, lateral part	0.81	0.05	6	0.73	0.89	0.80	0.04	6	0.71	0.85
9; 10	Parahippocampal and ambient gyri ant.p.	0.80	0.06	7	0.67	0.86	0.82	0.03	4	0.74	0.87
25; 24	Parahippocampal and ambient gyri post.p.	0.72	0.05	8	0.60	0.80	0.66	0.08	13	0.43	0.78
11; 12	Superior temporal gyrus, middle part	0.84	0.04	5	0.75	0.89	0.82	0.04	5	0.75	0.88
31; 30	Superior temporal gyrus, post.p.	0.64	0.11	17	0.35	0.78	0.61	0.12	20	0.35	0.78
13; 14	Middle and inferior temporal gyrus ant.p.	0.85	0.03	4	0.80	0.88	0.82	0.04	5	0.71	0.87
29; 28	Middle and inferior temporal gyrus post.p.	0.78	0.05	7	0.68	0.85	0.76	0.07	9	0.64	0.84
15; 16	Fusiform gyrus ant.p.	0.75	0.10	13	0.45	0.84	0.75	0.09	13	0.54	0.86
27; 26	Fusiform gyrus post.p.	0.69	0.09	12	0.52	0.80	0.62	0.14	23	0.36	0.78
<b>Posterior Fossa</b>											
17; 18	Cerebellum	0.93	0.02	2	0.88	0.95	0.93	0.02	2	0.89	0.96
19	Brainstem	0.93	0.01	1	0.90	0.95	<i>(unpaired)</i>				
<b>Insula and Cingulate gyri</b>											
21; 20	Insula	0.82	0.05	6	0.71	0.87	0.83	0.05	6	0.72	0.88
33; 32	Cingulate gyrus, anterior part	0.79	0.10	13	0.51	0.88	0.77	0.08	10	0.59	0.88
35; 34	Cingulate gyrus, posterior part	0.80	0.05	6	0.68	0.86	0.80	0.05	6	0.67	0.85
<b>Frontal Lobe</b>											
37; 36	Frontal lobe	0.93	0.02	2	0.87	0.95	0.93	0.02	2	0.89	0.95
<b>Occipital Lobe</b>											
23; 22	Occipital lobe	0.82	0.07	8	0.68	0.89	0.83	0.05	6	0.69	0.90
<b>Parietal Lobe</b>											
39; 38	Parietal lobe	0.88	0.02	2	0.85	0.90	0.88	0.02	2	0.84	0.90
<b>Basal Ganglia and Thalamus</b>											
41; 40	Caudate nucleus	0.83	0.04	4	0.73	0.87	0.81	0.05	6	0.71	0.86
43; 42	Thalamus	0.89	0.04	5	0.77	0.93	0.89	0.02	2	0.86	0.92
45; 44	Sub-thalamic nucleus	0.71	0.07	9	0.59	0.82	0.72	0.04	6	0.62	0.77
47; 46	Lentiform nucleus	0.84	0.04	5	0.76	0.91	0.83	0.04	5	0.74	0.88
<b>Corpus Callosum</b>											
48	Corpus callosum	0.75	0.04	6	0.65	0.81	<i>(unpaired)</i>				
<b>Ventricles</b>											
49; 50	Lateral ventricles	0.83	0.03	4	0.77	0.87	0.83	0.04	4	0.77	0.88

doi:10.1371/journal.pone.0059990.t004

anterior part of the middle and inferior temporal gyrus and the parietal lobe (Tables 7, 9). MPNA\_04\_Terms performed better than ALBERTs\_4 in the posterior part of the parahippocampal gyrus, the thalamus and the lentiform nucleus (Tables 8–9).

In Figure 5 we display the preliminary results of an automatic segmentation of developing brain MRIs that do not belong to the cohort of twenty used for the creation of the manually defined ALBERTs. In this instance, automatic segmentation was achieved via a single registration of a template constructed from term-borns (Template\_04), whereas the MPNA was obtained through fusion of all ALBERTs transformed into the space of Template\_04 (MPNA\_04, see Tables 1 and 2). The results are visually acceptable.

ALBERTs and MPNAs with corresponding MRIs and templates will become available through our website [www.brain-development.org](http://www.brain-development.org).

### Discussion

We present two methods for automatic segmentation of neonatal brain MR images into 50 ROIs. The first approach is based on fusion of anatomical prior information from various manually constructed neonatal atlases after one pairwise registration per atlas used. The second approach is based on propagation of labels from various neonatal MPNAs, requiring only one registration. In both cases we evaluated the performance of

**Table 5.** Dice statistics for 50 ROIs with fusion approach ALBERTs\_14 for preterms.

No	Name of Structure	Right Hemisphere					Left Hemisphere				
		Mean Dice	SD	CV (%)	Min	Max	Mean Dice	SD	CV (%)	Min	Max
<b>Temporal Lobe</b>											
1; 2	Hippocampus	0.78	0.07	9	0.53	0.85	0.75	0.09	12	0.49	0.86
3; 4	Amygdala	0.77	0.09	12	0.54	0.88	0.81	0.04	5	0.72	0.87
5; 6	Anterior temporal lobe, medial part	0.83	0.04	5	0.75	0.90	0.82	0.07	8	0.66	0.89
7; 8	Anterior temporal lobe, lateral part	0.79	0.07	9	0.64	0.89	0.76	0.09	12	0.47	0.85
9; 10	Parahippocampal and ambient gyri ant.p.	0.80	0.05	6	0.67	0.86	0.81	0.05	6	0.66	0.87
25; 24	Parahippocampal and ambient gyri post.p.	<b>0.72*</b>	0.05	7	0.60	0.80	0.64	0.10	15	0.42	0.78
11; 12	Superior temporal gyrus, middle part	0.83	0.04	5	0.75	0.89	0.77	0.10	13	0.51	0.88
31; 30	Superior temporal gyrus, post.p.	0.66	0.11	16	0.35	0.78	0.56	0.16	29	0.22	0.78
13; 14	Middle and inferior temporal gyrus ant.p.	0.84	0.03	4	0.77	0.88	0.78	0.07	9	0.60	0.87
29; 28	Middle and inferior temporal gyrus post.p.	0.77	0.06	7	0.66	0.85	0.70	0.13	19	0.38	0.84
15; 16	Fusiform gyrus ant.p.	0.74	0.09	12	0.45	0.84	0.73	0.11	15	0.46	0.86
27; 26	Fusiform gyrus post.p.	0.69	0.09	13	0.52	0.80	0.61	0.14	23	0.36	0.78
<b>Posterior Fossa</b>											
17; 18	Cerebellum	0.92	0.02	2	0.88	0.95	0.92	0.03	3	0.85	0.96
19	Brainstem	<b>0.93*</b>	0.01	1	0.90	0.95	<i>(unpaired)</i>				
<b>Insula and Cingulate gyri</b>											
21; 20	Insula	0.82	0.05	6	0.72	0.88	0.82	0.05	6	0.72	0.88
33; 32	Cingulate gyrus, anterior part	0.79	0.09	12	0.51	0.88	<b>0.77*</b>	0.08	10	0.59	0.88
35; 34	Cingulate gyrus, posterior part	0.81	0.05	6	0.68	0.88	0.81	0.05	6	0.67	0.87
<b>Frontal Lobe</b>											
37; 36	Frontal lobe	0.93	0.02	2	0.87	0.95	<b>0.93*</b>	0.01	1	0.89	0.95
<b>Occipital Lobe</b>											
23; 22	Occipital lobe	0.82	0.06	7	0.68	0.89	0.83	0.05	6	0.69	0.90
<b>Parietal Lobe</b>											
39; 38	Parietal lobe	0.88	0.02	2	0.85	0.90	0.88	0.02	2	0.84	0.90
<b>Basal Ganglia and Thalamus</b>											
41; 40	Caudate nucleus	0.80	0.07	9	0.61	0.87	0.82	0.04	5	0.71	0.86
43; 42	Thalamus	0.89	0.04	4	0.77	0.93	0.88	0.04	5	0.75	0.92
45; 44	Sub-thalamic nucleus	0.72	0.06	8	0.59	0.82	0.72	0.04	6	0.62	0.77
47; 46	Lentiform nucleus	0.83	0.05	6	0.68	0.91	0.83	0.04	5	0.74	0.88
<b>Corpus Callosum</b>											
48	Corpus callosum	0.75	0.04	6	0.65	0.82	<i>(unpaired)</i>				
<b>Ventricles</b>											
49; 50	Lateral ventricles	0.81	0.06	7	0.64	0.87	0.81	0.07	9	0.59	0.88

