Classification of Cartilage Degradation and Quantification of Matrix Composition through Multiparametric Support Vector Machine Analysis

P-C. Lin[1], O. Irrechukwu[1], R. Roque[1], and R. G. Spencer[1]
[1]National Institute on Aging, National Institutes of Health, Baltimore, MD, United States

Introduction. The development of MRI techniques to detect early osteoarthritis and monitor therapeutic response to potential interventions has been the subject of intense activity over recent years. However, individual MRI parameters exhibit limited sensitivity and specificity to cartilage pathology. Even when a statistically significant difference is observed in the mean values of a given parameter between e.g. control and degraded cartilage, there remains a substantial degree of overlap in the parameter values obtained for samples of the two groups [1, 2]. In view of these limitations of univariate classification, we undertook a multivariate support vector machine (SVM) analysis of bovine nasal cartilage (BNC) samples with pathomimetic degradation using trypsin and collagenase. The SVM is an extension of linear discriminant analysis that distinguishes between classes according to a separating hyperplane in a transformed parameter space [3]. In our case, the classes were defined by exposure to enzymatic degradation, and the parameters investigated were T_1, T_2, magnetization transfer rate (k_m) and apparent diffusion coefficient (ADC). SVM classification results were compared with results obtained by conventional univariate analysis. Further analysis explored the correlation of cartilage biochemical content with group assignment probabilities obtained from the SVM procedure in order to account for the graded nature of cartilage degradation.

Materials and Methods *Sample Preparation:* BNC discs harvested from the nasal septa of 6 month-old calves were imaged before and after 24 hr digestion in trypsin (1mg/ml; n=36) or collagenase II (30 units/ml; n=66). *MRI:* Imaging was performed using a 9.4T/105-mm Bruker DMX. T_2 mapping was performed using a 64-echo CPMG pulse sequence (TE/TR = 12.8 ms/ 5 s). T_1 mapping was performed using a progressive saturation sequence with TE = 12.8 ms and TR varying from 100 ms to 15 s. MT data were acquired using the same sequence (TE/TR = 12.8 ms/ 5 s) preceded by a 6 kHz off-resonance saturation pulse incremented from 0.1 to 4.6 s. ADC was measured using a spin-echo sequence with δ = 5 ms and Δ =12.5 ms and gradient strength between 0 and 320 mT/m. Other parameters included NEX = 2, BW = 50 kHz, FOV = 4.0 × 1.5 cm, matrix size = 256 × 128 and slice thickness = 0.5 mm. *Biochemical Tests*: Water, sulfated glycosaminoglycan (sGAG), and collagen content in each sample was quantified using standard biochemical techniques. *Classification of Validation Set Samples:* Univariate: Classification using conventional univariate analysis was performed with group membership of a sample being determined by closeness of MRI parameter values to training set parameter means, with assignment probability determined by Euclidean distance from these means. SVM: Classification through multivariate SVM analysis was performed in feature space after transformation by a radial kernel function [6]. Here, as with the univariate analysis, performance was evaluated with10-fold cross-validation. Group assignment probabilities for each sample were calculated from the sigmoidal distance function describing sample position from the separating hyperplane. ROC curves were constructed from the sensitivity and specificity of correct assignment of validation set samples, with position along the curves parameterized by assignment probabilities. The classification accuracy of a particular analysis is defined by the area under the corresponding ROC curve. Finally, the relationship between sample biochemical content and multivariate SVM assignment classification probability was determined in the validation set to assess the ability of the SVM analysis to characterize matrix composition.

Results Enzymatic degradation resulted in statistically significant changes in MRI and biochemical parameters (data not shown). Fig. 1: Results from the SVM classification of a single trial of control and trypsin-degraded samples in (T_1, k_m, ADC) space. The nonlinear separating hyperplane is indicated. No misclassified samples were observed in this trial. Fig. 2 shows the ROC curves constructed from SVM classifiers previously identified as effective [7] and, for comparison, from the best univariate classifier, T_1,. All multivariate SVM classifiers performed better than T_1. The ability of the SVM construction to predict tissue biochemistry on a sample-by-sample basis is shown in Fig. 3. Fig. 3A demonstrates the correlation between T_1 values and sGAG concentration in the combined trypsin-degraded and control group. While the correlation is fairly strong, Fig. 3B shows that the SVM yields a much stronger predictive accuracy using the triplet (T_1, k_m, ADC). Similar results were obtained for collagen concentration.

