Available online at www.sciencedirect.com
SCIENCE DIRECT
Magnetic Resonance Imaging 24 (2006) 793–800
MAGNETIC RESONANCE IMAGING

Principal component and linear discriminant analysis of T_1 histograms of white and grey matter in multiple sclerosis

Daniel J. Tozer[a,*], Gerard R. Davies[a], Daniel R. Altmann[a,b], David H. Miller[a], Paul S. Tofts[a]

[a]NMR Research Unit, Department of Neuroinflammation, Institute of Neurology, University College London, WC1N 3BG London, UK
[b]Medical Statistics Unit, London School of Hygiene and Tropical Medicine, Keppel Street, WC1E 7HT London, UK

Received 17 May 2005; accepted 1 August 2005

Abstract

Twenty-three relapsing remitting multiple sclerosis (RRMS) patients and 14 controls were imaged to produce normal-appearing white and grey matter T_1 histograms. These were used to assess whether histogram measures from principal component analysis (PCA) and linear discriminant analysis (LDA) out-perform traditional histogram metrics in classification of T_1 histograms into control and RRMS subject groups and in correlation with the expanded disability status score (EDSS). The histograms were classified into one of two groups using a leave-one-out analysis. In addition, the patients were scanned serially, and the calculated parameters correlated with the EDSS.

The classification results showed that the more complex techniques were at least as good at classifying the subjects as histogram mean, peak height and peak location, with PCA/LDA having success rates of 76% for white matter and 68%/65% for grey matter. No significant correlations were found with EDSS for any histogram parameter. These results indicate that there is much information contained within the grey matter as well as the white matter histograms. Although in these histograms PCA and LDA did not add greatly to the discriminatory power of traditional histogram parameters, they provide marginally better performance, while relying only on data-driven feature selection.

Keywords: Magnetic resonance imaging; Principal component analysis; Linear discriminant analysis; Multiple sclerosis

1. Introduction

Histogram analysis of magnetic resonance parameters is now an established tool, and many parameters such as the magnetisation transfer ratio (MTR) [1] and fractional anisotropy from diffusion-weighted imaging [2] have been investigated in a variety of conditions. Histograms are used in multiple sclerosis (MS) as (in common with many conditions) pathological changes are not limited to the focal lesions seen on conventional images; abnormalities are seen in the surrounding brain tissue with many parameters such as MTR [3], apparent diffusion coefficient and fractional anisotropy [4,5] and T_1 [6]. In studying the normal-appearing tissues, histograms have many advantages over region of interest-based approaches, the biggest being that bias in the positioning of regions is avoided and that information from the whole of the dataset is used.

In most cases, traditional histogram metrics such as the peak height (PH) and peak location (PL) are used to attempt to either separate patient and control groups or to correlate with clinical status indicators such as the expanded disability status scale (EDSS). However, recently, more complex tools such as principal component analysis (PCA) and linear discriminant analysis (LDA) have been used to provide the same information [7–10]. These techniques extract information from the whole histogram, as opposed to more localised measures such as PL. It has been argued [7] that the choice of histogram metrics such as PH is arbitrary and based on ease of measurement and visibility, so it is unlikely to be optimal for either of the tasks mentioned above. To classify subjects, the discriminant needs to maximise the variation between groups, and for correlation with EDSS, it must maximise the variation within a group [7]. These may be maximised by using information from throughout the histograms, so the analysis technique must take account of this. Techniques such as PCA and LDA do this and should prove superior to the standard measures. In MS, MTR histograms have been widely studied and many

* Corresponding author. Tel.: +44 (0)20 7837 3611x4260; fax: +44 (0)20 7278 5616.
E-mail address: d.tozer@ion.ucl.ac.uk (D.J. Tozer).

0730-725X/$ – see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.mri.2005.08.002

interesting studies published [4,11–13]. Some have had success in the classification of subjects and in finding a relationship with disease progression [7,11]. However, they do not totally explain the clinical status of a patient.

Similarly, T_1 values in MS have been shown to be increased not only in MS lesions but also in normal appearing white matter (NAWM) compared to control white matter [6]. The recent papers by Griffin et al. [6] and Vaithianathar et al. [14] have shown that the traditional histogram metrics can be used to distinguish relapsing remitting (RR) MS patients and healthy controls. However, once again, there was only limited correlation between the histogram parameters and disability. As there is only a modest correlation between disease status and lesion load (one long-term study suggested that lesion load, as seen on T_2-weighted images, accounts for only about a third of the variability in clinical status [15]), processes other than macroscopic lesion formation play an important role in disability. An increase in T_1 relaxation time, as seen in NAWM and histogram parameters [6], is likely to be caused by an increase in the mobile pool of protons caused by any number of disease processes, such as myelin loss, gliosis or inflammation. If these processes are occurring subtly in regions of normal-appearing brain tissue or at future lesion sites, then histogram analysis may pick up changes which could be missed by regional analysis.

