Evaluation of White Matter in Preterm Infants With Fetal Growth Restriction

Olivier Commowick[1], Neil I. Weisenfeld[1], Heidelise Als[2], Gloria B. McAnulty[2], Samantha Butler[2], Lindsay Lightbody[2], Richard M. Robertson[1], and Simon K. Warfield[1]

[1] Department of Radiology, Children's Hospital Boston
Olivier.Commowick@childrens.harvard.edu
[2] Department of Psychiatry, Children's Hospital Boston
Harvard Medical School, Boston, MA, USA

Abstract. Disrupted placental function is the most common cause for Fetal Growth Restriction (FGR). This may lead to neurodevelopmental disabilities for example in motor skills and/or cognitive function. The infant's sensory experience in the newborn intensive care unit (NICU) may have an influence on the immature brain and alter its development. To support brain development, an approach called NIDCAP has been developed. It relies on observation and interpretation of the infant's behavior for adaptation of the environment and care implementation.
The NIDCAP program aims to support brain structure maturation, particularly in white matter. Diffusion tensor images (DTI) provide an indirect way of studying the white matter structure through the study of water molecules diffusion. Therefore our goal was to evaluate locally the whole brain's white matter in FGR preterm-born infants through the study of DTI populations. We therefore developed a new algorithm to identify group differences in white matter from DTI. A local implementation of the continuous STAPLE algorithm is utilized to build at the same time a robust DTI reference standard from the populations and parameters characterizing each image with respect to the reference standard. Subsequently the algorithm is coupled with a statistical test (Cramer's test) through the use of an appropriate distance on the STAPLE parameters to obtain a robust comparison algorithm.
We compared 20 FGR infants, aged 42 weeks post-menstrual age (PMA), randomized into two populations: 9 patients treated with NIDCAP and 11 with standard best care. We demonstrate that white matter structure appears more mature in the posterior limb of the right and left internal capsule in infants in the NIDCAP group. Thus, it appears that NIDCAP is effective in FGR preterm infants and that the new whole-brain DTI algorithm is sensitive to these differences.

1 Introduction

Placental function is a crucial component of the intrauterine environment, and when disrupted is the most common cause of fetal growth restriction (FGR).

Resulting neurodevelopmental disabilities have been identified in motor skills, cognitive function and school performance. These disabilities are especially pronounced in FGR preterm-born infants. The infant's sensory experience in the Newborn Intensive Care Unit (NICU), such as exposure to bright lights, high sound levels and frequent interventions, may affect the immature brain and alter its development. To decrease the discrepancy between the immature brain expectation and the environment in a NICU, and support brain development, an approach named Newborn Individualized Developmental Care and Assessment Program (NIDCAP) has been developed. It relies on the adaptation of care and environment based on the observation and interpretation of the infant's behavior.

The NIDCAP program aims to support brain structure maturation. An earlier study randomized controlled trial of NIDCAP effects on brain development of appropriate in growth for gestational age (AGA) preterm-born infants, and showed more mature fiber structure in the NIDCAP group compared with the control group, who received standard care [1]. This study's DTI measures however were based on scalar parameters derived from the tensors, thereby losing some information, and were performed only on some manually delineated regions of interest. Moreover, errors in acquisitions such as movement artifacts or distortion might introduce bias in the studied images, especially in the case of newborn brain images. To test NIDCAP effectiveness in the white matter structure maturation, we set out to accomplish local evaluation of whole brain white matter in FGR preterm-born infants. Moreover, to take into account the aforementioned sources of errors, it is important to extend the investigation to whole brain automatic evaluation in order to detect all potential differences.

In order to solve these problems, we propose an algorithm that relies on the whole diffusion tensor information to detect locally significant differences between two populations of DT images. The algorithm uses a geometrically unbiased reference frame computed from the DTI of the two populations, coupled with recent methods for the detection of group differences between two tensor populations [2]. We furthermore improved the robustness of this approach by introducing a local continuous STAPLE algorithm to take into account locally the sources of errors coming from the image acquisition or registration errors when building the reference frame.