Discussion. Univariate classification, as is implicitly used in analyses of cartilage matrix using MRI parameters, exhibits limited ability to discriminate between control and degraded tissue [2]. However, multivariate SVM analysis resulted in substantial improvement. The SVM algorithm constructs a separating hyperplane reflecting training set sample distribution in parameter space. Our 10-fold cross-validation analysis indicated that this construction was also effective for validation sets, consistent with our previous results using Gaussian cluster-based discriminant analysis [7]. Our finding that the sets (T_1, k_m), (T_1, T_2, k_m) and (T_1, k_m, ADC) exhibit particularly favorable classification properties is also consistent with our previous study, indicating that these parameter combinations may emerge as particularly useful in multivariate cartilage matrix characterization [2, 7]. The most successful univariate classifier, T_1, was markedly inferior to multivariate analysis both in classification and in detecting and predicting cartilage matrix changes. This may reflect the effects of particular features of the degraded tissue preferentially affecting particular MR parameters, as well as the fact that the complex tissue changes with degradation may result in offsetting changes in any one particular MR outcome. Fuzzy group assignment, as appropriate for the graded nature of cartilage degradation, is readily implemented in the SVM formalism and showed excellent correlation with tissue biochemistry. We conclude from this that the SVM approach may represent a substantial improvement over the conventional uniparametric analysis for detection of cartilage degradation, for defining degree of partial degradation, and for predicting matrix composition from MR measurements. Finally, we reiterate that certain MR parameter combinations, e.g. (T_1, k_m, ADC), have been found to be consistently effective for classification in a number of settings. This ability to pre-specify an effective parameter set is required for the multivariate approach to be useful in the noninvasive classification and discriminant analysis of cartilage matrix.

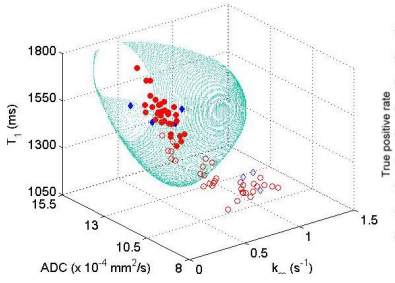

Fig.1: Representative single trial of the SVM classification of control and degraded cartilage samples. In both cases, empty symbols indicate controls and filled symbols degraded samples. Training set:○; validation set: ◊; The separating hyperplane is indicated.

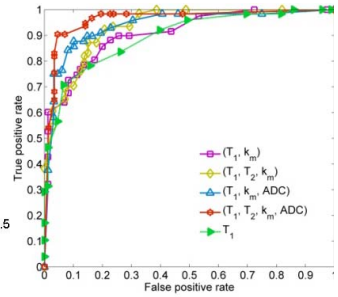

Fig. 2: ROC curves constructed from the best univariate classifier (T_1) and from the multivariate SVM formalism.

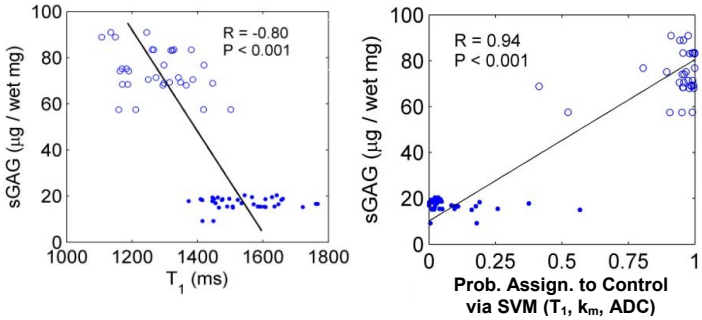

Fig. 3 Accuracy of the best univariate classifier (T_1) (Panel A) and of the parameter set (T_1, k_m, ADC) (Panel B) in predicting tissue sGAG.

References: 1. Laurent D et al. Magn. Reson. Imaging 19 (2001) 1279; **2.** Lin et al. Magn. Reson. Med. 62 (2009) 131 ; **3.** Hastie T et al. J. Mach. Learn. Res. 5 (2004) 1391 ; **4.** Farndale RW et al. Biochem. Biophys. Acta. 883 (1986) 173 ;**5.** Redding GK et al. Clin. Biochem. 29 (1996) 225 ; **6.** Meyer D. R News 1 (2001) 23;**7.** Lin PC et al. J. Magn. Reson. 201 (2009) 61