Dehmeshki et al. [9,16] have shown that grey matter MTR histograms of primary progressive MS patients were correlated with EDSS using the histogram mean and PCA and that a correlation was seen using PCA in whole-brain histograms as well, meaning there was no need for segmentation. Histograms of the apparent diffusion coefficient or mean diffusivity (MD) have also been shown to differ in MS patients. Nusbaum et al. [17] and Cercignani et al. [2] showed an increased MD in MS patients, although this depended somewhat on clinical subtype. Wilson et al. [18] showed that the PH of the ADC histogram correlated with disease duration, disability and cerebral atrophy, although this finding has not been widely repeated [19]. These mixed results may be due to the arbitrary nature of the choice of histogram metric, and the use of global measures such as PCA and LDA may allow a more complete investigation of the histogram. This work looks at applying these techniques to T_1 histograms to obtain measures that discriminate between patient and control groups and correlate with measures of disease progression such as EDSS, with the hope of improving on results from the traditional histogram metrics.

2. Methods

Twenty-three MS patients diagnosed with the early RR form of the disease and 14 controls were imaged using the method of Parker et al. [20] to produce T_1 maps of the whole brain. The baseline EDSS and disease duration were (mean±S.D.) 1.11±0.72 and 724±239 days, respectively. In addition, the subjects were imaged longitudinally at gaps of 6 months for up to 3 years having between two and seven studies (mean±S.D.=4.9±1.6) to assess temporal changes in their T_1 and EDSS. The imaging methods can be seen in more detail elsewhere [20], however briefly PD-and T_1-weighted gradient echo images were acquired on 28 slices with the following parameters: slice thickness=5 mm, field of view=24 cm×24 cm, matrix=256^2, flip angle=45°, TE=11 ms, TR(PD)/TR(T_1)=1500/50 ms and NEX(PD)/NEX(T_1)=1.5/3. These images were corrected for B_1 and flip angle inhomogeniety and their ratio taken to produce the T_1 maps. The PD-weighted images of the control subjects were then segmented into white matter (WM), grey matter (GM) and cerebrospinal fluid (CSF) segments using SPM99 [21]; while the patients had any lesions removed and were segmented into NAWM, normal-appearing GM and CSF segments, these segmentations were then applied to the T_1 maps. To reduce the effect of spurious pixel classifications, the WM segments were thresholded to exclude values above 1000 ms and the GM to 1700 ms (although in fact the histograms have come down to zero well below 1000 ms for WM and are close to zero at 1700 ms for GM (Figs. 1 and 2), so no bias was introduced by this operation). Both were then eroded (WM twice and GM once) to reduce partial volume effects. From these resultant maps, histograms were produced of WM and GM for each subject. These were normalised to the eroded brain volume to account for differences in brain volume and were smoothened with a moving average window width of 5 ms. The bin width was 1 ms. The resulting histograms had y values expressed in units of % volume/ms and areas of 100% volume (Figs. 1–4). There were approximately 20,000 voxels in the white matter histograms and 60,000 in the GM histograms.

PCA and LDA were both performed on the histograms to answer two questions: firstly, how good was each method at classifying a histogram correctly into the control or patient groups, using only the first time point for each subject, and secondly, how well did the measures obtained from each analysis correlate with clinical measures for the patients.

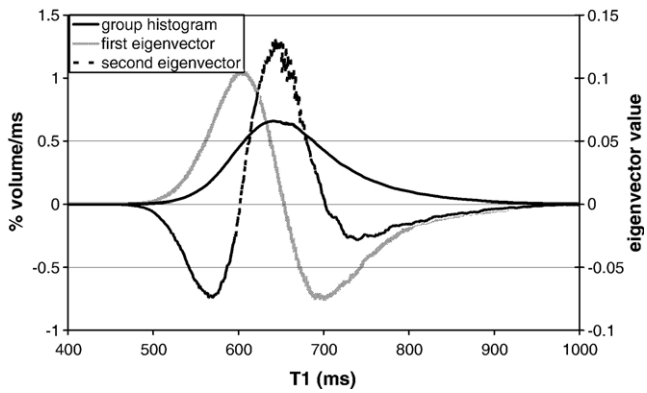

Fig. 1. The group histogram and first two eigenvectors for the patient group in WM, which were obtained using PCA. The primary y-axis related to the histogram and the secondary, to the eigenvectors.

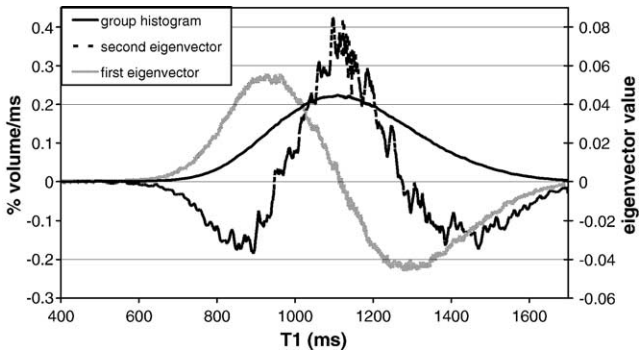

Fig. 2. GM patient group histogram and associated eigenvectors from PCA. Again, the primary y-axis related to the histogram and the secondary, to the eigenvectors.