\*Significant difference (two-tailed paired TTEST,  $p < 0.05$ ) with corresponding value in Table 4. Not significant after correction for multiple comparisons with Bonferroni correction.

doi:10.1371/journal.pone.0059990.t005

different selections neonatal atlases and MPNAs obtained via different strategies. The approaches were compared with the manual “gold standard” segmentations by means of the Dice overlap coefficient. The maximum Dice values obtained, averaged across all regions, were  $0.81 \pm 0.02$  using label propagation and fusion for the preterm population, and  $0.81 \pm 0.02$  using the single registration of a MPNA based on term controls only, in combination with a template based on term controls only, too. Such Dice overlaps are in line with results using maximum probability maps in adults, and somewhat lower than multi-atlas propagation and label fusion in adults [14]. Segmentations of

unlabeled ex-cohort target images yielded segmentations of high quality on visual inspection.

In terms of atlases used, pre-processing pipeline and parameters used, we present the first detailed evaluation of several strategies for the automatic segmentation of neonatal and developing brains. This was made possible through the availability of manual priors. It took 18 person-months to delineate the 1000 ( $50 \times 20$ ) structures and thus create the first cohort of neonatal manual priors, and another four to check them for consistency with the protocols; it is unlikely that larger single-investigator datasets will ever become available. The ALBERTs will become available through our

**Table 6.** Dice statistics for 50 ROIs with MPNA\_04 approach for preterms.

No	Name of Structure	Right Hemisphere					Left Hemisphere				
		Mean Dice	SD	CV (%)	Min	Max	Mean Dice	SD	CV (%)	Min	Max
<b>Temporal Lobe</b>											
1; 2	Hippocampus	<b>0.74<sup>†</sup></b>	0.08	11	0.51	0.82	0.74*	0.05	7	0.64	0.80
3; 4	Amygdala	0.77*	0.03	4	0.70	0.83	0.72*	0.10	13	0.51	0.81
5; 6	Anterior temporal lobe, medial part	<b>0.79<sup>†</sup></b>	0.06	7	0.64	0.86	<b>0.68<sup>†</sup></b>	0.12	18	0.49	0.85
7; 8	Anterior temporal lobe, lateral part	0.76	0.07	9	0.64	0.84	0.68*	0.17	36	0.16	0.67
9; 10	Parahippocampal and ambient gyri ant.p.	0.77*	0.06	7	0.61	0.82	<b>0.77<sup>†</sup></b>	0.03	4	0.71	0.82
25; 24	Parahippocampal and ambient gyri post.p.	<b>0.66<sup>†</sup></b>	0.07	10	0.54	0.76	0.59*	0.08	14	0.42	0.74
11; 12	Superior temporal gyrus, middle part	0.80*	0.05	6	0.66	0.85	0.73*	0.05	7	0.60	0.79
31; 30	Superior temporal gyrus, post.p.	0.60	0.12	21	0.40	0.78	0.55	0.14	25	0.20	0.69
13; 14	Middle and inferior temporal gyrus ant.p.	<b>0.78<sup>††</sup></b>	0.03	4	0.73	0.85	<b>0.74<sup>††</sup></b>	0.06	8	0.60	0.82
29; 28	Middle and inferior temporal gyrus post.p.	0.76	0.05	7	0.65	0.85	0.66*	0.10	15	0.45	0.75
15; 16	Fusiform gyrus ant.p.	0.68*	0.08	12	0.45	0.78	0.69*	0.08	11	0.55	0.78
27; 26	Fusiform gyrus post.p.	0.65*	0.08	13	0.52	0.79	0.57	0.09	16	0.37	0.70
<b>Posterior Fossa</b>											
17; 18	Cerebellum	<b>0.89<sup>††</sup></b>	0.03	3	0.84	0.93	<b>0.86<sup>††</sup></b>	0.03	3	0.78	0.90
19	Brainstem	<b>0.92<sup>††</sup></b>	0.01	1	0.90	0.94	(unpaired)				
<b>Insula and Cingulate gyri</b>											
21; 20	Insula	<b>0.76<sup>†</sup></b>	0.06	7	0.60	0.83	0.68*	0.15	22	0.39	0.82
33; 32	Cingulate gyrus, anterior part	0.73	0.09	12	0.55	0.84	0.71*	0.08	12	0.49	0.81
35; 34	Cingulate gyrus, posterior part	0.78	0.05	7	0.68	0.83	0.77	0.04	6	0.70	0.84
<b>Frontal Lobe</b>											
37; 36	Frontal lobe	<b>0.91<sup>††</sup></b>	0.02	2	0.87	0.93	0.90*	0.03	4	0.81	0.93
<b>Occipital Lobe</b>											
23; 22	Occipital lobe	0.82	0.04	5	0.73	0.89	0.79*	0.05	6	0.70	0.87
<b>Parietal Lobe</b>											
39; 38	Parietal lobe	0.86*	0.01	1	0.83	0.88	<b>0.86<sup>††</sup></b>	0.02	2	0.82	0.88
<b>Basal Ganglia and Thalamus</b>											
41; 40	Caudate nucleus	<b>0.76<sup>††</sup></b>	0.04	5	0.69	0.82	0.71*	0.10	14	0.51	0.81
43; 42	Thalamus	0.89	0.03	4	0.81	0.92	<b>0.86<sup>†</sup></b>	0.02	3	0.80	0.90
45; 44	Sub-thalamic nucleus	0.64*	0.09	15	0.41	0.77	0.69	0.07	10	0.54	0.79
47; 46	Lentiform nucleus	<b>0.78<sup>†</sup></b>	0.06	8	0.63	0.84	0.73*	0.11	15	0.53	0.86
<b>Corpus Callosum</b>											
48	Corpus callosum	<b>0.70<sup>††</sup></b>	0.05	7	0.58	0.77	(unpaired)				
<b>Ventricles</b>											
49; 50	Lateral ventricles	<b>0.78<sup>††</sup></b>	0.04	5	0.70	0.84	<b>0.74<sup>††</sup></b>	0.03	5	0.67	0.80

