An Automated Probabilistic Tractography Tool with Anatomical Priors for Use in the Newborn Brain

Lilla Zöllei[1], Isabelle Filipiak[2], Elie Saliba[2], Laurent Barantin[2], Christophe Destrieux[2], Hugo Dupuis[2], Maria Cottier[2], Jessica Owen[3], Yangming Ou[1], Ani Varjabedian[1], Camilo Jaimes[1], Ellen Grant[4], Anastasia Yendiki[1]

[1]Massachusetts General Hospital, Boston, United States; [2]Université François-Rabelais de Tours, Tours, France; [3]Queensland Academies, Brisbane, Australia; [4]Boston Children's Hospital, Boston, United States

lzollei@nmr.mgh.harvard.edu, Isabelle.Filipiak@etu.univ-tours.fr, elie.saliba@univ-tours.fr, laurent.barantin@univ-tours.fr, christophe.destrieux@univ-tours.fr, hugo.dupuis@etu.univ-tours.fr, mariacottier@gmail.com, jessicaowen@hotmail.com, yangming.ou@mgh.harvard.edu, aniv@nmr.mgh.harvard.edu, CJAIMESCOBOS@PARTNERS.ORG, Ellen.Grant@childrens.harvard.edu, ayendiki@nmr.mgh.harvard.edu

Abstract. The ongoing myelination of white-matter fiber bundles plays a significant role in connectivity during development. However, reliable and consistent identification of tracts is often challenging due to motion artifacts and inherently low diffusion anisotropy in infant brain MRIs. In this paper we introduce a new atlas-based probabilistic tractography tool specifically designed for newborn infants. Our tool incorporates prior information on tracts and the underlying morphology from a training data set. In our experiments, we analyze a set of both full term and prematurely born infants and demonstrate that we are able to consistently, robustly and accurately recover known white-matter tracts in both groups. As such our tool paves the way for performing a host of sophisticated analyses in newborns that have been heretofore possible only in adults, such as pointwise analysis along tracts and longitudinal analysis, in both health and disease.

Keywords: probabilistic tractography, infant brain MRI, atlas-guided annotation, premature and full-term infants

1 Introduction

Robust and reproducible tools to study normal brain development and early brain injury are of great interest. During the later stages of pregnancy and in the early neonatal period the brain is developing very rapidly and is also very vulnerable to injury [1, 2]. Diffusion-weighted MRI can be used to characterize normal or atypical brain development, as myelination, pre-myelination, and white matter (WM) injury can be

characterized by diffusion parameters [3]. A major challenge in studying infant brains is the region-specific rates of development; different WM tracts, for instance, myelinate at different ages and rates. Region of interest (ROI)-based measurements have been used to study developmental changes and injury [2, 4]. However, this approach is highly susceptible to observer bias, inter-observer variability and is less anatomically meaningful, as it does not incorporate measurements along the entire length of a tract. Tractography-based experiments have shown improved reproducibility of diffusion metrics compared to ROI-based methods, suggesting that this approach may also provide more informative biomarkers when characterizing brain injury.

Tractography tools can be divided into local or global, depending on how they integrate the diffusion orientations across voxels, and deterministic or probabilistic, depending on how they model those orientations locally. Local ("streamline") methods grow a pathway one voxel at a time, relying on the local diffusion orientation around that voxel only. Global methods fit the pathway to the diffusion orientations over its entire trajectory, making it less sensitive to a localized region of high uncertainty (e.g., a crossing or lesion) somewhere along the path. Deterministic methods use the most probable diffusion orientation at each step of the reconstruction, while probabilistic ones also model the variability in the orientation. Probabilistic methods explore more connections but are very sensitive to the placement of the seed ROIs.

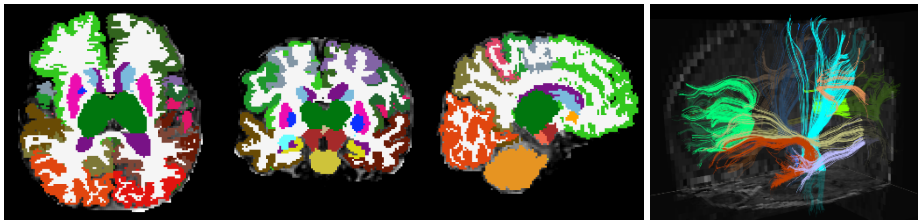

Fig. 1. (**Left**) Cortical and subcortical segmentation labels displayed on the structural MRI of an infant. (**Right**) Manually annotated tracts of a sample subject: CST (light blue), FMIN (green), FMAJ (bright green), IFOF (beige), ILF (purple), UF (red) and SLF (light brown).