2.1. Principal component analysis

PCA is a technique for reducing the dimensionality of an object while concentrating on the aspects that provide the most variation within a group. Scores, called principal components (PCs), are extracted, which explain decreasing amounts of variation within a group, with the proviso that it is orthogonal to the first. In general, the vast majority of the variation is contained in a few PCs so, e.g., a 1000-bin histogram can be reduced to a few parameters, without significantly reducing the value of the information stored within it. This is, in principle, the same as extracting a few parameters such as the PH and PL. However, the choice of PH or PL is often arbitrary and made for convenience of analysis. The features extracted using PCA are free from selection bias.

A full mathematical description of PCA can be found elsewhere [7,9,22,23]; simply, it involves performing an eigen analysis on a covariance matrix constructed from all the histograms within a sample. Each PC can be defined as shown in Eq. 1

$$\mathrm{PC}_j = \underline{v}_j \cdot \underline{h} = v_1 h_1 + v_2 h_2 + \cdots + v_N h_N \tag{1}$$

where PC_j is the jth PC; h_N, the Nth bin of histogram h; N, the number of bins in the histogram and v_j, the jth eigenvector containing elements $v_1, v_2, \ldots, v_N$.

Once eigenvectors and eigenvalues have been obtained for each group, a new subject can be classified using a Bayesian classification technique. The variation of the histogram data can be modelled for each group as a Gaussian distribution, as shown in Eq. 2:

$$p\left(\underline{h}^p | C^i\right) = \frac{1}{\prod_{j=1}^{m} \sqrt{2\pi\lambda_j^i}} \exp\left(- \sum_{j=1}^{m} \frac{\left(\left(\underline{h}^p - \underline{\overline{h}}^i \right) \cdot \underline{v}_j^i \right)^2}{2\lambda_j^i} \right) \tag{2}$$

where i and C^i represent the group, $\underline{h}^p$, the histogram of the new subject; $\underline{\overline{h}}^i$, the group histogram of i; m, the number of PCs included in the analysis and $\underline{v}^i$ and λ^i, the eigenvectors and eigenvalues of i. The Bayesian post probability of a histogram (p) belonging to each group is then given by Eq. 3:

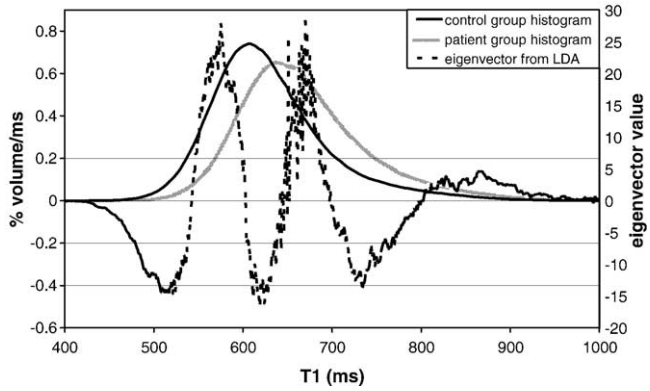

Fig. 3. The WM group histograms and the eigenvector obtained from the LDA. As for the PCA figures, the primary y-axis relates to the histograms and the secondary, to the eigenvector.

$$P(C^i | \underline{h}^p) = \frac{p(\underline{h}^p | C^i)\, \overset{i}{\Pi}}{\sum_g p\left(\underline{h}^p | C^g\right) \overset{g}{\Pi}} \tag{3}$$

where g sums over all groups and Π^i is the prior probability of the new subject belonging to group i. The histogram is then classified to the group to which it most likely belongs. It was assumed that the prior probability of a histogram belonging to a specific group was equal for all groups (1/number of groups). Probability thresholds and levels of confidence in the classifications can be assigned if required.

2.2. Linear discriminant analysis

LDA tries to maximise the ratio of between-group variance to within-group variance. This means that in theory, it should be superior to PCA in classifying subjects into groups. The technique was first devised by Fisher [24] and is sometimes known as Fisher's linear discriminant. Again, the method reduces the dimensionality of an object by producing a series of linear discriminant functions, analogous to the PCs. In contrast to PCA, the number of

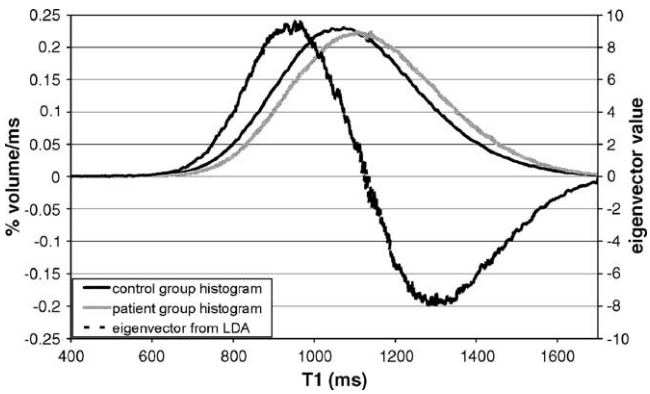

Fig. 4. The GM group histograms and the eigenvector obtained from the LDA. The two y-axes are as in the other figures.