<sup>††</sup>Significant difference (two-tailed paired TTEST, p<0.0001) with corresponding value in Table 4, after Bonferroni correction for multiple comparisons.

<sup>†</sup>Significant difference (two-tailed paired TTEST, p<0.001) with corresponding value in Table 4, after Bonferroni correction for multiple comparisons.

\*Significant difference (two-tailed paired TTEST, p<0.05) with corresponding value in Table 4, after Bonferroni correction for multiple comparisons.

doi:10.1371/journal.pone.0059990.t006

website ([www.brain-development.org](http://www.brain-development.org)). While validation of automatic labeling methods is only possible within-sample, where labels created manually with the same protocol are available for calculating overlaps, the availability of the atlases will make it possible to assess automatic segmentation of other cohorts, e.g. NIHPD (<http://pediatricmri.nih.gov/nihpdp>).

The validation of the two different methods is based on leave-one-out approaches, which have been widely used by researchers in the past, including our team [14,16,26,35,36]. In such an approach, for the validation of the ALBERTs\_19 performance for

example, for each of the target brains we use as priors the remaining 19. This means that the manual segmentation of the target brain and the automatic segmentation we obtain after registration, propagation and fusion are two totally independent segmentations. The results of the validation by all means highlight the potential to segment unseen brains.

Automatic segmentation is commonly used in studies of adult MR brain images but has been challenging in infants. An initial spatial normalization to a template or average brain in a standard stereotaxic space [37,38] can be problematic. Spatial normaliza-

**Table 7.** Dice statistics for 50 ROIs with fusion approach ALBERTs\_19 for terms.

No	Name of Structure	Right Hemisphere					Left Hemisphere				
		Mean Dice	SD	CV (%)	Min	Max	Mean Dice	SD	CV (%)	Min	Max
<b>Temporal Lobe</b>											
1; 2	Hippocampus	0.75	0.03	4	0.71	0.77	0.64	0.09	14	0.49	0.74
3; 4	Amygdala	0.65	0.09	14	0.54	0.77	0.85	0.01	2	0.84	0.87
5; 6	Anterior temporal lobe, medial part	0.81	0.05	6	0.75	0.85	0.77	0.10	13	0.66	0.89
7; 8	Anterior temporal lobe, lateral part	0.72	0.08	11	0.64	0.82	0.66	0.12	18	0.47	0.79
9; 10	Parahippocampal and ambient gyri ant.p.	0.79	0.02	3	0.76	0.82	0.80	0.08	10	0.66	0.85
25; 24	Parahippocampal and ambient gyri post.p.	0.74	0.02	3	0.72	0.77	0.59	0.12	21	0.42	0.71
11; 12	Superior temporal gyrus, middle part	0.81	0.05	6	0.75	0.86	0.63	0.08	13	0.51	0.72
31; 30	Superior temporal gyrus, post.p.	0.69	0.09	13	0.54	0.77	0.50	0.18	44	0.32	0.73
13; 14	Middle and inferior temporal gyrus ant.p	0.81	0.04	5	0.77	0.86	0.68	0.06	8	0.60	0.74
29; 28	Middle and inferior temporal gyrus post.p.	0.76	0.07	9	0.66	0.83	0.52	0.10	20	0.38	0.64
15; 16	Fusiform gyrus ant.p.	0.73	0.07	10	0.60	0.79	0.65	0.14	22	0.46	0.81
27; 26	Fusiform gyrus post.p.	0.69	0.10	14	0.54	0.78	0.57	0.14	25	0.39	0.72
<b>Posterior Fossa</b>											
17; 18	Cerebellum	0.92	0.01	2	0.90	0.94	0.90	0.03	3	0.85	0.93
19	Brainstem	0.93	0.01	1	0.92	0.95	<i>(unpaired)</i>				
<b>Insula and Cingulate gyri</b>											
21; 20	Insula	0.75	0.06	7	0.68	0.82	0.80	0.04	5	0.74	0.84
33; 32	Cingulate gyrus, anterior part	0.79	0.05	7	0.70	0.83	0.77	0.08	11	0.67	0.85
35; 34	Cingulate gyrus, posterior part	0.82	0.06	7	0.73	0.88	0.85	0.02	3	0.81	0.87
<b>Frontal Lobe</b>											
37; 36	Frontal lobe	0.92	0.01	1	0.91	0.93	0.93	0.01	1	0.92	0.94
<b>Occipital Lobe</b>											
23; 22	Occipital lobe	0.82	0.03	4	0.79	0.85	0.83	0.04	5	0.77	0.90
<b>Parietal Lobe</b>											
39; 38	Parietal lobe	0.86	0.01	2	0.85	0.88	0.86	0.01	1	0.85	0.87
<b>Basal Ganglia and Thalamus</b>											
41; 40	Caudate nucleus	0.72	0.08	11	0.61	0.83	0.82	0.03	4	0.78	0.86
43; 42	Thalamus	0.89	0.02	3	0.85	0.91	0.82	0.05	6	0.75	0.86
45; 44	Sub-thalamic nucleus	0.72	0.05	6	0.65	0.75	0.73	0.04	5	0.67	0.77
47; 46	Lentiform nucleus	0.78	0.06	8	0.68	0.84	0.83	0.04	5	0.77	0.88
<b>Corpus Callosum</b>											
48	Corpus callosum	0.76	0.05	7	0.67	0.82	<i>(unpaired)</i>				
<b>Ventricles</b>											
49; 50	Lateral ventricles	0.77	0.09	12	0.64	0.87	0.73	0.10	14	0.59	0.87

doi:10.1371/journal.pone.0059990.t007

tion requires an appropriate template [39], and when cerebral images of children are aligned using an adult template the variation of anatomical landmarks is increased [40,41], and greater nonlinear local deformations are required for registration [42]. Indeed the use of the adult MNI template [43] in infants and children has been criticized [44,45,46], and pediatric templates recommended for the analysis of pediatric images [47,48,49]. Transforming neonatal rather than pediatric cerebral images to an adult template has additional difficulties [37] including age-dependent differences in regional brain size [50] and unmyelinated white matter with different MR characteristics in neonates [51], so that several groups have constructed specific neonatal templates [37,52,53,54,55,56].