One can annotate specific WM bundles manually on the output of whole-brain tractography, but this can be a biased and time-consuming process. Alternatively, there exist automated clustering tools that rely on either feature similarities among neighborhood tracts (such as shape and distance) or information from previously curated training data to identify tract bundles [5-7]. These tools are unfortunately often sensitive to distance metric definition and tract reconstruction in presence of WM anomaly, making point-to-point correspondence along the tractography solutions difficult to establish when comparing multiple subjects. In this paper, we propose to develop an atlas-based probabilistic tractography tool that is capable of consistently extracting a set of major WM bundles even in the case of lower image quality and potential structural abnormalities. Our approach is inspired by TRACULA, a publicly available tool for automated reconstruction of a set of major WM pathways in FreeSurfer [8]. It augments global probabilistic tractography with anatomical priors. Prior distributions on the neighboring anatomical structures of each pathway are derived

from an atlas and combined with cortical and sub-cortical segmentation labels to constrain the tractography solutions. It has been demonstrated that TRACULA can accurately reconstruct tracts in disease populations, even if the atlas includes only healthy subjects [9]. Although TRACULA and the FreeSurfer tools were designed and extensively tested on adult data sets, they have been successfully applied to children as young as 4.5 years of age [10]. However, due to vast morphological differences and reversed grey-white contrast, they are not directly applicable to our newborn population.

2 Methods

We have developed a framework where all computational components have been adapted or newly developed specifically to address the challenging image processing requirements of newborn MRI data. Our global tractography tool benefits from prior anatomical knowledge on the pathways that we want to reconstruct, unlike exploratory tractography methods that aim to discover which brain region is connected to which. Importantly, however, our model does not assume that the pathways have the same shape, size, or integrity in the study subjects as in the atlas subjects; only that they go through the same broad areas of anatomy.

Training data set (atlas) construction:

Structural data processing: In order to obtain the cortical and subcortical segmentation labels of anatomical areas of interest, which are included in the priors used by our tractography, we first used our novel double-consensus skull-stripping approach to identify the brain region. Even though it is a challenging task in the newborn population, with this tool we achieved over 90% Dice coefficient overlap measure when compared to expert-delineated brain masks. Then, for the segmentation, we implemented a Bayesian multi-label atlas fusion framework based on [11]. Instead of relying on the intensities of the training images, we used the consistency of voxel intensities within target regions and their relation to the propagated labels. This was critical for segmenting scans with varying degrees of myelination. In total, we identified 23 subcortical regions as well as 7 cortical areas per hemisphere (including the prefrontal, premotor, primary motor, primary somatosensory, posterior parietal, occipital, and temporal regions) (Fig. 1, left). For this automated segmentation task we relied on DRAMMS [12] to register our input data sets. In order to align the corresponding structural and diffusion acquisitions of a subject we applied FLIRT.

Manual tract annotation: We adapted rules from [13] to infants to annotate twelve major WM pathways, all crucial in development and relatively well-characterized by neuroanatomists: the forceps major and minor of the corpus callosum (FMIN, FMAJ), and bilateral corticospinal tract (CST), inferior longitudinal (ILF), inferior fronto-occipital (IFOF), superior longitudinal (SLF), and uncinate fasciculi (UF) (Fig 1, right). Our annotations were performed using Trackvis (http://trackvis.org) by a set of trained experts under the supervision of a neuroradiologist (CJ) with 5 years of experience in image analysis. In a subset of our data, we also estimated the inter-rater reli-

ability of the manual annotations. This resulted in a highly satisfying average Dice overlap score of 83.6%, indicating consistent reproducibility of our annotation rules.
Common spatial coordinate space: We established an unbiased common spatial coordinate space from our training subjects with a robust template construction tool using affine registration [14]. This serves as a reference space for the training data used in our tractography reconstruction, although the tractography itself takes place in the individual's native space. Our method does not rely on perfect spatial alignment between the individual and the atlas, because it uses the positions of the tracts relative to the segmentation labels, rather than their exact spatial coordinates. Thus an affine registration was sufficient for accurate tract reconstructions.

