A Whole Brain Morphometric Analysis of Changes Associated With Pre-Term Birth

C.E. Thomaz[1], J.P. Boardman[3], S. Counsell[3], D.L.G. Hill[4], J.V. Hajnal[3], A.D. Edwards[5], M.A. Rutherford[3], D.F. Gillies[2], and D. Rueckert[2]

[1]Department of Electrical Engineering, Centro Universitario da FEI, Sao Paulo, Brazil
[2]Department of Computing, Imperial College, London, UK
[3]Imaging Sciences Department, Imperial College, London, UK
[4]Centre for Medical Image Computing, University College London, UK
[5]Division of Paediatrics, Obstetrics and Gynaecology, Imperial College, London, UK
cet@fei.edu.br

Abstract. Pre-term birth is strongly associated with subsequent neuropsychiatric impairment. To identify structural differences in preterm infants we have examined a dataset of magnetic resonance (MR) images containing 88 preterm infants and 19 term born controls. We have analyzed these images by combining image registration, deformation based morphometry (DBM), multivariate statistics, and effect size maps (ESM). The methodology described has been performed directly on the MR intensity images rather than on segmented versions of the images. The results indicate that the approach described makes clear the statistical differences between the control and preterm samples, showing a leave-one-out classification accuracy of 94.74% and 95.45% respectively. In addition, finding the most discriminant direction between the groups and using DBM features and ESM we are able to identify not only what are the changes between preterm and term groups but also how relatively relevant they are in terms of volume expansion and contraction.

1 Introduction

Pre-term birth is associated with neuropsychiatric impairment that is more severe with more prolonged exposure to the extrauterine environment [2]. Half of all surviving infants born at 25 weeks or less show neurodevelopmental impairment at 30 months of age [16], and among less immature infants over one third develop neurocognitive and behavioral problems [2, 10, 11]. The most common known cerebral abnormality in surviving preterm infants is non-localized white matter disease seen in about two thirds of survivors and increased apparent diffusion coefficient values on diffusion weighted imaging [4, 9]. The effect of this diffuse lesion on neural systems is unknown, but defining its neuroanatomical correlates could advance understanding about the neural basis of functional impairment.

To identify structural differences in preterm infants we have examined an extensive dataset of magnetic resonance (MR) images containing 88 preterm infants and 19 term born controls. We have analyzed these images by combining image registration [5, 12], deformation based morphometry [1, 6], multivariate statistics [14], and effect size maps [8]. In this way, differently from analyzing the morphological differences given by the average volume change described by each sample group, we are able to analyze regional brain contraction and expansion of statistical extremes of preterm and term born infants and make a non-hypothesis survey of the whole brain to identify structural alterations apparent at term equivalent age.

2 Subjects and Data Acquisition

The MR images of the 88 preterm infants (mean 29.8, range 24-34 weeks post-menstrual age) were acquired at term equivalent age (38 to 42 weeks), together with 19 term-born infants (mean 39.0, range 36-43 weeks post-menstrual age). All images were reviewed by a perinatal neuroradiologist (MR) and infants with focal lesions or congenital cerebral malformations were excluded. The mean birth weight of the preterm at term equivalent group was 1302g (range 610-2178g), compared with a mean birth weight of 3390g (range 2448-4780g) among the control group. Preterm infants at term equivalent age were sedated with chloral hydrate for the examination but did not require mechanical ventilation at the time of MR imaging. Control infants were examined in natural sleep. Pulse oximetry, electrocardiographic and televisual monitoring were used throughout the examination which was attended by a neonatologist.

A 1.5 T Eclipse MR System (Philips Medical Systems, Best, The Netherlands) was used to acquire high resolution T1 weighted images (TR=30ms, TE=4.5ms, flip angle = 30^{o}). In addition to conventional T1 weighted image acquisition, volume datasets were acquired in contiguous sagittal slices (in-plane matrix size 256 x 256, FOV = 25cm) with a voxel size of 1.0 x 1.0 x 1.6 mm3. The skull, scalp, and other extra-cranial tissues were removed from each image before performing the analysis. Ethical permission for this study was granted by the Hammersmith Hospital Research Ethics Committee and informed parental consent was obtained for each infant.