Tissue segmentation is also difficult due to the different and highly variable tissue characteristics [18,57]. Prastawa et al. (2005) reported an automated method using a three-subject atlas for GM, CSF and myelinated and unmyelinated WM, but did not attempt subcortical GM segmentation [20]. Warfield et al. (2000) use a specific template for newborn brains with predefined classifications for myelinated and unmyelinated WM [58]. Huppi et al. (1998) and Inder et al. (2005) showed tissue class segmentation results of newborn infants using this method [59,60]. Kazemi et al. (2011) presented a neonatal brain phantom that consists of 9 different tissue types: skin, fat, muscle, skull, dura mater, gray matter, myelinated white matter, nonmyelinated white matter and cerebrospinal fluid [61].

**Table 8.** Dice statistics for 50 ROIs with fusion approach ALBERTs\_4 for terms.

No	Name of Structure	Right Hemisphere					Left Hemisphere				
		Mean Dice	SD	CV (%)	Min	Max	Mean Dice	SD	CV (%)	Min	Max
<b>Temporal Lobe</b>											
1; 2	Hippocampus	0.73	0.07	10	0.51	0.82	0.71	0.08	11	0.55	0.80
3; 4	Amygdala	<b>0.74*</b>	0.07	10	0.57	0.83	0.75	0.09	13	0.51	0.85
5; 6	Anterior temporal lobe, medial part	0.79	0.05	7	0.64	0.86	0.72	0.12	17	0.49	0.85
7; 8	Anterior temporal lobe, lateral part	0.74	0.08	11	0.60	0.86	<b>0.52*</b>	0.17	33	0.16	0.76
9; 10	Parahippocampal and ambient gyri ant.p.	<b>0.77*</b>	0.05	6	0.61	0.82	0.77	0.04	5	0.69	0.82
25; 24	Parahippocampal and ambient gyri post.p.	0.66	0.06	9	0.54	0.76	0.59	0.08	13	0.42	0.74
11; 12	Superior temporal gyrus, middle part	0.80	0.04	6	0.66	0.85	<b>0.70*</b>	0.08	11	0.51	0.79
31; 30	Superior temporal gyrus, post.p.	0.62	0.11	19	0.40	0.78	<b>0.51*</b>	0.14	28	0.20	0.69
13; 14	Middle and inferior temporal gyrus ant.p.	0.79	0.03	4	0.73	0.85	<b>0.73*</b>	0.06	8	0.60	0.82
29; 28	Middle and inferior temporal gyrus post.p.	0.76	0.05	7	0.65	0.85	<b>0.63*</b>	0.10	16	0.45	0.75
15; 16	Fusiform gyrus ant.p.	0.68	0.08	11	0.45	0.78	0.68	0.08	12	0.54	0.78
27; 26	Fusiform gyrus post.p.	0.66	0.08	13	0.52	0.79	0.58	0.08	15	0.37	0.70
<b>Posterior Fossa</b>											
17; 18	Cerebellum	<b>0.90*</b>	0.02	3	0.84	0.93	<b>0.87*</b>	0.03	4	0.78	0.91
19	Brainstem	0.93	0.01	1	0.90	0.94	(unpaired)				
<b>Insula and Cingulate gyri</b>											
21; 20	Insula	<b>0.76*</b>	0.05	6	0.60	0.83	<b>0.69*</b>	0.13	19	0.39	0.82
33; 32	Cingulate gyrus, anterior part	0.73	0.08	11	0.55	0.84	0.73	0.08	11	0.49	0.82
35; 34	Cingulate gyrus, posterior part	0.78	0.05	7	0.68	0.87	0.78	0.04	6	0.70	0.84
<b>Frontal Lobe</b>											
37; 36	Frontal lobe	0.91	0.01	2	0.87	0.93	0.91	0.03	4	0.81	0.94
<b>Occipital Lobe</b>											
23; 22	Occipital lobe	0.82	0.04	5	0.73	0.89	<b>0.81*</b>	0.05	6	0.70	0.89
<b>Parietal Lobe</b>											
39; 38	Parietal lobe	0.86	0.02	2	0.83	0.89	<b>0.86*</b>	0.02	2	0.82	0.88
<b>Basal Ganglia and Thalamus</b>											
41; 40	Caudate nucleus	0.75	0.05	7	0.62	0.82	0.73	0.09	13	0.51	0.84
43; 42	Thalamus	0.89	0.03	3	0.81	0.92	0.85	0.03	4	0.75	0.90
45; 44	Sub-thalamic nucleus	0.65	0.09	13	0.41	0.77	0.69	0.06	9	0.54	0.79
47; 46	Lentiform nucleus	0.78	0.05	7	0.63	0.84	0.75	0.11	14	0.53	0.86
<b>Corpus Callosum</b>											
48	Corpus callosum	0.72	0.05	7	0.58	0.80	(unpaired)				
<b>Ventricles</b>											
49; 50	Lateral ventricles	0.77	0.05	7	0.63	0.84	0.73	0.06	8	0.55	0.82