Automated tractography:
We incorporated the neonate segmentations and manually labeled tracts into a Bayesian framework, following the approach of TRACULA [8]. Our tool relies on global probabilistic tractography with prior distributions on the anatomical structures that each pathway intersects or neighbors, derived from the training data. In this framework the unknown pathway $\mathcal{F}$ in any new test subject is estimated from the diffusion-weighted images Y of that subject via the posterior probability distribution: $P(\mathcal{F}|Y) \propto P(Y|\mathcal{F})p(\mathcal{F})$. The likelihood term $P(Y|\mathcal{F})$ models the variability in the measured data given the shape of the pathway in the specific subject and the prior distribution $P(\mathcal{F})$ the variability in the pathway shape among subjects. We use the same formulation to compute the likelihood term as [15], which assumes Gaussian noise and uses the "ball-and-stick" model of diffusion. We used the *bedpostx* tool in FSL to estimate the distributions of the parameters of this model at each voxel given the diffusion data. Our departure from [15] is that, instead of assuming equal prior probability for all possible paths connecting two regions of interest, we use a prior of the form: $p(\mathcal{F}) = p(\mathcal{F}|A, \{F_k\}_{k=1}^n, \{A_k\}_{k=1}^n)$, where A is the anatomical segmentation map of the test subject, and $F_k, k = 1, \dots, n$ is the pathway of interest in each of the n training subjects, and $A_k, k = 1, \dots, n$ is the anatomical segmentation map of each training subject. Our reconstruction took 1-2 min per tract on a Linux computer with 2 Quad Core Xeon processors, 3.0 GHz CPU and 32GB of RAM.

Infant data sets: We used two sets of infant MRI data for our atlas building and validation. Full-terms: Ten healthy, full-term infants born at 38-41 weeks were scanned within the first week of life at Boston Children's Hospital using a Siemens 3T Skyra with the following sequences: motion-compensated multi-echo 1 mm^3 MPRAGE and 2 mm^3 diffusion (TE = 104 ms, TR = 3700 ms; b=1000 s/mm^2, 30 dirs) accelerated with 2xGRAPPA and simultaneous multi-slice factor of 2. Pre-terms: Fifteen premature subjects, at three different imaging centers, born at 27-32 weeks were imaged at 41 weeks on either a 3T Philips or a 1.5T Siemens scanner with an 8-channel coil: fast field echo (1.2 mm coronal slices, in-plane resolution 0.27 mm, TR = 18.9 ms, TE = 4.6 ms, flip angle 8°, T_{acq} = 2:34 min) and diffusion-weighted EPI (1 b = 0 volume, 32 b = 1000 volumes, 3 mm axial slices, in-plane resolution 1.56 mm, TR = 2.98 s, TE = 68 ms, T_{acq} = 2:35 min) in the case of the former and FLASH (4.0 mm sagittal slices, in-plane resolution 0.43 mm, TR = .38 ms, TE = 5.52 ms, flip angle 90°, T_{acq} = 2:40

min) and diffusion-weighted EPI (2 b = 0 volumes, 24 b = 1000 volumes, 2.2 mm axial slices, in-plane resolution 1.97 mm, TR = 7.6 s, TE = 98 ms, T_{acq} = 2:43 min) in the case of the latter.

3 Experiments and results

An important question for any atlas-based method is whether the inclusion of subjects from different populations affects the accuracy of the output. In the case of infants, in particular, where data is scarce and quality variable, it would be crucial to be able to pool data from different acquisition sites and studies in our atlas. To investigate this, we performed automated tractography on all subjects using three different training sets: full-term subjects only, pre-term subjects only, and all subjects. Whenever the test subject belonged to a population included in the training set, we removed the test subject from the training set. In the results we use "F", "P", and "A" to denote, respectively, the set of full-term only, pre-term only, or all subjects. Each experiment is denoted as t|T, where t the test set and T the training set (t, T ∈ {F, P, A}). We organize results into five experiments: F|F, A|A, P|P, F|P and P|F.