3 Image Registration

To map the anatomy of each subject into the anatomy of the reference subject we used a non-rigid registration algorithm [12] that has been successfully applied to registration tasks involving deformable human tissue [5]. This algorithm uses a combined transformation T, which consists of a global transformation and a local transformation. The global transformation, represented by an affine transformation, models global differences in rotation, scaling, shearing and skew between the reference and subject images. The local transformation describes any local deformation required to achieve alignment of anatomical structures across subjects. This is modelled using a free-form deformation (FFD) based on B-splines, which is a powerful tool for modelling 3D deformable objects. In essence FFDs deform an object by manipulating an underlying mesh of control points. The resulting deformation controls the shape of the 3D object and can be written as the 3D tensor product of the 1D cubic B-splines as follows

$$T_{local}(x) = \sum_{s1=0}^{3}\sum_{s2=0}^{3}\sum_{s3=0}^{3} B_{s1}(u)B_{s2}(v)B_{s3}(w)c_{i+s1,j+s2,k+s3}\ , \qquad (1)$$

where c denotes the control points which parameterise the transformation.

The optimal transformation is found by minimizing a cost function associated with the global transformation parameters as well as the local transformation parameters. The cost function comprises two competing goals: the first term represents the cost associated with the voxel-based similarity measure, in this case normalized mutual information, while the second term corresponds to a regularization term which constrains the transformation to be

smooth. The resulting transformation T maps each point in the anatomy of the reference subject to the corresponding point in the anatomy of the subject.

4 Deformation-based Morphometry (DBM) and Volume Change

An MR image taken from a term born control infant that was normal by visual inspection was selected as the reference subject. The MR datasets from 88 preterm at term equivalent age infants and 19 term born controls were brought into alignment with the reference subject using the non-rigid registration algorithm with a multi-resolution scheme of control spacings of 20mm, 10mm, 5mm and 2.5mm. The resulting control point mesh defines a continuous and analytic representation of the deformation field which describes the point-wise 3D displacement vectors that are required to warp each dataset to the reference image. Boardman et al. [3] have previously shown that the algorithm performs consistently irrespective of the chosen reference subject. In all cases registration achieved a visually plausible alignment of anatomical structures. There were 7 subjects, in addition to the 88 preterms used, which were excluded due to misregistration (motion artefact in raw image or representing a structural extreme).

To calculate regional volume changes, the determinant of the Jacobian of the deformation field was used to quantify local volume differences between the registered images and the reference subject [13]. The determinant of the Jacobian for any given location x in the reference coordinate system provides an estimate of the point-wise volume change for each subject with respect to the reference. Values above 1 indicate tissue expansion and values below 1 indicate tissue contraction. As pointed out by Studholme et al. [13], assuming an unbiased registration, the choice of reference anatomy simply defines the origin of comparison of the relative size of local anatomy among subjects, and does not directly influence a statistical-based comparison of populations.

5 Multivariate Statistical Analysis

The volume change maps for subjects in each group were analysed using the general multivariate statistical methodology (PCA+LDA) to identify the most discriminating hyper-plane separating two populations. We used a recent approach proposed by Thomaz et al. [14] to overcome the well-known instability of the LDA within-class scatter matrix and increase the computational efficiency of the analysis.

5.1 Principal Components Analysis (PCA)

There are a number of reasons for using PCA to reduce the dimensionality of the 3D MR volume changes. PCA is a linear transformation that is not only simple to compute and analytically tractable but also extracts a set of features that is optimal with respect to representing the data back into the original domain. Moreover, using PCA as an intermediate step will reduce dramatically the computational and storage requirements for the subsequent LDA-based covariance method. Since our application is a small sample size problem where the number of training patterns N (or images) is much smaller than the number of features n (voxel-wise volume change map), it is possible to transform data in a way that patterns occupy as compact regions in a lower dimensional feature space as possible with far fewer degrees of freedom to estimate.

Although much of the sample variability can be accounted for by a smaller number of principal components, and consequently a further dimensionality reduction can be accomplished by selecting the principal components with the largest eigenvalues, there is no guarantee that such additional dimensionality reduction will not add any artefacts on the images when mapped back into the original image space. Since one of our main concerns here is to map the classification results back to the image domain for further visual interpretation, we must be certain that any modification on the images, such as blurring or subtle differences, is not related to an "incomplete" or perhaps "misleading" feature extraction intermediate procedure. Therefore, in order to reproduce the total variability of the preterm and control samples, we composed the PCA transformation matrix by selecting all principal components with non-zero eigenvalues. In other words, assuming that all the N training patterns are linearly independent the rank of the total covariance matrix is $N-1$ and the number of principal components selected is $m = N-1$.

5.2 Maximum Uncertainty LDA (MLDA)

The primary purpose of LDA is to separate samples of distinct groups by maximising their between-class separability while minimising their within-class variability. It is well known, however, that the performance of the standard LDA can be seriously degraded if there are only a limited number of total training observations N compared to the dimension of the feature space m. Since the within-class scatter matrix S_w is a function of $(N-g)$ or less linearly independent vectors, where g is the number of groups, its rank is $(N-g)$ or less. Therefore in the problem under investigation where the number of training patterns is comparable to the number of features, S_w might be singular or mathematically unstable and the standard LDA cannot be used to perform the task of the classification stage.