We searched for white matter differences in 20 FGR infants, aged 42 weeks post-menstrual age (PMA), randomized to a NIDCAP group (9 patients) or standard care group (11 patients). The results of the investigation demonstrated significant differences in regions that are developing rapidly at this age, namely the posterior limb of the internal capsule, right and left, which showed an increase in maturation of the white matter structure in the experimental group, when compared to the control group.

2 Methods

In order to compare the DTI from the two populations, we need to bring all the images into the same reference frame. Using any subject as a reference or an

external atlas as reference would introduce a bias due to the specific anatomy of the respective reference image. Therefore, we chose to build a geometrically unbiased DTI atlas from the data of the two groups to serve as a reference standard for comparison (see section 2.1). In a second step, we computed robust statistics for the detection of differences between the registered DT images (see sections 2.2 and 2.3).

2.1 Atlas Construction for Group Differences Evaluation

We based the construction of the geometrically unbiased atlas on the method proposed by Guimond et al. [3]. This algorithm iterates over the following two steps until convergence. First, all images I_k are non linearly registered onto a reference image R, using a block-matching based non linear registration method [4]. These registrations produce a set of deformation fields T_k. An average image M is then built from all registered images by taking the mean of all images. In a second step, the deformation fields T_k are averaged to get a mean transformation $\overline{T}$, which inverse is applied to M. This new average image is then used as the new reference image for the next iteration, i.e. $R = M \circ \overline{T}^{-1}$. We used the Log-Euclidean framework for diffeomorphisms [5] when averaging deformation fields as it provides a fast approach to compute the average transformation $\overline{T}$ in the manifold of diffeomorphisms. Transformations were calculated based on DT images using an extension to tensor based registration of our block-matching based non linear registration algorithm also used in [4].

When warping diffusion tensor images, special attention must be paid to the reorientation of tensors. In our registration algorithm, tensors are reoriented using the finite strain reorientation strategy [6]. Tensors were resampled and averaged using Log-Euclidean metrics [7] as they provide a fast and accurate way to interpolate tensors. In summary, a set of transformations T_k was obtained for each DT image I_k along with an average DTI $\bar{I}$.

2.2 A Local Continuous STAPLE for the Robust Comparison of Vector Images

Once the images were transformed into a common reference frame, we sought to detect local differences between the two DTI populations. For this step, it is important to take into account several factors that may introduce bias and variance in the analysis of these images. First, the quality of the construction of the reference frame highly depends on the non linear registration of the images. However, differences in acquisition protocols or anatomies may lead to registration discrepancies, therefore resulting in a bias on the transformed DTI. Further sources of bias and variance are the artifacts due to abnormal inter-subject brain differences, and as indicated earlier, due to acquisition problems, for example movement artifacts or DTI distortion.

In previous work [8], we have introduced a robust algorithm based on predefined regions of interest, called continuous STAPLE. It is based on an Expectation-Maximization algorithm, iterating over two steps:

- Estimation of unbiased reference tensors from the input DT images and current parameters (see Eqs. (9) and (10) in [8]),
- Computation of a bias vector β_j and covariance matrix Λ_j for each input image (see Eqs. (13) to (15) in [8]), which describe a multivariate Gaussian distribution $N(\beta_j, \Lambda_j)$ characterizing their adequacy with respect to the reference tensors (see Eq. (1) in [8]).

This method is of interest for the robust comparison of images, for which voxels may be represented as vectors, such as DTI in the Log-Euclidean space [7]. In our case however, we wished to compute local differences between populations. Moreover, the continuous STAPLE algorithm was proposed as a way to compare a patient with respect to a group of images. We however sought to compare populations of images. To solve these problems, we therefore developed a new method called local continuous STAPLE. This method allows computation of unbiased reference tensors as well as parameters in a voxel-wise manner. Below, we discuss the combination of the obtained parameters with a statistical test to use them for robust group comparison.

The voxel-wise computations were performed with a sliding-window strategy. A block was defined around each voxel in the image. For each of these blocks, the reference tensors and the parameters were then estimated and their values kept only for the voxel of interest.

