Classification of prostatic diseases by means of multivariate analysis on in vivo proton MRSI and DCE-MRI data

Mariacristina Valerioa*, Valeria Panebiancob, Alessandro Sciarrač, Marcello Osimanib, Stefano Salsicciac, Lorena Casciania, Alessandro Giulianid, Mariano Bizzarrie, Franco Di Silverioc, Roberto Passariellob and Filippo Contia

Multivariate analysis has been applied on proton magnetic resonance spectroscopic imaging (1H-MRSI) and dynamic contrast enhanced MRI (DCE-MRI) data of patients with different prostatic diseases such as chronic inflammation, fibrosis and adenocarcinoma. Multivariate analysis offers a global view of the entire range of information coming from both the imaging and spectroscopic side of NMR technology, leading to an integrated picture of the system relying upon the entire metabolic and dynamic profile of the studied samples. In this study, we show how this approach, applied to 1H-MRSI/DCE-MRI results, allows us to differentiate among the various prostatic diseases in a non-invasive way with a 100% accuracy. These findings suggest that multivariate analysis of 1H-MRSI/DCE-MRI can significantly improve the diagnostic accuracy for these pathological entities. From a more theoretical point of view, the complementation of a single biomarker approach with an integrated picture of the entire metabolic and dynamic profile allows for a more realistic appreciation of pathological entities. Copyright © 2009 John Wiley & Sons, Ltd.

Keywords: cancer; differential diagnosis; dynamic contrast enhanced MRI; medical imaging; metabolomics; multivariate analysis; prostatic diseases; proton magnetic resonance spectroscopic imaging

INTRODUCTION

Differential radiological diagnosis of both benign and malignant prostatic diseases, such as benign prostatic hyperplasia (BPH), prostatitis and prostate cancer (PrC) is often difficult despite their epidemiological and clinical relevance (1,2). The non-invasive differential diagnosis of prostate cancer, BPH, prostatitis and normal tissue is of utmost importance for cancer staging and for follow-up after therapy (3).

The serum prostate-specific antigen (PSA) has been identified as a sensitive biological marker for prostate cancer diagnosis in recent times. However, the PSA levels, although significantly linked to prostate cancer, are devoid of any discriminatory power for deciding among different prostatic diseases (4,5). In fact, elevated serum PSA levels can be caused by benign conditions that are mainly prevalent in older men (6). To date, histological analysis of biopsy specimens has been the only reliable procedure to distinguish normal, benign and malignant prostatic tissues. However, due to the heterogeneous and frequent multifocal nature of prostate cancer, biopsy methods may not include an adequate specimen sampling of the prostate. In fact, clinical studies have indicated that the usual systematic sextant biopsy technique shows a positive predictive value of only 30% for detection of prostate cancer (7). The use of novel biopsy schemes significantly increases the diagnostic yield of prostate biopsy in finding the malignant disease (8–10), but still does not reach a fully satisfactory accuracy of cancer mapping within the prostate (11).

* Correspondence to: M. Valerio, Department of Chemistry, University of Rome ‘La Sapienza’, P.le A. Moro 5, 00185 Rome, Italy. E-mail: mvalerio@caspur.it

a M. Valerio, L. Casciani, F. Conti
Department of Chemistry, University of Rome ‘La Sapienza’, Rome, Italy

b V. Panebianco, M. Osimani, R. Passariello
Department of Radiological Sciences, University of Rome ‘La Sapienza’, Rome, Italy

c A. Sciarra, S. Salsiccia, F. Di Silverio
Department of Urology, University of Rome ‘La Sapienza’, Rome, Italy

d A. Giuliani
Department of Environment and Health, Istituto Superiore di Sanità, Rome, Italy

e M. Bizzarri
Department of Experimental Medicine, University of Rome ‘La Sapienza’, Rome, Italy

† Deceased

Abbreviations used: AC, adenocarcinoma; BPH, benign prostatic hyperplasia; Cho, choline; CI, chronic inflammation; Cit, citrate; CO, control; Cr, creatine; DCE-MRI, dynamic contrast-enhanced magnetic resonance imaging; FB, fibrosis; 1H-MRSI, proton magnetic resonance imaging; LVs, latent variables; OT, onset time; PCA, principal component analysis; PCs, principal components; PE, peak enhancement; PLS-DA, partial least square discriminant analysis; PrC, prostate cancer; PSA, prostate-specific antigen; ROI, region-of-interest; SI-T, signal intensity-time; TRUS, transrectal ultrasonography; TTP, time to peak; VOI, volume of interest.
Improvements in prostate imaging provide more accurate mapping of cancer allowing for biopsy optimized plans. Needle biopsy guided by transrectal ultrasonography (TRUS) is the most commonly used method for histological diagnosis of prostate cancer. On the other hand, this technique is limited by the lack of accuracy in the estimation of cancer tissue extension (12).