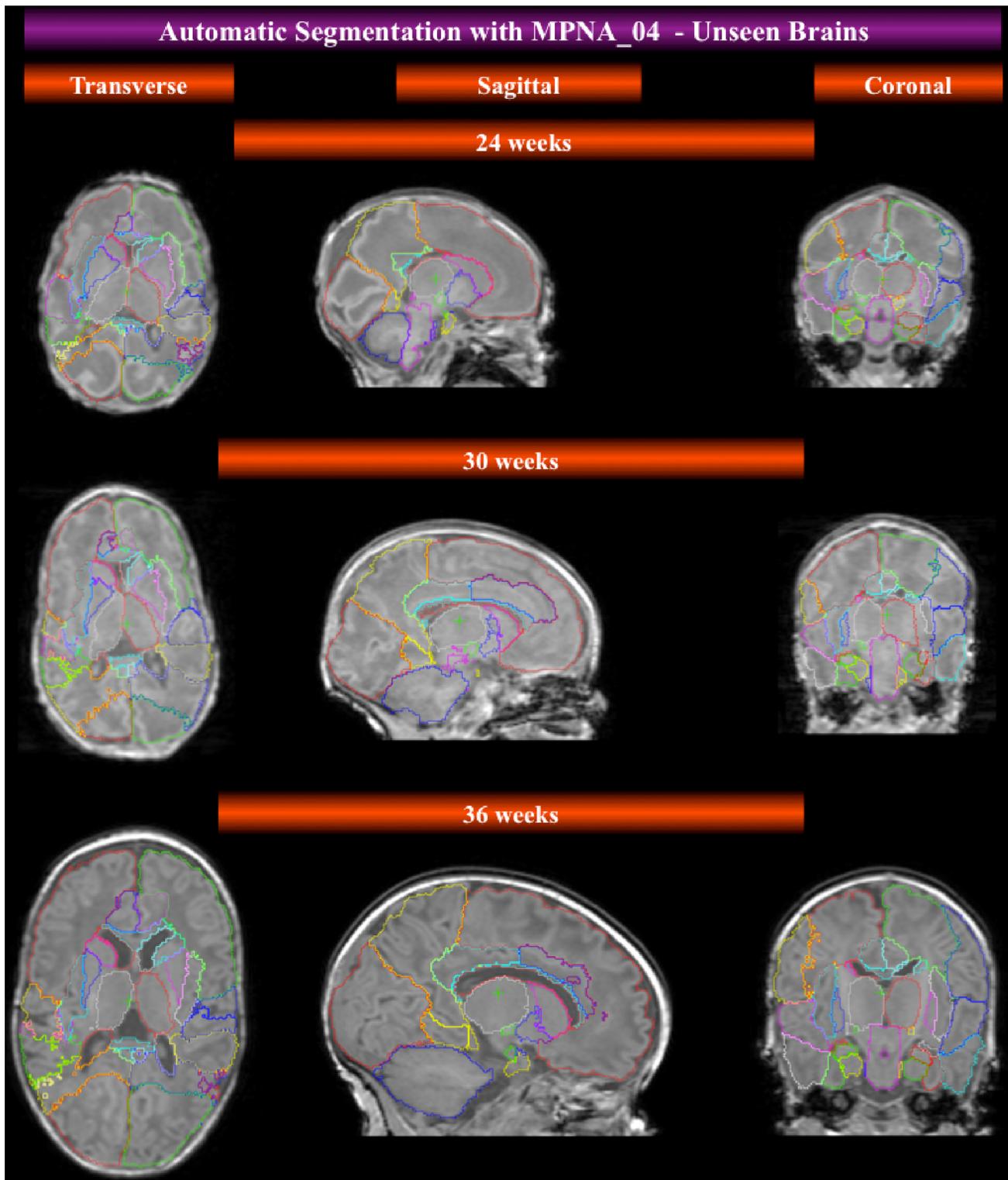
\*Significant difference (two-tailed paired TTEST, p<0.05) with corresponding value in Table 7. Not significant after correction for multiple comparisons with Bonferroni correction.

doi:10.1371/journal.pone.0059990.t008

Despite the difficulties there have been previous reports of automatic segmentation methods for newborn infants. Nishida et al. (2006) presented a semi-automated method for segmentation of preterm infants at term corrected age into anatomical ROI. Unfortunately, their cohort did not include any term controls and they did not validate with any gold standard segmentation and hence the comparison with our method is difficult [62]. There are also approaches based on Diffusion Tensor Imaging, resulting in segmentations with numerous regions with or without clear anatomical or functional correspondence [63,64]. These approaches yield results that are visually plausible, but have not

yet been compared or validated against external standards in the neonatal population, as for example defined anatomical protocols. It is hence difficult to compare our work with these two studies.

The average spaces required for spatial normalisation were created using an approach similar to that of Guimond et al. (2000) and Rueckert et al. (2003) for averaging local deformations [39,65], also used in the pediatric population [32]. In some studies, contrary to the main trend of using a standard reference template like MNI, a single subject data set of the image group is selected as the reference or template image [66,67,68]. A disadvantage of this atlas construction method is that the resulting



**Figure 5. Performance of MPNA\_04 in the automatic segmentation of three randomly chosen unlabeled developing brains at various ages, which did not form part of the cohort of priors.** The segmentation is the result of a single step registration and propagation of the MPNA.

doi:10.1371/journal.pone.0059990.g005

atlas can inherently contain unique features of the selected initial reference image, which results in local topological bias [69]. Group-wise registration, based on the minimization of the average

deformation field, could be a solution to the problem [70,71]. However, the presence of a few very unusually shaped brains (cf. Figure 1 of Gousias et al. (2012) [21]) coupled with the small

**Table 9.** Dice statistics for 50 ROIs with MPNA\_04\_Terms approach for terms.

No	Name of Structure	Right Hemisphere					Left Hemisphere				
		Mean Dice	SD	CV (%)	Min	Max	Mean Dice	SD	CV (%)	Min	Max
<b>Temporal Lobe</b>											
1; 2	Hippocampus	0.76	0.03	4	0.70	0.78	0.74	0.08	11	0.60	0.82
3; 4	Amygdala	0.79	0.07	9	0.67	0.84	0.82	0.03	4	0.78	0.86
5; 6	Anterior temporal lobe, medial part	0.84	0.04	5	0.78	0.88	0.84	0.05	5	0.79	0.89
7; 8	Anterior temporal lobe, lateral part	0.77	0.04	5	0.70	0.79	0.81	0.06	8	0.72	0.88
9; 10	Parahippocampal and ambient gyri ant.p.	0.81	0.02	2	0.79	0.84	0.80	0.05	6	0.74	0.85
25; 24	Parahippocampal and ambient gyri post.p.	0.74	0.07	9	0.62	0.79	<b>0.72<sup>†</sup></b>	0.04	6	0.66	0.76
11; 12	Superior temporal gyrus, middle part	0.83	0.03	3	0.79	0.86	0.82	0.04	4	0.77	0.86
31; 30	Superior temporal gyrus, post.p.	0.72	0.06	8	0.64	0.78	0.76	0.03	4	0.72	0.79
13; 14	Middle and inferior temporal gyrus ant.p.	0.82	0.02	3	0.79	0.85	0.82*	0.06	8	0.71	0.87
29; 28	Middle and inferior temporal gyrus post.p.	0.78	0.04	5	0.72	0.82	0.79	0.04	6	0.73	0.85
15; 16	Fusiform gyrus ant.p.	0.74	0.06	8	0.69	0.81	0.71	0.06	9	0.61	0.78
27; 26	Fusiform gyrus post.p.	0.71	0.07	11	0.63	0.80	0.75	0.07	10	0.63	0.82
<b>Posterior Fossa</b>											
17; 18	Cerebellum	0.94	0.01	1	0.91	0.95	0.94	0.02	2	0.90	0.95
19	Brainstem	0.94	0.01	1	0.93	0.94	(unpaired)				
<b>Insula and Cingulate gyri</b>											
21; 20	Insula	0.85	0.01	2	0.84	0.87	0.85	0.02	2	0.83	0.87
33; 32	Cingulate gyrus, anterior part	0.75	0.06	8	0.65	0.81	0.79	0.05	6	0.71	0.82
35; 34	Cingulate gyrus, posterior part	0.82	0.04	6	0.75	0.86	0.83	0.02	3	0.79	0.86
<b>Frontal Lobe</b>											
37; 36	Frontal lobe	0.93	0.01	1	0.91	0.94	0.93	0.00	1	0.93	0.94
<b>Occipital Lobe</b>											
23; 22	Occipital lobe	0.87	0.04	5	0.79	0.90	0.89	0.01	1	0.88	0.91
<b>Parietal Lobe</b>											
39; 38	Parietal lobe	0.89	0.01	1	0.88	0.90	0.89*	0.01	1	0.87	0.90
<b>Basal Ganglia and Thalamus</b>											
41; 40	Caudate nucleus	0.78	0.04	5	0.73	0.81	0.80	0.04	5	0.77	0.86
43; 42	Thalamus	0.90	0.02	2	0.87	0.92	<b>0.88<sup>†</sup></b>	0.03	4	0.84	0.92
45; 44	Sub-thalamic nucleus	0.70	0.05	6	0.65	0.78	0.72	0.05	7	0.63	0.75
47; 46	Lentiform nucleus	<b>0.83<sup>†</sup></b>	0.05	6	0.77	0.88	0.85	0.03	3	0.81	0.88
<b>Corpus Callosum</b>											
48	Corpus callosum	0.78	0.03	3	0.74	0.80	(unpaired)				
<b>Ventricles</b>											
49; 50	Lateral ventricles	0.75	0.08	11	0.64	0.83	0.76	0.08	10	0.63	0.84