Table 1
Percentage variation of PCs

	WM		GM	
No. of PC	1	2	1	2
Controls (%)	80	97	79	91
Patients (%)	80	97	93	96

The percentage of variation were given by the first PC and first 2 PCs in the patients and control groups for WM and GM.

these functions is limited to at most $g-1$ where g is the number of groups. As the method also involves eigen analysis of a matrix, this means that there are at most $g-1$ nonzero eigenvalues. Again, a detailed analysis of the technique can be found elsewhere [25,26].

The linear discriminant scores are then calculated, as with the PC, as the dot product of the relevant eigenvector, or linear discriminant function, with the histogram. This then allows the calculation of a nearest-mean group classifier, where the subject is classified by the minimum distance between its transform scores and a group centre point.

The individual PC and LDA transform scores for each histogram can be used to correlate with measures such as the EDSS; however, to take account of more than one PC, a multiple regression model must be constructed with EDSS as the dependent variable and the PCs as the independent variables.

For the histogram classification, the leave-one-out [27,28] analysis technique was used. Training of the classifier is performed using group statistics calculated leaving 1 histogram out in turn. This histogram is thus classified using the measures from the other subjects. This removes the bias inherent in trying to classify a histogram using training groups that include that histogram but does not significantly reduce the sample size used for training. Classification was performed using the linear discriminant function from LDA (only 1 function was obtained as there are only 2 groups), combinations of the first n PCs and, for comparison, the three conventional histogram parameters PH, PL and mean.

Associations between the EDSS scores at the various time points and the PCs and linear discriminant scores obtained at these times were examined by multilevel regression of EDSS score, with the PCs as predictors and duration (scan time from onset) as covariate; the models employed random intercepts and random slopes on duration. Although there are problems interpreting numerical estimation of EDSS results when EDSS score is treated as a continuous variable, in this context, the negative results can be interpreted with some confidence as other methods for analysing EDSS as an ordinal variable are likely to be much less sensitive to association.

The study was given ethical approval by the joint ethics committee of the Institute of Neurology and National Hospital for Neurology and Neurosurgery, and all patients gave informed written consent prior to imaging.

3. Results

Cross-sectional patient group histograms for WM and GM can be seen in Figs. 1 and 2; also shown are the first 2 eigenvectors for each group obtained from the PCA. The control group histograms and eigenvectors were very similar and, so, are not shown. The areas of the eigenvectors with the greatest magnitude relate to areas in the histogram with the greatest within-group variation. The equivalent data for the LDA are shown in Figs. 3 and 4. Here, the same group histograms are shown with the eigenvector calculated from their combination, as described above. Here, the areas of greatest magnitude in the eigenvector show the areas of the histograms with the largest ratio of between-group variance to within-group variance.

3.1. Subject classification

Table 1 shows the percentage of variation in the histogram groups accounted for by the first two PCs for both WM and GM in the patients and controls. As can be seen, the first two PCs contain over 90% of the within-group variation for both groups and tissues, so it was decided to use only these for the subject classification. Subsequent use of more PCs did not improve on the results shown below.

Using the Student's t test, there were significant differences between the control and patient groups for many of the parameters investigated. The P values are shown in Table 2. The PCs were calculated from the eigenvectors obtained from all the subjects to ensure that any variations were not due to differences in the calculation of the eigenvectors, whereas (as was discussed above) for the classification, the eigenvectors were calculated for each group leaving each histogram out in turn, resulting in a different eigenvector for each classification. As can be seen, there were highly significant differences for the first PC between the controls and patients for both tissue types. The

Table 2
Histogram parameters and P values

Parameter	WM Co	WM Pa	GM Co	GM Pa	WM P value	GM P value
PH (% volume/ms)	0.88±0.08	0.82±0.09	0.25±0.01	0.24±0.01	**.037**	.415
PL (ms)	609±27	644±28	1085±46	1116±39	**.01**	**.038**
Mean (ms)	630±23	670±29	1099±24	1138±40	**.0001**	**.001**
LDA score	0.8±1.4	−0.6±1.1	0.9±0.7	−0.4±1.3	**.002**	**.002**
PC1	0.026±0.022	−0.006±0.023	0.007±0.004	0.0002±0.007	**.0002**	**.001**
PC2	0.031±0.014	0.034±0.013	0.020±0.002	0.020±0.001	.491	.954

The mean parameters values (±S.D.) from the histograms for the control and patient groups and the P values obtained from the Student's t test for their comparison. Significant results at the $P=.05$ level are shown in bold. Pa, patient groups; Co, control groups.