2.3 Combining Local Continuous STAPLE and Statistical Tests

The final step was then to compare the parameters in order to obtain a robust detection of DTI differences. One algorithm for such a comparison, called the Cramer's test, has been proposed for DTI comparison by Whitcher et al. [2]. It evaluates for each voxel the Cramer's statistic, based on the computation of inter- and intra-group distances between the objects in the two groups:

$$T_{n_1,n_2} = \frac{n_1 n_2}{n_1+n_2}\left[\frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} d(P_{1,i},P_{2,j}) - \frac{1}{2n_2^2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_1} d(P_{1,i},P_{1,j}) - \frac{1}{2n_1^2}\sum_{i=1}^{n_2}\sum_{j=1}^{n_2} d(P_{2,i},P_{2,j})\right] \quad (1)$$

where n_1 and n_2 are the numbers of images in each group, $d(A,B)$ is a distance to be defined (in the case of tensors, Whitcher et al. proposed the use of the Euclidean distance on Log-Euclidean representation of tensors [7]), $P_{1,.}$ is a sample of the first group, $P_{2,.}$ is a sample of the second group. This statistic is computed for permutations of group compositions in order to compute its probability density function (PDF) and deduce from it a p-value describing the likelihood of the existence of a difference between the two specified groups. This algorithm is particularly valuable as it does not assume any PDF for the statistic but builds

it by testing permutations. Moreover, it may be used on the tensors themselves using the Log-Euclidean framework on tensors [7] but also to any data, as long as a distance may be defined between two values.

Since our local continuous STAPLE provides bias and covariance parameters for each subject, we chose to identify significant differences with the Cramer's test applied to these local parameters. The two groups, if different, should indeed have distinct distributions of parameters. Moreover, comparing these parameters will allow to take into account the inter-individual variability as well as the potential erroneous tensors and therefore obtain robust statistics. Each set of parameters in continuous STAPLE (bias β_j and covariance matrix Λ_j) represents a multivariate Gaussian $N(\beta_j, \Lambda_j)$. To apply the Cramer's test to this data, we defined an appropriate distance d between these Gaussian distributions. Such a distance has been proposed by Calvo and Oller [9] as an explicit solution of the geodesic between multivariate normal distributions (Eq. (18) in [9]). We therefore applied this distance in the Cramer's test statistic for comparison of the Gaussian parameters between groups.

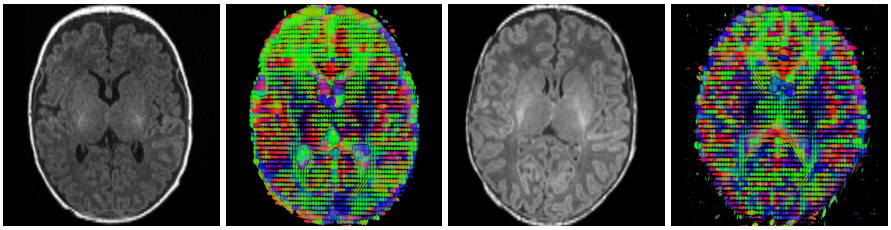
3 Results

3.1 Image Database

For this study, we have used a database of 20 FGR preterm infants, born between 26 and 34 weeks gestational age (GA) at birth, with birth weights and head circumferences $< 5^{th}$ percentile, and delivered after Doppler diagnosed compromised umbilical artery blood flow, randomized to standard care group (first group, 11 infants) or NIDCAP group (second group, 9 infants). All infants were assessed at 42 weeks post-menstrual age (PMA), i.e. 2 weeks after their expected due date, had they not been born prematurely.

Images were acquired at two time points on a 1.5T scanner: the first at birth and the second at 42 weeks PMA. For each patient, structural MR images (T1 SPGR, and T2 FSE) were acquired in the coronal plane (slice thickness 1.3 mm, matrix-size 256x256, in-plane resolution 0.78125 mm), as well as diffusion weighted images (slice thickness 2.5 mm, matrix size 256x256, b value of 750 s/mm^2) with between 6 and 35 directions acquired depending on the cooperation of the patients. Diffusion tensor images were then computed from the DWI. Examples of the images from each group are illustrated in Fig. 1.