In order to avoid both the singularity and instability critical issues of the within-class scatter matrix S_w when LDA is used in such limited sample and high dimensional problem, we used a maximum uncertainty LDA-based approach (MLDA) based on a straightforward covariance selection method for the S_w matrix [15]. Thomaz et al. have shown recently that the MLDA approach provides an appropriate separation between the control and preterm voxel-based features without over-fitting the limited sample data [14].

6 Experimental Results

We performed two main tasks: classification and visual analysis. In the classification task the PCA+MLDA linear classifier was generated and the leave-one-out method was used to test and evaluate the classifier's performance. The main characteristic of the classification step is to find the most discriminant hyper-plane that not only separates the control from preterm infants but also indicates the statistical direction where patterns of morphological group differences are most expressive.

Since the dataset under investigation comes with a higher proportion of preterm images relatively to the control ones it is important to assess the PCA+MLDA classifier using different prior probabilities for the control and preterm groups as well as different misclassification costs. In order to avoid the specification of these values that are not known in practice, we used the equivalent threshold-free ROC (Receiver Operating Characteristic) approach that examines the effect of setting the decision line at various points between the samples and consequently varying the bias of the classifier.

Varying the decision threshold such that a given example will be classified as either control or preterm when its discriminant feature is compared to the threshold value leads to a tradeoff between sensitivity and specificity. The ROC curve is a plot of sensitivity (or true positive rate) versus 1–specificity (or false positive rate). In our problem, the true positive rate is the fraction (or %) of control labelled test data correctly classified as control. Analogously, the false positive rate is the fraction of truly preterm examples incorrectly classified as control. The ROC curve of the PCA+MLDA classifier using the leave-one-out method is plotted in Figure 1. It highlights the effect of setting the decision line at various points from the control to the preterm samples. When 0% of false positive rate is assumed, that is all the preterm examples are correctly classified, the PCA+MLDA classifier achieves a top of 89.47% of controls correctly classified. For a false positive rate of 4.55%, or preterm classification accuracy of 95.45%, the approach achieves a control classification accuracy of 94.74%. These results confirm the PCA+MLDA classifier's ability to discriminate brains of controls from those of preterm infants with high classification rates.

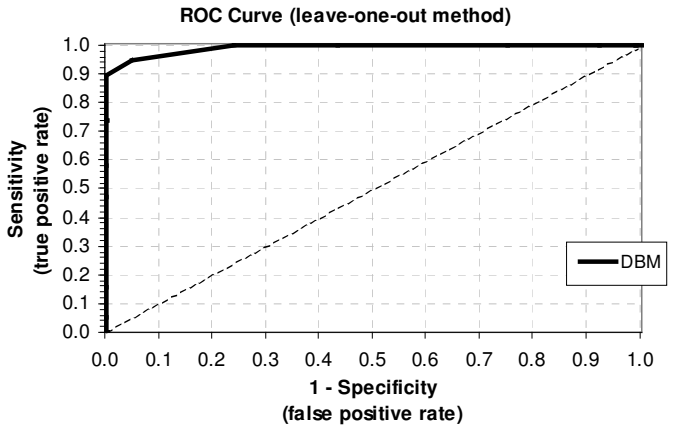

Fig.1. ROC curve based on leave-one-out method.

Figure 2 presents the PCA+MLDA most discriminant feature of the infant database using all the control and preterm samples as training data. The vertical axis is illustrative only and represents the corresponding index of each image in the sample set. Circles and crosses describe respectively the control and preterm observations. The post-menstrual age (PMA) at birth and scan of the infants on the extremes of correct classification and near the boundary of misclassification are also presented in the plot. For instance, the control infant on the right most place of the plot has 41.86 and 42.00 PMAs at birth and scan respectively. Analogously, the nearest preterm infant to the control group has 34.00 and 41.29 PMAs at birth and scan respectively. As reference values, we have indicated in the plot with arrows the infants of each group with maximum and minimum PMAs at birth and scan. As can be seen, when we move from the preterm (left) to the control (right) samples the infants on the extremes of each group are not the ones with extreme ages.