Table 3
Histogram parameter classification rates

	PH		PL		Mean		PCA		LDA	
Tissue	WM	GM	WM	GM	WM	GM	WM	GM	WM	GM
Controls (%)	64	50	**79**	57	**79**	71	**79**	**86**	71	71
Patients (%)	57	65	70	**74**	65	61	74	56	**78**	61
Total (%)	59	59	73	**68**	70	65	**76**	**68**	**76**	65

The success rates for the various histogram parameters at classifying the histograms into the correct group using a leave-one-out analysis (100%=14 in the control group, 23 in the patient group and 37 in total). The figures in bold are the best results for each tissue and group.

second PC was not significantly different for either tissue type; this is expected, as it accounts for a lot less within-group variation than the first PC.

The subject classification success rates for the PCA, LDA and conventional histogram parameters are shown in Table 3. For the WM histograms, both PCA and LDA produced the best success rate, while for the GM histograms, there was equal success using the PCA and PL. While this is encouraging in that it suggests that the more complex techniques may add to the ability of the histograms to separate the groups, it is difficult to conclude that they are significantly better, as the results for the PL and histogram means are very similar to those of both PCA and LDA. This is especially so when it is remembered that there is a limited sample size, so an improvement of 3% only equates to one more subject being correctly classified.

The parameters from the various analyses were then correlated with each other to look for relationships between them. Table 4 shows that the parameters that perform the best, PL, mean, the LDA score and the first PC, are all highly correlated, indicating that the features extracted by the PCA and LDA, being the most important, are linked to the conventional histogram parameters. In contrast, the PH, which is much less successful at classifying the histograms, only correlates with the second PC. There are also only slight differences between the results for GM and WM, indicating that the features extracted by PCA and LDA are similar for both tissues.

It is interesting to look at the classification rates for PCA in terms of the sums of their eigenvalues. The greater the sum of the eigenvaglues for a group, the greater the within-group variation [7] (it is the relative size that is important, as the eigenvalues will differ in size between data sets, such as WM and GM histograms, but this does not mean one has more variation than the other). The control and patient group eigenvalues summed to 7.02×10^{-4} and 7.17×10^{-4}, respectively, for the WM histograms. This shows that there is only approximately 2% more variation in the patient group than the control group, whereas for the GM histograms, the sums were 2.0×10^{-5} and 4.8×10^{-5}, respectively, an increase of 140% in the patient group. These differences are borne out by the classification rates for patients and controls in WM, which are similar to the sum of eigenvalues, but for the GM, the classification rate for controls is much higher than for patients, indicating that the reduced variation in the control group makes them easier to classify correctly.

3.2. Correlation with clinical parameters

The correlations between EDSS and PCs and the linear discriminant scores showed no significant correlations ($P \geq .52$). The high P values obtained suggest that there is no hint of a relationship; however, these may also be caused by the small group sizes or the small range in disability scores.

4. Discussion

Although LDA and PCA classified the histograms as well as, if not better than, all other histogram parameters, the improvement was not such that it should be considered proven. It is not known whether this result is due to the small group sizes or whether some of the conventional histogram parameters are indeed close to optimal for separating the two groups. Certainly, the fact that there are significant differences between the groups for all the

Table 4
Correlation of histogram parameters and PCA/LDA

	Group	PC1	PC2	LDA score
WM				
PH	All	0.179 (0.288)	0.431 (**0.008**)	0.232 (0.167)
	Controls	−0.238 (0.413)	0.342 (0.231)	−0.013 (0.664)
	Patients	0.367 (0.085)	0.619 (**0.002**)	0.21 (0.338)
PL	All	−0.821 (**4e−10**)	0.102 (0.548)	−0.927 (**2e−16**)
	Controls	−0.992 (**3e−12**)	0.007 (0.982)	−0.916 (**4e−6**)
	Patients	−0.871 (**6e−8**)	0.035 (0.872)	−0.907 (**2e−9**)
Mean	All	−0.803 (**2e−9**)	0.074 (0.662)	−0.923 (**5e−16**)
	Controls	−0.984 (**2e−10**)	0.095 (0.748)	−0.947 (**3e−7**)
	Patients	−0.913 (**1e−9**)	−0.075 (0.734)	−0.904 (**3e−9**)
GM				
PH	All	−0.091 (0.593)	0.622 (**4e−5**)	−0.012 (0.945)
	Controls	−0.044 (0.881)	0.637 (**0.014**)	−0.024 (0.935)
	Patients	−0.142 (0.518)	0.672 (**4e−4**)	−0.12 (0.585)
PL	All	−0.685 (**3e−6**)	−0.19 (0.26)	−0.738 (**2e−7**)
	Controls	−0.687 (**0.007**)	−0.141 (0.631)	−0.65 (**0.012**)
	Patients	−0.764 (**2e−5**)	0.092 (0.677)	−0.767 (**2e−5**)
Mean	All	−0.894 (**9e−14**)	−0.293 (0.079)	−0.993 (**3e−34**)
	Controls	−0.985 (**1e−10**)	−0.102 (0.729)	−0.99 (**1e−11**)
	Patients	−0.996 (**1e−23**)	−0.028 (0.9)	0.994 (**5e−22**)