Preprocessing: For each infant, a brain mask was computed from the aligned T1 and T2 images using newborn atlas-based segmentation methods. Then, the DTI was aligned on the T1 image. First, the B0 image was rigidly registered to the T1 image, and the obtained transform was applied to the diffusion tensors. Then, to correct for acquisition distortion in DTI, a non rigid registration was applied between the B0 image and the T1 using a B-Splines method [10] with few control points ($5 \times 5 \times 5$).

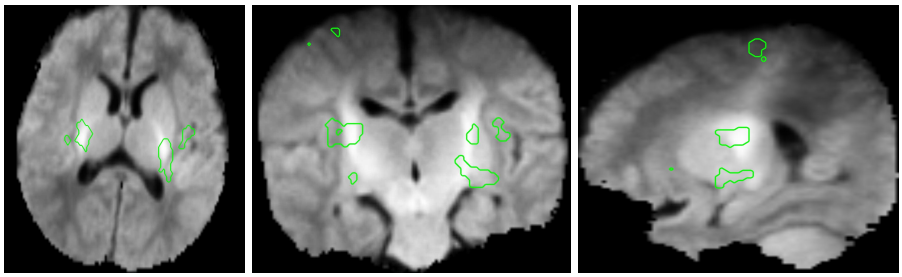


(a) (b) (c) (d)

Fig. 1. Examples of Patients from our Database. Axial slices showing T1 images and DTI for each of group ((a, b): controls group, (c, d): NIDCAP group).

3.2 Evaluation of Brain Differences after Treatment

Differences were evaluated between the two populations at time point 2 (42 weeks PMA). Fig. 2 illustrates the contours of regions where differences were detected at the 95% significance level (to obtain the contours, the p-value map was thresholded, very small regions were removed as they were likely due to artifacts, and an isosurface was extracted).



(a) (b) (c)

Fig. 2. Differences Detection Results. Illustration of the differences detected using the Cramer's test combined to the local continuous STAPLE. The contours are obtained after thresholding the p-value maps at the 95% confidence level, removing very small regions and extracting an isosurface.

Significant differences have been detected between the groups of patients mainly in the posterior limb of internal capsule (PLIC) on both sides, which correspond to areas related to motor functions (cortico-spinal tracts). These results are consistent with those obtained in the previous study [1] based on the comparison with the AGA preterm-born infants. This finding is also consistent with the fact that these internal brain regions are maturing in the last trimester of pregnancy, and therefore may be affected by the difference in care received.

To provide further insight into the manner in which the detected regions are different, we have performed a study of scalar parameters derived from the ten-

sors. We have therefore compared average mean diffusivity (MD) and fractional anisotropy (FA) values on the regions detected using the modified Cramer's test. An analysis of variance (ANOVA) was performed for FA and MD (using the *anova1* function in Matlab). Comparison of the average MD and FA values for the regions detected showed no significant difference in the FA values ($p = 0.265$). However, as presented in Fig. 3, which shows the box plot of the MD values for the two groups, MD values were significantly decreased ($p = 0.045$) in the NIDCAP group compared to the controls. The maturation of brain structure may therefore be accelerated for the NIDCAP group infants in the detected regions, therefore decreasing the MD.

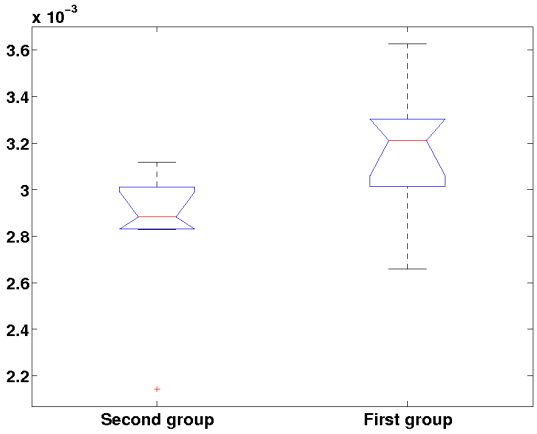


Fig. 3. Illustration of the ANOVA Study on MD Values. First group corresponds to the controls group and second group to the NIDCAP group. The two groups are significantly different in terms of MD ($p = 0.045$).