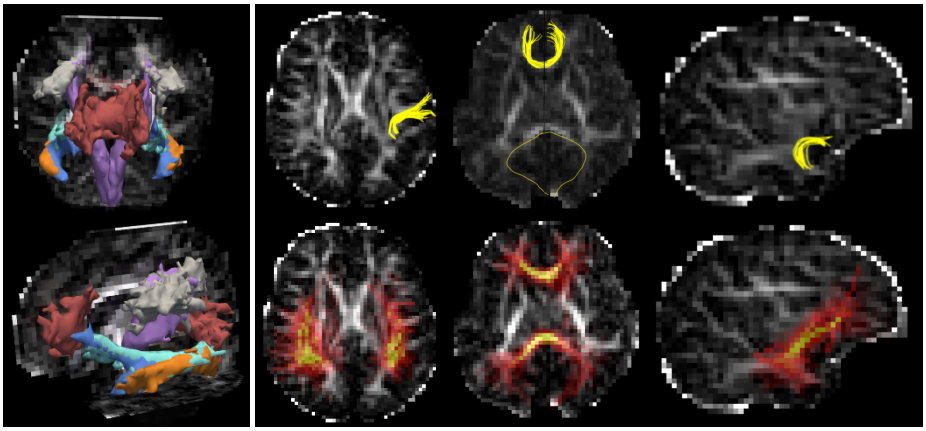

Fig. 2. (**Left**) Automatically reconstructed pathways in a premature subject. The estimated posterior distribution of each pathway is displayed as an isosurface in coronal and sagittal views: FMAJ / FMIN (red), CST (purple), ILF (orange), UF (blue), SLF (grey) and IFOF (cyan). (**Right - Top**) Example pathways difficult or impossible to obtain during manual annotation (**Right - Bottom**) Maximum-intensity projections of the automatically reconstructed posterior distributions of the same pathways.

Fig 2 displays, on the left, the full set of 12 automatically reconstructed pathways in a premature subject, using the remaining 24 subjects as the training set and, on the right, a set of examples of pathways that were difficult or impossible to obtain during manual annotation, but were successfully reconstructed with our automated approach. Quantitatively, we evaluated the accuracy of our automatically reconstructed pathways by comparing them to the corresponding manually labeled pathways from the

same subject. We quantified accuracy as the modified Hausdorff distance (MHD) between the automatically reconstructed pathway and its manually labeled equivalent. We define the MHD between two labels as the minimum distance of each point on one label from the other label, averaged over all points on the two labels. Before computing the MHD, the posterior probability distribution estimated by our automated tractography method was thresholded by masking out all values below 20% of the maximum. Thus the comparison is based on the center of the distribution and not its tails, as we expect the center and not the tails to overlap with the manual labels.

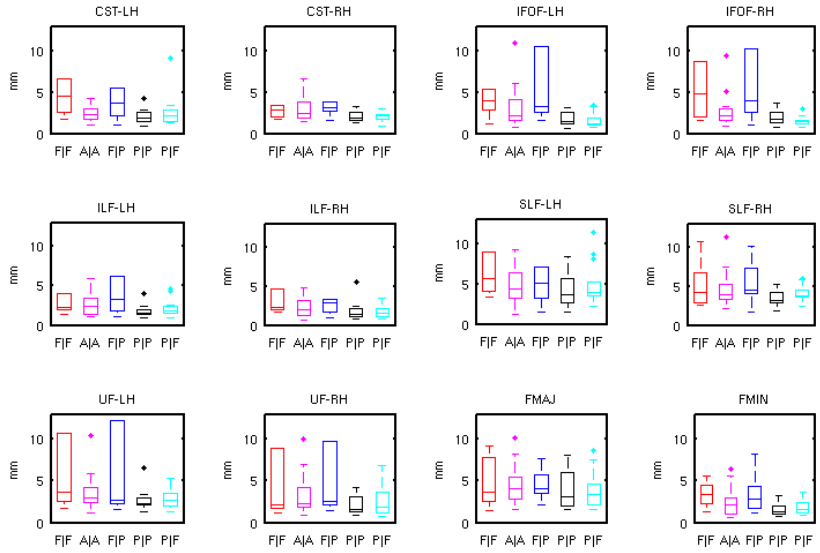

Fig. 3. Modified Hausdorff distances (in mm) between automatically reconstructed and manually labeled pathways. For all 12 pathways, we show results with 5 combinations of test and training data, each denoted as (test set) | (training set), F: full-term, P: pre-term, A: all subjects.

Figure 3 shows box plots of the MHD between automatically reconstructed and manually labeled tracts for each combination of test and training data. The horizontal line inside each box marks the median, whereas the edges of the box mark the 25th and 75th percentiles. The median MHD over all the analyses performed here was 2.61mm, and the interquartile range was 2.39mm. This indicates that, in most cases, the distance between automatically reconstructed and manually labeled tracts was less than 2 diffusion voxels. Note that some differences between the manual and automated approaches may be desirable. This would be the case, for example, when the manual approach cannot find a specific tract but the automated approach, aided by the training data and anatomical segmentation, can (see Fig 2, right).