Correlation of the conventional histogram parameters with those calculated from PCA and LDA for all subjects, the patients and the controls. Shown are the Pearson correlation coefficients, R, and the (P values); correlations significant at the $P=.05$ level are given in bold. Note how parameters with similar success rates in the classification of the histograms are highly correlated (4e−10 ≡ 4×10^{-10}).

parameters except PH and the high level of correlation of PL and histogram mean with the LDA scores and the first PC indicate that all these measures should be similarly successful at separating the groups. Work looking at applying PCA and LDA to MTR histograms found that LDA was the better discriminator, [7,9] as it should be from the definition of the quantity it maximises (to classify subjects into groups requires a maximised between-group variance as opposed to the within-group variance maximised by PCA). Given this, it is disappointing that LDA does not outperform PCA or, indeed, the other histogram parameters. However, it may be that on this limited selection of histograms, the areas that optimally separate the groups are related to those which maximise the within-group variation, as suggested by the heavy correlation between the LDA score and the first PC ($r=.765$, $P=3\times10^{-8}$ and $r=.921$, $P=6\times10^{-16}$ for WM and GM, respectively).

The accuracy of the techniques may also be limited by the changes which occur in the normal-appearing brain matter. It is not surprising that the techniques were all more successful in classifying the WM histograms than the GM. Although there is evidence that GM is affected in MS [29], T_1 abnormality may be more pronounced in WM [30]. This means that the histogram analysis techniques should find it easier to separate the WM histograms, something which is borne out by these results. However, it is clear that there is a significant amount of information in the GM histograms as their discriminatory power is almost as good as for the WM. Although it is not possible to detect focal cortical lesions using these techniques, it may be that such lesions, which are known to occur in MS [31,32], are still having an influence on the histograms.

Several patients were also at the early stages of the disease, and any changes which may have taken place could well be too small for the histogram to pick up. Small regional changes such as microscopic lesions [33] may occur early in the disease course but be masked when the whole tissue is considered. This is supported by the fact that the GM histograms were less successfully classified, as they are less visibly affected by disease. It is quite possible that a cohort with more developed disease could show an improvement in classification rate, and it would be of interest to determine the relative discriminatory value of the analysis methods in both WM and GM when comparing distinct clinical subgroups, e.g., nondisabled RR versus disabled secondary progressive MS.

Given all these factors, it is reasonable to state that the hypothesis that LDA and PCA should perform at least as well as standard histogram measures is upheld by these results, although it is also true that they do not perform significantly better.

It is interesting, however, that Table 3 seems to indicate that some parameters achieved better results in the control group, while others were better in the patient group, and that this varied between the two tissues. This variation was examined above for the PCA, where it can be seen that the greater sum of eigenvalues in the patient group for GM leads to a reduced successful classification rate, whereas the similar sums in WM are reflected in similar success rates. For all the other parameters, the variation in success rate between the groups can be explained by looking at the standard deviation of the parameters in each group. For the WM histograms, the standard deviation of the PH, PL and histogram mean is smaller for the controls than the patients, while for the LDA scores, this is reversed. In the GM, controls have a smaller standard deviation for the histogram mean and LDA score, while both PL and PH show a smaller standard deviation in the patient group. This fits in with the percentages shown in Table 3, where in each case, the group with the smaller standard deviation has more of its members correctly classified, which fits in with the intuitive reasoning that groups showing less within-group variation should be easier to classify correctly.

One further interesting point from Table 3 is that the parameters with the best classification success rates are not the same as those with the smallest P values. In particular, the PL classifies more cases correctly than the histogram mean, but the P value of the former is much larger. This apparent discrepancy is due to differences in the formulation of the tests. The function of the t test is to assess whether the difference in means observed in the sample could have arisen by chance when there is no true difference. Classification techniques assess the success of criteria at correctly classifying the data points within the sample, based on sample values; therefore, they do not directly comment on any underlying generality likely to be repeated in similar studies sampling from the same population.

The pictorial representations of the first eigenvector in Figs. 1 and 2 seem to imply that the within-group variation is maximum in the areas at either side of the peak, whereas the variation accounted for by the second eigenvector is at a maximum at the histogram peak itself. These histograms also show that, although there is some variation (as seen by the eigenvectors) at the tails of the group histograms, the vast majority occurs at values where the histogram is relatively large, indicating that any partial volume or segmentation errors are unlikely to have affected the results to a great extent. This fits with the classification results presented above. The fact that the first eigenvector is maximal at either side of the peak indicates further that it is related to the PL, whereas the second eigenvector appears to be related to the PH.