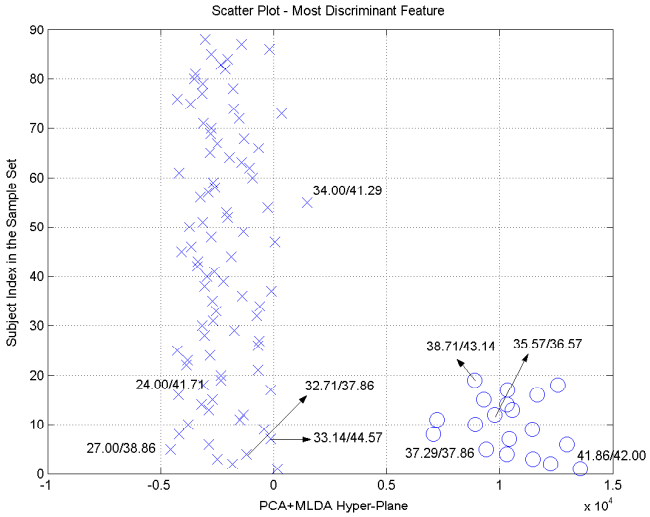

Fig. 2. Scatter plot of the most discriminant feature (all controls and preterms as training data).

The other main task that has been carried out by this two-stage multivariate statistical classifier is to visually analyse the most discriminant feature described by the PCA+MLDA hyper-plane. Assuming that the clouds of the control and preterm points follow a multidimensional Gaussian distribution and applying limits to the variance of each cloud, such as ± 3 standard deviations, we can move along this most discriminant feature and map the result back into the image domain. We can understand this mapping procedure as a way of defining volume changes that come from "definitely preterm" and "definitely control" samples, and consequently investigating differences captured by the classifier that are beyond the average volume change described by each sample group.

In order to avoid the use of raw units, which are quite arbitrary or lack meaning outside the investigation, we have expressed regional volume changes between two groups in units of variability by calculating the effect size $\varepsilon(x)$ of the differences, that is

$$\varepsilon(x) = \frac{v_p(x) - v_c(x)}{\sigma_{p \cup c}(x)}. \qquad (2)$$

Here $v_p(x)$ and $v_c(x)$ denote the volume changes at x for "definitely preterm" and "definitely control" samples (statistical extremes) calculated at 3 standard deviations from each corresponding sample group on the most discriminant feature space, while $\sigma_{p \cup c}(x)$ denotes the standard deviation of the volume changes at x for the corresponding pooled group.

Figure 3 illustrates the spatial distribution of the effect size $|\varepsilon(x)| \geq 0.5$ of the volume changes in both statistical extremes superimposed on a reference image. In red the tissues contained within the isoline are more contracted in the preterm at term equivalent age group compared to the term controls. The isolines in blue show areas of relative tissue expansion in the preterm group compared to the term infants.

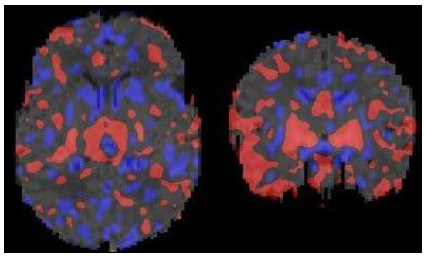

Fig.3. Detection of volumetric changes associated with preterm birth ($|\varepsilon(x)| \geq 0.5$).

Figures 4 and 5 show respectively the effect size maps (ESMs) of the volumetric contraction and expansion areas previously illustrated. In both figures the colour-scale shows relative volume change as a range of the effect size $\varepsilon(x)$. That is, tissues contained within the isoline in the spectrum of blue show areas of relatively small effect size $0.5 \leq |\varepsilon(x)| < 1.0$. The isolines in the spectrum of green and red show areas with more expressive changes and correspond to $1.0 \leq |\varepsilon(x)| < 2.0$ and $|\varepsilon(x)| \geq 2.0$ respectively. As can be seen, the area of most volume loss is within basal ganglia and thalami. This most expressive result is consistent with previous results published by other investigators [3, 7] and the less expressive ones highlight patterns of morphological group differences that are often not detectable.

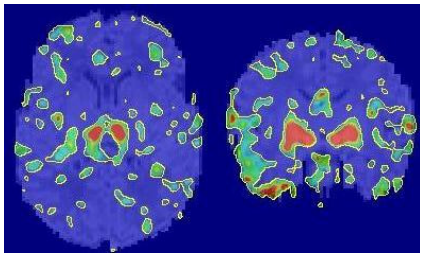

Fig.4. ESM - Contraction areas in detail.

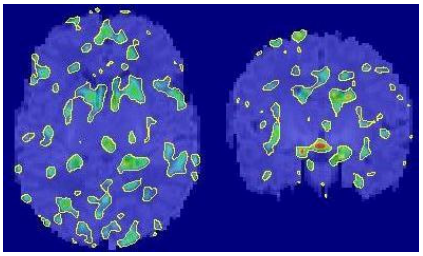

Fig.5. ESM - Expansion areas in detail.