\*Significant difference (two-tailed paired TTEST, p<0.05) with corresponding value in Table 7, after Bonferroni correction for multiple comparisons.

†Significant difference (two-tailed paired TTEST, p<0.05) with corresponding value in Table 8, after Bonferroni correction for multiple comparisons.

doi:10.1371/journal.pone.0059990.t009

number of subjects available due to the phenomenal effort required for manual delineation, leads us to believe that our strategy of explicitly choosing “normal looking” brains is appropriate in this situation. Extremely dolichocephalic subjects or subjects with obvious major asymmetries were not selected as candidate targets. Group-wise registration remains an area for future study.

The average Dice indices for the various approaches (Table 3) and the Dice indices for the best approaches (Tables 4, 5, 6, 7, 8, 9) indicate that fusing ALBERTs of the same group (in terms of degree of myelination) yields better or similar results than fusing

more classifiers from different groups: term priors performed better for a term target and preterm priors better for a preterm target, confirming our previous findings [17]. Also, single propagations from twin pair, expected to be more similar than brain MRIs from unrelated subjects, perform at a level comparable to fusion. This finding highlights the importance of resemblance in sulcal and gyral patterns between the source and the target brain, as it has been shown before in corresponding scans between different ages in the context of longitudinal segmentation [17]. Optimal template selection approaches have previously been shown to be effective in atlas-based segmentation

of confocal microscopy images of bee brains [15], as well as in human brain segmentation [72,73].

An MPNA was created for the term (MPNA\_04\_Term) and the preterm brain (MPNA\_04\_Preterm). This type of atlases [26,32] has shown its potential in the absence of a bigger database or for computational time savings [14]. In the present study, their application results in segmentation accuracies comparable with the segmentation using fusion of transformed ALBERTs. The results of the MPNA\_04 template registration, between source template and target image, show the need for crispier and not extremely smooth templates (MPNA\_02, 03), which incorporate the basic anatomical information from a smaller number of images and not necessarily the whole cohort (MPNA\_01). Besides, the results illustrated in Figure 5 highlight the effectiveness of the MPNA\_04 template registration through its intrinsic smoothness to capture the lack of prominent cortical anatomical landmarks in the extremely preterm population. The latter findings highlight the importance of the feature of smoothness, which has to be present but not to an extreme level. Template selection is important, because it has been shown that the choice of the template affects region-based volumetric analysis, either when the template does not correspond to the age cohort [49] or when multiple templates are used [74].

In neonates, the selection of the candidate target was also based on symmetry and normality criteria. The first template (term candidate target – all subjects) is slightly rounder on transverse sections than the second (term candidate target – preterms) (Figure 3). This happens because the term brains seem to have a more round/spherical brain shape. The difference between the second template and the third (preterm candidate target – preterms) is more obvious, especially in the subcortical tissues, because of the different candidate target (Figure 3). This could indicate that a good template, in terms of representation of anatomy and corresponding tissue properties, should be limited to a cohort of data sets of tight gestational age range, due to the rapid progression of myelination of the WM and the contrast issues arising as a consequence. In case of a wider gestational age range

the template may become extremely blurry. This may be the reason template 3, which was based on the whole preterm cohort and not some images of tighter age range, did not perform as expected for the corresponding preterm population. The fact that template 04, based on a term control candidate average and MRI averaging of the transformed ALBERTs of the remaining four term controls (ALBERTs\_4\_Term), gave the best results for the term population also supports this statement.

Atlases containing such detailed segmentation can be useful in the monitoring of developmental growth of different brain regions in longitudinal studies or aid group comparisons between normal controls and pathological cases. The associated templates can be used as a reference in functional and connectivity studies and will benefit from the anatomical annotations contained in the associated MPNAs. Both methods presented here yield very plausible and comparable results, ALBERTs performing slightly better in absolute Dice measurements for the preterm. However, MPNAs have the advantage of requiring only one registration per target brain and will require fewer computational power resources (8 hours compared to  $20 \times 8 = 160$  hours for all ALBERTs, even if the latter process can be calculated in parallel on a cluster of computers).

### Acknowledgments

The authors would like to thank the parents and children whose images were used in the study and the nurses and staff of the Hammersmith Hospital Neonatal Unit and Neonatal Imaging Centre. They would like to thank the Imperial College Comprehensive Biomedical Research Centre for the support. Dr. Ioannis Spyridon Gousias would also like to thank Mr. Alvertos Gousias.

### Author Contributions

Drafted the article or revised it critically for important intellectual content: ISG AH SJC LS RAH DR ADE. Conceived and designed the experiments: ISG. Performed the experiments: ISG. Analyzed the data: ISG. Contributed reagents/materials/analysis tools: ISG AH SJC LS MAR RAH JVH DR ADE. Wrote the paper: ISG.