The structure of the eigenvectors is also similar to that seen in previous work looking at functional magnetic resonance imaging [34] and first-pass perfusion [35] data. The first eigenvector seems to represent the derivative of the group histogram and the second eigenvector, the second derivative. In electroencephalography data, this is thought to relate to a jitter in the time delay for individual data [36]. If the same principle is applied to this work, then it suggests that there are shifts in the histogram location between subjects, which mimic the effect of a variable time delay. This is supported by the high correlations seen between the

first PC and the histogram location statistics such as PL and histogram mean. This natural variation in histogram location would partially explain why the classification results are slightly disappointing. If the PCA picks out this feature of the histograms rather than any shape change, it is not surprising that (a) its performance is no better than the PL and (b) that a significant number of histograms are misclassified using these metrics due to the overlap in values between the two groups. Analysis of the LDA eigenvector is more complex due to its composite nature, but the similar correlations seen in Table 4 and the fact that the eigenvector is zero at both group histogram peaks suggests that LDA follows a similar pattern to PCA.

As far as the correlation with EDSS is concerned, again, the results for PCA and LDA are disappointing. There was no evidence of a relationship between the parameters, and although other studies [6,7,16] looking at the association of T_1 with clinical parameters have shown only limited correlations, that is still superior to the results seen here. The lack of a positive result may be due to the fact that the patient group analysed is in the early stages of MS and are not particularly disabled, meaning that there is not significant enough range in the EDSS for an accurate assessment of any correlation. Indeed, the mean EDSS for all patients at all time points was 1.6 ± 1. This is coupled with the discrete nature of the measure to mean that it is difficult to obtain an accurate assessment of any correlation that may exist. These results should not be seen as conclusive due to the limited range of clinical dysfunction and the small group size.

5. Conclusions

Although the PCA and LDA techniques seem to offer a slight improvement on the traditional measures, it was not big enough to conclude that they are significantly better in the study of minimally disabled early MS patients with T_1 histograms; larger group sizes are needed to confirm this result. The use of PCA and LDA does, however, remove the need for feature selection and gives confidence that optimal analysis is being used. While the WM histograms were more successfully classified, the GM histograms also produced a success rate indicating that they contain useful information. No correlations were seen between the EDSS and the PCA and LDA parameters.

Acknowledgments

The authors would like to thank the MS society of Great Britain and Northern Ireland for support of this work and the NMR research unit.

References

[1] van Buchem MA, McGowan JC, Kolson DL, Polansky M, Grossman RI. Quantitative volumetric magnetization transfer analysis in multiple sclerosis: estimation of macroscopic and microscopic disease burden. Magn Reson Med 1996;36:632–6.

[2] Cercignani M, Inglese M, Pagani E, Comi G, Filippi M. Mean diffusivity and fractional anisotropy histograms of patients with multiple sclerosis. Am J Neuroradiol 2001;22:952–8.

[3] Ge YL, Grossman RI, Udupa JK, Babb JS, Mannon LJ, McGowan JC. Magnetization transfer ratio histogram analysis of normal-appearing gray matter and normal-appearing white matter in multiple sclerosis. J Comput Assist Tomogr 2002;26:62–8.

[4] Cercignani M, Bozzali M, Iannucci G, Comi G, Filippi M. Magnetisation transfer ratio and mean diffusivity of normal appearing white and grey matter from patients with multiple sclerosis. J Neurol Neurosurg Psychiatry 2001;70:311–7.

[5] Rashid W, Hadjiprocopis A, Griffin CM, Chard DT, Davies GR, Barker GJ, et al. Diffusion tensor imaging of early relapsing-remitting multiple sclerosis with histogram analysis using automated segmentation and brain volume correction. Mult Scler 2004;10:9–15.

[6] Griffin CM, Dehmeshki J, Chard DT, Parker GJ, Barker GJ, Thompson AJ, et al. T1 histograms of normal-appearing brain tissue are abnormal in early relapsing-remitting multiple sclerosis. Mult Scler 2002;8:211–6.

[7] Dehmeshki J, Barker GJ, Tofts PS. Classification of disease subgroup and correlation with disease severity using magnetic resonance imaging whole-brain histograms: application to magnetization transfer ratios and multiple sclerosis. IEEE Trans Med Imag 2002;21:320–31.

[8] Dehmeshki J, Silver NC, Leary SM, Tofts PS, Thompson AJ, Miller DH. Magnetisation transfer ratio histogram analysis of primary progressive and other multiple sclerosis subgroups. J Neurosci 2001; 185:11–7.

[9] Dehmeshki J, Ruto AC, Arridge S, Silver NC, Miller DH, Tofts PS. Analysis of MTR histograms in multiple sclerosis using principal components and multiple discriminant analysis. Magn Reson Med 2001;46:600–9.

[10] Dehmeshki J, Van Buchem M, Bosma GPT, Huizinga TWJ, Tofts PS. Systemic lupus erythematosus: diagnostic application of magnetization transfer ratio histograms in patients with neuropsychiatric symptoms — initial results. Radiology 2002;222:722–8.

[11] Rovaris M, Bozzali M, Santuccio G, Ghezzi A, Caputo D, Montanari E, et al. In vivo assessment of the brain and cervical cord pathology of patients with primary progressive multiple sclerosis. Brain 2001;124: 2540–9.