7 Conclusion

This paper describes the idea of using deformation-based morphometry (DBM), multivariate statistical hyper-plane (PCA+MLDA), and effect size maps to analyse changes associated with pre-term birth. The methodology described has been performed directly on the MR intensity images rather than on segmented versions of the images. The pivotal step of this multivariate statistical analysis is registration.

Using a neonatal MR brain data set that contains images of 88 preterm infants at term equivalent age and 19 term controls, the model described makes clear the statistical differences between the control and preterm samples, showing a leave-one-out classification accuracy of 94.74% and 95.45% respectively. In addition, finding the most discriminant direction between the groups and using DBM features and effect size maps we are able to identify not only what are the changes between preterm and term groups but also how relatively relevant they are in terms of volume expansion and contraction. We believe that

applying such multivariate statistical analysis we can detect subtle and perhaps spatially more complex patterns of morphological group differences that are often not detectable.

Since the methodology described is not restricted to any particular set of features, there is thus scope for augmenting or replacing the input features. We are currently investigating and comparing the performance of the multivariate statistical model on voxel and deformation-based morphometries. Initial experiments carried out have shown complementary results between the two methods of morphometries.

Acknowledgments. This work is part of the UK EPSRC e-science project "Information eXtraction from Images" (IXI).

References

1. Ashburner J., Hutton C., et al. Identifying global anatomical differences: Deformation-based morphometry. *Hum Brain Mapp* 6, 638-57, 1998.
2. Bhutta,A.T., Cleves,M.A., Casey,P.H., Cradock,M.M. & Anand,K.J. Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. *JAMA* 288, 728-737, 2002.
3. J. P. Boardman et al. "An evaluation of deformation-based morphometry applied to the developing human brain and detection of volumetric changes associated with preterm birth", In Proc. *MICCAI*, LNCS 2878, pp. 697-704, 2003.
4. Counsell,S.J. *et al.* Diffusion-weighted imaging of the brain in preterm infants with focal and diffuse white matter abnormality. *Pediatrics* 112, 1-7, 2003.
5. E. R., Denton, "Comparison and evaluation of rigid, affine, and nonrigid registration of breast MR images", *J. Comput. Assist. Tomogr.* 23, pp. 800-805, 1999.
6. Gaser C., Nenadoc I., et al. Deformation-based morphometry and its relation to conventional volumetry of brain lateral ventricles in MRI. *NeuroImage* 13, 1140-1145, 2001.
7. Inder et al. Abnormal cerebral structure is present at term in premature infants. *Pediatrics*, January, 2005.
8. T. L. Jernigan, A. C. Gamst, C. Fennema-Notestine, and A. L. Ostergaard. More Mapping in Brain Mapping: statistical comparison of effects. *Hum Brain Mapp* 19, 90-95, 2003.
9. Maalouf,E.F. *et al.* Magnetic resonance imaging of the brain in a cohort of extremely preterm infants. *J. Pediatr.* 135, 351-357, 1999.
10. Marlow,N., Roberts,L. & Cooke,R. Outcome at 8 years for children with birth weights of 1250 g or less. *Arch. Dis. Child* 68, 286-290, 1993.
11. McCormick,M.C., Workman-Daniels,K. & Brooks-Gunn,J. The behavioral and emotional well-being of school-age children with different birth weights. *Pediatrics* 97, 18-25, 1996.
12. D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, "Non-rigid registration using free-form deformations: Application to breast MR images", *IEEE Transactions on Medical Imaging*, vol. 18, no. 8, pp. 712-721, 1999.
13. C. Studholme, V. Cardenas, R. Blumenfeld, N. Schuff, H. J. Rosen, B. Miller, and M. Weiner, "Deformation tensor morphometry of semantic dementia with quantitative validation", *NeuroImage* 21, 1387-1398, 2004.
14. C. E. Thomaz et al. "Using a Maximum Uncertainty LDA-based Approach to Classify and Analyse MR Brain Images". In Proc. MICCAI, LNCS 3216, pp. 291-300, 2004.
15. C. E. Thomaz and D. F. Gillies, "A Maximum Uncertainty LDA-based approach for Limited Sample Size problems - with application to Face Recognition", *Technical Report TR-2004-01*, Department of Computing, Imperial College, London, UK, January 2004.
16. Wood, N.S., Marlow,N., Costeloe,K., Gibson,A.T. & Wilkinson,A.R. Neurologic and developmental disability after extremely preterm birth. EPICure Study Group. *N. Engl. J. Med.* 343, 378-384, 2000.