### References

- Bajcsy R, Lieberman R, Reivich M (1983) A computerized system for the elastic matching of deformed radiographic images to idealised atlas images. *J Comput Assist Tomogr* 7: 618–625.
- Holden M, Schnabel JA, Hill DL (2002) Quantification of small cerebral ventricular volume changes in treated growth hormone patients using nonrigid registration. *IEEE Trans Med Imaging* 21: 1292–1301.
- Iosifescu DV, Shenton ME, Warfield SK, Kikinis R, Dengler J, et al. (1997) An automated registration algorithm for measuring MRI subcortical brain structures. *Neuroimage* 6: 13–25.
- Miller MI, Christensen GE, Amit Y, Grenander U (1993) Mathematical textbook of deformable neuroanatomies. *Proc Natl Acad Sci USA* 90: 11944–11948.
- Fischl B, van der Kouwe A, Destrieux C, Halgren E, Segonne F, et al. (2004) Automatically parcellating the human cerebral cortex. *Cereb Cortex* 14: 11–22.
- Wu J, Chung AC (2009) A novel framework for segmentation of deep brain structures based on Markov dependence tree. *Neuroimage*.
- Collins DL, Zijdenbos AP, Baaré WFC, Evans AC (1999) ANIMAL+INSECT: Improved cortical structure segmentation. *LNCS* 1613: 210–223.
- Marroquin JL, Vemuri BC, Botello S, Calderon F, Fernandez-Bouzas A (2002) An accurate and efficient bayesian method for automatic segmentation of brain MRI. *IEEE Trans Med Imaging* 21: 934–945.
- Pohl KM, Fisher J, Grimson WE, Kikinis R, Wells WM (2006) A Bayesian model for joint segmentation and registration. *Neuroimage* 31: 228–239.
- Powell S, Magnotta VA, Johnson H, Jammalamadaka VK, Pierson R, et al. (2008) Registration and machine learning-based automated segmentation of subcortical and cerebellar brain structures. *Neuroimage* 39: 238–247.
- Ashburner J, Friston KJ (2005) Unified segmentation. *Neuroimage* 26: 839–851.
- Hammers A, Koeppe MJ, Free SL, Brett M, Richardson MP, et al. (2002) Implementation and application of a brain template for multiple volumes of interest. *Hum Brain Mapp* 15: 165–174.
- Svarer C, Madsen K, Hasselbalch SG, Pinborg LH, Haugbol S, et al. (2005) MR-based automatic delineation of volumes of interest in human brain PET images using probability maps. *Neuroimage* 24: 969–979.
- Heckemann RA, Hajnal JV, Aljabar P, Rueckert D, Hammers A (2006) Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. *Neuroimage* 33: 115–126.
- Rohlfing T, Brandt R, Menzel R, Maurer CR, Jr. (2004) Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. *Neuroimage* 21: 1428–1442.
- Gousias IS, Rueckert D, Heckemann RA, Dyet LE, Boardman JP, et al. (2008) Automatic segmentation of brain MRIs of 2-year-olds into 83 regions of interest. *Neuroimage* 40: 672–684.
- Gousias IS, Hammers A, Heckemann RA, Counsell SJ, Dyet LE, et al. (2010) Atlas selection strategy for automatic segmentation of pediatric brain MRIs into 83 ROIs; Imaging Systems and Techniques (IST), IEEE International Conference 2010 1–2 July 2010. pp. 290–293.
- Weisenfeld NI, Mewes AU, Warfield SK (2006) Highly accurate segmentation of brain tissue and subcortical gray matter from newborn MRI. *Med Image Comput Assist Interv* 9: 199–206.
- Rutherford M (2002) MRI of the Neonatal Brain: W. B. Saunders.
- Prastawa M, Gilmore JH, Lin W, Gerig G (2005) Automatic segmentation of MR images of the developing newborn brain. *Med Image Anal* 9: 457–466.
- Gousias IS, Edwards AD, Rutherford MA, Counsell SJ, Hajnal JV, et al. (2012) Magnetic resonance imaging of the newborn brain: manual segmentation of labelled atlases in term-born and preterm infants. *Neuroimage* 62: 1499–1509.
- Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, et al. (2004) Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23 Suppl 1: S208–219.
- Mitchell TN, Free SL, Merschhemke M, Lemieux L, Sisodiya SM, et al. (2003) Reliable callosal measurement: population normative data confirm sex-related differences. *AJNR Am J Neuroradiol* 24: 410–418.

24. Ashburner J, Friston KJ (1997) Multimodal image coregistration and partitioning - A unified framework. *NeuroImage* 6: 209–217.

25. Ashburner J, Friston KJ (1999) Nonlinear spatial normalization using basis functions. *Hum Brain Mapp* 7: 254–266.

26. Hammers A, Allom R, Koeppe MJ, Free SL, Myers R, et al. (2003) Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Hum Brain Mapp* 19: 224–247.

27. Ahsan RL, Allom R, Gousias IS, Habib H, Turkheimer FE, et al. (2007) Volumes, spatial extents and a probabilistic atlas of the human basal ganglia and thalamus. *Neuroimage* 38: 261–270.

28. Hammers A, Chen CH, Lemieux L, Allom R, Vossos S, et al. (2007) Statistical neuroanatomy of the human inferior frontal gyrus and probabilistic atlas in a standard stereotaxic space. *Hum Brain Mapp* 28: 34–48.

29. Rueckert D, Sonoda LI, Hayes C, Hill DL, Leach MO, et al. (1999) Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans Med Imaging* 18: 712–721.

30. Studholme C, Hill DLG, Hawkes DJ (1997) Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures. *Med Phys* 24: 25–35.

31. Kittler J, Hatef M, Duin RPW, Matas J (1998) On combining classifiers. *IEEE Trans Pattern Anal Mach Intell* 20: 226–239.

32. Gousias IS, Hammers A, Counsell SJ, Edwards AD, Rueckert D (2012) Automatic segmentation of pediatric brain MRIs using a maximum probability pediatric atlas; *Imaging Systems and Techniques (IST)*, IEEE International Conference on 2012 16–17 July 2012, pp. 95–100.

33. Dice LR (1945) Measure of the amount of ecological association between species. *Ecology* 26: 297–302.

34. Shattuck DW, Sandor-Leahy SR, Schaper KA, Rottenberg DA, Leahy RM (2001) Magnetic resonance image tissue classification using a partial volume model. *Neuroimage* 13: 856–876.

35. Morra JH, Tu Z, Apostolova LG, Green AE, Avedissian C, et al. (2008) Validation of a fully automated 3D hippocampal segmentation method using subjects with Alzheimer’s disease mild cognitive impairment, and elderly controls. *Neuroimage* 43: 59–68.

36. Eskildsen SF, Coupe P, Fonov V, Manjon JV, Leung KK, et al. (2012) BEaST: brain extraction based on nonlocal segmentation technique. *Neuroimage* 59: 2362–2373.

37. Kazemi K, Moghaddam HA, 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: 463–473.

38. Ashburner J, Friston KJ (2000) Voxel-based morphometry—the methods. *Neuroimage* 11: 805–821.

39. Guimond A, Meunier J, Thirion JP (2000) Average brain models: a convergence study. *Comput Vis Image Underst* 77: 192–210.

40. Burgund ED, Kang HC, Kelly JE, Buckner RL, Snyder AZ, et al. (2002) The feasibility of a common stereotaxic space for children and adults in fMRI studies of development. *Neuroimage* 17: 184–200.