[12] Traboulsee A, Dehmeshki J, Peters KR, Griffin CM, Brex PA, Silver N, et al. Disability in multiple sclerosis is related to normal appearing brain tissue MTR histogram abnormalities. Mult Scler 2003;9:566–73.

[13] Rocca MA, Mastronardo G, Rodegher M, Comi G, Filippi M. Long-term changes of magnetization transfer-derived measures from patients with relapsing-remitting and secondary progressive multiple sclerosis. AJNR Am J Neuroradiol 1999;20:821–7.

[14] Vaithianathar V, Tench CR, Morgan PS, Lin X, Blumhardt LD. White matter T1 time histograms and cerebral atrophy in multiple sclerosis. J Neurosci 2002;197:45–50.

[15] Molyneux PD, Barker GJ, Barkhof F, Beckmann K, Dahlke F, Filippi M, et al. Clinical-MRI correlations in a European trial of interferon beta-1b in secondary progressive MS. Neurology 2001;57: 2191–7.

[16] Dehmeshki J, Chard DT, Leary S, Thompson AJ, Miller DH, Tofts PS. Magnetisation transfer histograms in primary progressive multiple sclerosis: grey matter changes relate to disability and principal component analysis shows this most sensitively. Book of abstracts: tenth annual meeting International Society Magnetic Resonance in Medicine. vol. 1. Hawaii: ISMRM; 2002. p. 179.

[17] Nusbaum AO, Tany CY, Wei T, Buchsbaum MS, Atlas SW. Whole-brain diffusion MR histograms differ between MS subtypes. Neurology 2000;54:1421–7.

[18] Wilson M, Morgan PS, Lin X, Turner BP, Blumhardt LD. Quantitative diffusion weighted magnetic resonance imaging, cerebral atrophy and

disability in multiple sclerosis. J Neurol Neurosurg Psychiat 2001;70: 318–22.
[19] Bammer R, Fazekas F. Diffusion imaging in multiple sclerosis. Neuroimaging Clin N Am 2002;12:71–106.
[20] Parker GJM, Barker GJ, Tofts PS. Accurate multislice gradient echo T1 measurement in the presence of non-ideal RF pulse shape and RF field nonuniformity. Magn Reson Med 2001;45:838–45.
[21] Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RSJ. Statistical parametric maps in functional imaging: a general linear approach. Hum Brain Mapp 1995;2:189–210.
[22] Jackson JE. A user's guide to principal components. New York: John Wiley & Sons; 1991.
[23] Flury B. Common principal components and related multivariate models. New York: John Wiley & Sons; 1988.
[24] Fisher RA. The use of multiple measurements in taxonomic problems. Annals Eugenics 1936;7:179–88.
[25] Webb A. Statistical pattern recognition. London: Arnold; 1999.
[26] Hand DJ. Construction and assessment of classification rules. Chichester: John Wiley & Sons; 1997.
[27] Hills M. Allocation rules and their error rates (with Discussion). J R Stat Soc 1966;28:1–31.
[28] Lunts Al, Braillovsky VL. Evaluation of attributes obtained in statistical decision rules. Eng Cybern 1967;3:98–109.
[29] De Stefano N, Matthews PM, Filippi M, Agosta F, De Luca M, Bartolozzi ML, et al. Evidence of early cortical atrophy in MS relevance to white matter changes and disability. Neurology 2003;60:1157–62.
[30] Griffin CM, Chard DT, Parker GJM, Barker GJ, Thompson AJ, Miller DH. The relationship between lesion and normal appearing brain tissue abnormalities in early relapsing remitting multiple sclerosis. J Neurol 2002;249:193–9.
[31] Kidd D, Barkhof F, McConnell R, Algra PR, Allen IV, Revesz T. Cortical lesions in multiple sclerosis. Brain 1999;122:17–26.
[32] Peterson JW, Bo L, Mork S, Chang A, Trapp BD. Transected neurites, apoptotic neurons, and reduced inflammation in cortical multiple sclerosis lesions. Ann Neurol 2001;50:389–400.
[33] Barbosa S, Blumhardt LD, Roberts N, Lock T, Edwards RH. Magnetic resonance relaxation time mapping in multiple sclerosis: normal appearing white matter and the "invisible" lesion load. Magn Reson Imaging 1994;12:33–42.
[34] Andersen AH, Rayens WS. Structure-seeking multilinear methods for the analysis of fMRI data. Neuroimage 2004;22:728–39.
[35] Andersen AH, Rayens WS, Greene-Avison R, Pettigrew C, Berger J, Avison MJ. PCA for mapping time delays in the transient response from functional MRI studies. Book of abstracts: ninth annual meeting International Society Magnetic Resonance in Medicine, vol. 3. Glasgow (UK): ISMRM; 2001. p. 1720.
[36] Möcks J. The influence of latency jitter in principal component analysis of event-related potentials. Psychophysiology 1986;23: 480–4.