A three-way analysis of variance on the distance, where the three factors were tract (12 levels), test set (2 levels: full-term or pre-term), and training set (3 levels: full-term, pre-term, or all) showed a significant effect of tract ($p=0.01$) and test set ($p<0.0001$), but not of training set ($p=0.95$). That is, the accuracy of the reconstruc-

tion depends on the test data but not on the training data. This implies that, for a given test subject, the accuracy is similar whether the atlas consists of full-term subjects only, pre-term subjects only, or subjects from both groups. As the full-term and pre-term subjects were scanned in different scanners with different acquisition protocols, we cannot determine if any differences in accuracy between test subjects from the two groups are due to data acquisition differences or population differences. However, an important conclusion that we can draw from these results is that it is possible to include subjects from different populations and acquisition sites in the training data of our algorithm without affecting its accuracy. This robustness to variability in the training data is due to the fact that our algorithm does not expect a specific image contrast, and does not rely on perfect alignment of subjects in a template space. The priors used in the algorithm involve only the position of each tract relative to a set of anatomical segmentation labels.

4 Discussion

We proposed a novel method for the automated reconstruction of WM tracts in newborn subjects. Even though short acquisition times and low contrast due to delayed or incomplete maturation make data from this population very challenging, our tool reconstructed the pathways of interest successfully. This work will make available, for the first time, a robust diffusion tractography tool that is applicable not only in the healthy control infant population, but also in subjects born prematurely. As part of this development, we also established, to the best of our knowledge, the largest manually annotated tractography data set for newborn infants.

In the future, we plan to run experiments with a more extensive training data set in order to establish the necessary and sufficient number of subjects that will be needed for our tool to be robust in a wide variety of settings. We are currently in the process of incorporating additional WM bundles to the set of 12 that we presented here, as well as extensively comparing the performance of our tool to that of publicly available clustering solutions. Upon completion, we plan to release the atlas along with the automated tractography tool for research use.

Acknowledgments. This work was funded by NIH R00 HD061485-03; EPIRMEX: National Hospital Clinical Research (PHRC) 2011; and Martinos Computing facilities: NIH S10RR023401, S10RR019307, S10RR019254, S10RR023043.

5 References

1. Dobbing, J., *The later growth of the brain and its vulnerability.* Pediatrics, 1974. **53**: p. 2-6.
2. Partridge, S.C., et al., *Tractography-based quantitation of diffusion tensor imaging parameters in white matter tracts of*

preterm newborns. J Magn Reson Imaging, 2005. **22**(22): p. 467-474.
3. Miller, S.P., et al., *Serial quantitative diffusion tensor MRI of the premature brain: development in newborns with and without injury.* J Magn Reson Imaging, 2002. **2002**.
4. Dudink, J., et al., *Fractional anisotropy in white matter tracts of very-low-birth-weight infants.* Pediatr Radiol, 2007. **37**: p. 1216-1223.
5. O'Donnell, L.J. and C.-F. Westin, *Automatic Tractography Segmentation Using a High-Dimensional White Matter Atlas.* IEEE Transac Medical Imaging, 2007. **26**(11): p. 1562-1575.
6. Wang, X., W.E.L. Grimson, and C.-F. Westin, *Tractography segmentation using a hierarchical Dirichlet processes mixture model.* NeuroImage, 2011. **54**(1): p. 290–302.
7. Guevara, P., et al., *Robust clustering of massive tractography datasets.* NeuroImage, 2011. **54**: p. 1975-1993.
8. Yendiki, A., et al., *Automated probabilistic reconstruction of white-matter pathways in health and disease using an atlas of the underlying anatomy.* Front. Neuroinform, 2011. **5:23**.
9. Koldewyn, K., et al., *Differences in the right inferior longitudinal fasciculus but no general disruption of white matter tracts in children with ASD.* Proc. Nat. Acad. Sci., 2014. **111**(5): p. 1981-6.
10. Ghosh, S.S., et al., *Evaluating the validity of volume-based and surface-based brain image registration for developmental cognitive neuroscience studies in children 4 to 11 years of age.* NeuroImage, 2010. **53**: p. 85-93.
11. Iglesias, J.E., M.R. Sabuncu, and K. Van Leemput, *A Generative Model for Multi-Atlas Segmentation across Modalities.* Proc IEEE Int Symp Biomed Imaging, 2012: p. 888-891.
12. Ou, Y., et al., *DRAMMS: Deformable registration via attribute matching and mutual-saliency weighting.* Medical Image Analysis., 2011. **15**(4): p. 622-639.
13. Wakana, S., et al., *Reproducibility of quantitative tractography methods applied to cerebral white matter.* NeuroImage, 2007. **36**(3): p. 630-644.

14. Reuter, M., H.D. Rosas, and B. Fischl, *Highly Accurate Inverse Consistent Registration: A Robust Approach.* NeuroImage, 2010. **53**(4): p. 1181--1196.
15. Jbabdi, S., et al., *A Bayesian framework for global tractography.* Neuroimage, 2007. **37**(1): p. 116-29.