41. Muzik O, Chugani DC, Juhasz C, Shen C, Chugani HT (2000) Statistical parametric mapping: assessment of application in children. *Neuroimage* 12: 538–549.

42. Wilke M, Schmithorst VJ, Holland SK (2002) Assessment of spatial normalization of whole-brain magnetic resonance images in children. *Hum Brain Mapp* 17: 48–60.

43. Evans AC, Collins DL, Mills SR, Brown ED, Kelly RL, et al. (1993) 3D statistical neuroanatomical models from 305 MRI volumes; *Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging Conference*. pp. 1813–1817.

44. Wilke M, Holland SK (2003) Variability of gray and white matter during normal development: a voxel-based MRI analysis. *Neuroreport* 14: 1887–1890.

45. Hoeksma MR, Kenemans JL, Kemner C, van Engeland H (2005) Variability in spatial normalization of pediatric and adult brain images. *Clin Neurophysiol* 116: 1188–1194.

46. Machielsen B, d’Agostino E, Maes F, Vandermeulen D, Hahn HK, et al. (2007) Linear normalization of MR brain images in pediatric patients with periventricular leukomalacia. *Neuroimage* 35: 686–697.

47. Wilke M, Holland SK, Altaye M, Gaser C (2008) Template-O-Matic: a toolbox for creating customized pediatric templates. *Neuroimage* 41: 903–913.

48. Wilke M, Schmithorst VJ, Holland SK (2003) Normative pediatric brain data for spatial normalization and segmentation differs from standard adult data. *Magn Reson Med* 50: 749–757.

49. Yoon U, Fonov VS, Perusse D, Evans AC (2009) The effect of template choice on morphometric analysis of pediatric brain data. *Neuroimage* 45: 769–777.

50. Gaillard WD, Grandin CB, Xu B (2001) Developmental aspects of pediatric MRI: considerations for image acquisition, analysis, and interpretation. *Neuroimage* 13: 239–249.

51. Ketonen LM, Hiwatashi A, Sidhu R, Westesson PL (2005) *Pediatric Brain and Spine: An Atlas of MRI and Spectroscopy*. Berlin Heidelberg: Springer Verlag.

52. Dehaene-Lambertz G, Dehaene S, Hertz-Pannier L (2002) Functional neuroimaging of speech perception in infants. *Science* 298: 2013–2015.

53. Srinivasan L, Dutta R, Counsell SJ, Allsop JM, Boardman JP, et al. (2007) Quantification of deep gray matter in preterm infants at term-equivalent age using manual volumetry of 3-tesla magnetic resonance images. *Pediatrics* 119: 759–765.

54. Altaye M, Holland SK, Wilke M, Gaser C (2008) Infant brain probability templates for MRI segmentation and normalization. *Neuroimage* 43: 721–730.

55. Kuklisova-Murgasova M, Aljabar P, Srinivasan L, Counsell SJ, Doria V, et al. (2011) A dynamic 4D probabilistic atlas of the developing brain. *Neuroimage* 54: 2750–2763.

56. Shi F, Yap PT, Wu G, Jia H, Gilmore JH, et al. (2011) Infant brain atlases from neonates to 1- and 2-year-olds. *PLoS One* 6: e18746.

57. Peterson BS, Anderson AW, Ehrenkranz R, Staib LH, Tageldin M, et al. (2003) Regional brain volumes and their later neurodevelopmental correlates in term and preterm infants. *Pediatrics* 111: 939–948.

58. Warfield SK, Kaus M, Jolesz FA, Kikinis R (2000) Adaptive, template moderated, spatially varying statistical classification. *Med Image Anal* 4: 43–55.

59. Huppi PS, Warfield S, Kikinis R, Barnes PD, Zientara GP, et al. (1998) Quantitative magnetic resonance imaging of brain development in premature and mature newborns. *Ann Neurol* 43: 224–235.

60. Inder TE, Warfield SK, Wang H, Huppi PS, Volpe JJ (2005) Abnormal cerebral structure is present at term in premature infants. *Pediatrics* 115: 286–294.

61. Kazemi K, Moghaddam HA, Grebe R, Gondry-Jouet C, Wallois F (2011) Design and construction of a brain phantom to simulate neonatal MR images. *Comput Med Imaging Graph* 35: 237–250.

62. Nishida M, Makris N, Kennedy DN, Vangel M, Fischl B, et al. (2006) Detailed semiautomated MRI based morphometry of the neonatal brain: preliminary results. *Neuroimage* 32: 1041–1049.

63. Oishi K, Mori S, Donohue PK, Ernst T, Anderson L, et al. (2011) Multi-contrast human neonatal brain atlas: Application to normal neonate development analysis. *Neuroimage* 56: 8–20.

64. Faria AV, Zhang J, Oishi K, Li X, Jiang H, et al. (2010) Atlas-based analysis of neurodevelopment from infancy to adulthood using diffusion tensor imaging and applications for automated abnormality detection. *Neuroimage* 52: 415–428.

65. Rueckert D, Frangi AF, Schnabel JA (2003) Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration. *IEEE Trans Med Imaging* 22: 1014–1025.

66. Douaud G, Smith S, Jenkinson M, Behrens T, Johansen-Berg H, et al. (2007) Anatomically related grey and white matter abnormalities in adolescent-onset schizophrenia. *Brain* 130: 2375–2386.

67. Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, et al. (2006) Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage* 31: 1487–1505.

68. Jones DK, Griffin LD, Alexander DC, Catani M, Horsfield MA, et al. (2002) Spatial normalization and averaging of diffusion tensor MRI data sets. *Neuroimage* 17: 592–617.

69. Wang Q, Seghers D, D’Agostino E, Maes F, Vandermeulen D, et al. (2005) Construction and validation of mean shape atlas templates for atlas-based brain image segmentation. *Inf Process Med Imaging* 19: 689–700.

70. Bhatia KK, Aljabar P, Boardman JP, Srinivasan L, Murgasova M, et al. (2007) Groupwise combined segmentation and registration for atlas construction. *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv* 10: 532–540.

71. Bhatia KK, Hajnal J, Hammers A, Rueckert D (2007) Similarity metrics for groupwise non-rigid registration. *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv* 10: 544–552.

72. Wu M, Rosano C, Lopeiz-Garcia P, Carter CS, Aizenstein HJ (2007) Optimum template selection for atlas-based segmentation. *Neuroimage* 34: 1612–1618.

73. Aljabar P, Heckemann RA, Hammers A, Hajnal JV, Rueckert D (2009) Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. *Neuroimage* 46: 726–738.

74. Allen JS, Bruss J, Mehta S, Grabowski T, Brown CK, et al. (2008) Effects of spatial transformation on regional brain volume estimates. *Neuroimage* 42: 535–547.