4 Conclusion

We have introduced in this study a new algorithm to detect, in a robust manner, significant differences between DTI of two groups. This approach is based on a novel local continuous STAPLE algorithm extending a previous algorithm [8]. We demonstrated the local estimation of a reference standard diffusion tensor together with a local log-Euclidean tensor mean and a variance-covariance matrix for every subject describing the difference between each subject and the reference standard. This provides a local Gaussian model for the difference between each subject and the reference standard. We then formulated a novel test for group differences using a non-parametric permutation test based on Cramer's test [2] utilizing the Colver-Oller distance between Gaussians [9].

We applied this new approach to the evaluation of the influence of individualized developmental care (NIDCAP) on the white matter maturation of FGR preterm-born infants. The results show significant differences in the posterior limb of the internal capsule on both sides of the brain. We showed that these differences are likely related to an increased maturation of the brain, leading to a significant decrease in MD. This suggests a potential positive influence of the NIDCAP treatment on early brain development.

Future work should include the search for significant differences from the intrauterine stage of brain development e.g. shortly after the birth of these FGR preterm infants in order to characterize more fully the progression of brain structural development in the NICU with and without NIDCAP. Further insight will be gained by coupling the results with studies performed from other modalities, such as the comparison of structures volumes extracted from T1 images, gyrification indexes and cortical thickness comparison. Furthermore, white matter evolution might be tracked by white matter tractography studies. Correlation with behavioral and EEG findings in combination of all these parameters will allow for better understanding of the cortical regions influenced by the treatment.

Acknowledgments

This investigation was supported in part by a research grant from CIMIT and by NIH grants R01 HD46855, R03 CA126466, R01 RR021885, R01 GM074068 and R01 EB008015.

References

1. Als, H., Duffy, F.H., McAnulty, G.B., et al.: Early experience alters brain function and structure. Pediatrics **113**(4) (Apr 2004) 846–857
2. Whitcher, B., Wisco, J.J., Hadjikhani, N., Tuch, D.S.: Statistical group comparison of diffusion tensors via multivariate hypothesis testing. MRM **57** (2007) 1065–1074
3. Guimond, A., Meunier, J., Thirion, J.: Average brain models: A convergence study. Computer Vision and Image Understanding **77**(2) (2000) 192–210
4. Commowick, O., Malandain, G.: Evaluation of atlas construction strategies in the context of radiotherapy planning. In: SA2PM Workshop, MICCAI 2006. (2006)
5. Arsigny, V., Commowick, O., Pennec, X., Ayache, N.: A Log-Euclidean framework for statistics on diffeomorphisms. In: MICCAI (I). LNCS (4190) (2006) 924–931
6. Ruiz-Alzola, J., Westin, C.F., Warfield, S.K., Alberola, C., Maier, S., Kikinis, R.: Nonrigid registration of 3D tensor medical data. MedIA **6**(2) (June 2002) 143–161
7. Arsigny, V., Fillard, P., Pennec, X., Ayache, N.: Log-Euclidean metrics for fast and simple calculus on diffusion tensors. MRM **56**(2) (2006) 411–421
8. Commowick, O., Warfield, S.K.: A continuous STAPLE for scalar, vector and tensor images: An application to DTI analysis. IEEE TMI **28**(6) (2009) 838–846
9. Calvo, M., Oller, J.: An explicit solution of information geodesic equations for the multivariate normal model. Statistics and Decisions **9** (1991)
10. Rueckert, D., Sonoda, L.L., Hayes, C., et al.: Nonrigid registration using free-form deformations: Application to breast MR images. IEEE TMI **18**(8) (1999) 712–721