High-resolution endo-rectal/pelvic phased array MRI has demonstrated an enhanced sensitivity compared to clinical data, systematic biopsy, TRUS and MRI when considered alone, but it provides low specificity in detection and localization of prostate cancer due to other benign pathologies which cause low signal intensity on T2-weighted images similar to that of prostate cancer (13–15). Moreover, this conventional imaging method sometimes cannot distinguish accurately between healthy and malignant tissues subsequent to resection or treatments for prostate cancer because of induced changes in tissue structure (16). The ability to identify prostate tissue types can be significantly improved by the combined use of MRI and magnetic resonance spectroscopic imaging (MRSI) (17,18). Specifically, proton MRSI ($^1$H-MRSI) allows the recognition of some relevant metabolites like citrate (Cit), choline (Cho) and creatine (Cr), endowed with the discrimination power for different prostate diseases. In particular, the ratio choline-plus-creatine to citrate ((Cho + Cr)/Cit) has been widely investigated in the differential diagnosis of prostate diseases. The rationale at the basis of the use of this index stems from the observed increase of Cho-containing compounds and on the reduction or absence of Cit in PrC compared to BPH and the surrounding healthy peripheral zone tissue. However, in some cases, it has been reported that glandular BPH and normal, healthy peripheral zone tissues display similar Cit levels; on the other hand, stromal BPH regions can show reduced Cit levels similar to those observed in peripheral zone cancers (19–26). In a recent paper, Li et al. (27) analysed the spectral differences between PrC and BPH, evaluating the (Cho + Cr)/Cit and Cho/Cr ratios measured in each voxel with proven-biopsy cancer or BPH. The specificity, sensitivity and accuracy for the discriminant function were 98.6, 85.7, 92.9%, respectively. Finally, the addition of metabolic information provided by MRSI to morphologic information provided an enhanced specificity up to 95% for localizing cancer and has become important for detecting the extent of cancer within the prostate and its aggressiveness (17,18,28). Despite the very high accuracy of the MRI/MRSI combined approach in the detection of prostate cancers, this mixed technique is not very accurate in identifying cancers (only 30% are found) in both the central gland (29) and small (<0.5 cm$^3$) cancers within the peripheral zone (21,30).

Recent studies have shown that accuracy can be improved by performing MRI/MRSI at higher magnetic field strengths and through the addition of dynamic contrast enhanced MRI (DCE-MRI) (31–33).

DCE-MRI is known to be a powerful tool for visualizing the vascularity of prostate gland tumour and for providing additional information useful for both the detection and the staging of prostate cancer (31,34). Dynamic MRI was demonstrated to be able to differentiate cancer from normal prostatic tissue (35,36), an earlier and stronger enhancement in PrC versus normal tissue was found in these studies. DCE-MRI was also demonstrated to differentiate the cancer tissue from benign lesions (37,38). Recently, Ren et al. (39) reported that, based on T2-weighted imaging, DCE-MRI curves can discriminate PrC and BPH with a sensitivity, specificity and accuracy of 79.31, 66.67 and 74%, respectively. In particular, PrC showed stronger enhancement with an earlier peak time, higher enhancement and enhancement rate than those of BPH.

In this study, we developed a computational approach to the analysis of $^1$H-MRSI and DCE-MRI combined results, that was demonstrated to outperform both the classical ((Cho + Cr)/Cit) ratio and purely image-based parameters in the differential diagnosis of prostatic diseases.

Our data set was composed of 11 prostatic tissue samples coming from healthy control subjects, and of 40 pathologic tissue samples coming from patients affected by one of the following pathologies: (a) chronic inflammation, (b) fibrosis or (c) adenocarcinoma. The discrimination between healthy and disease samples and the differential diagnosis of the various prostatic diseases was the aim of the present study. Given the paucity of the data set, our work has mainly methodological value, giving a proof-of-concept to the feasibility of an integrative systemic view, made possible by multivariate approach, to the MRI/MRSI non-invasive differential diagnosis of prostatic diseases.

We adopted a data analysis strategy that mixed unsupervised and supervised approaches. Unsupervised is the general heading of techniques such as principal component analysis (PCA) and cluster analysis, whose final result is not guided by the maximization of ‘externally’ imposed classification goals, like the discrimination between different classes of disease or placebo and drug treated patients. The goal of unsupervised algorithms is to maximize some purely syntactical internal features of the data set at hand, like the obtaining of the most faithful projection of an initially high dimensional data set with the least number of axes (PCA), or the allocation of the statistical units to classes that are the most internally compact and separated (cluster analysis) (40). Unsupervised methods allow for an unbiased (not driven by the goal of diagnosis) description of the natural correlation structure present in the data. In contrast, supervised methods that we adopted in our strategy, like partial least squares (PLS) or discriminant analysis (DA), have the goal of maximizing an ‘externally imposed task’, such as the separation of two a priori classes (like healthy/disease) inside a given data field. In fact, supervised methods are driven by a specific goal, external to the intrinsic nature of the collected data, assuring the ‘best possible discrimination’ of the classes, at the expense of the appreciation of the natural correlations present in the data, which we exploited with the unsupervised approach.

We decided to complement the unsupervised and supervised approaches, with a classical strategy of data analysis (41) so as to get the maximal global efficiency of the model on both descriptive and diagnostic sides. The initial unsupervised extraction of principal components from the original data set allows for noise filtering of the data, maintaining only the correlated (and thus more reliable) portion of information and permitting a biological interpretation of the obtained results. Subsequently, the use of the extracted principal components as initial variables for the supervised portion of the procedure, avoids possible inconsistencies coming from the regressors’ mutual collinearity components which are orthogonally constructed. However, besides statistical subtleties, the important thing to stress is that the adopted strategy allows for all the information embedded in our data to be potentially exploited for the task of classification.

All in all, the obtained results allowed for both an efficient discrimination of the different diseases and for a biologically sound general picture of the system at hand.
METHODS

Patient population

In this study, we retrospectively reviewed a total of 51 prostate MR examinations, including morphologic imaging, \(^1\)H-MRSI and DCE-MRI protocols, carried out from June 2007 to February 2008. The MR investigations were performed before TRUS-guided needle biopsy. The mean time interval between MR examination and TRUS-guided biopsy was 7 ± 3 days. For each patient, medical histories including digital rectal examination, serum PSA level and a confirmed biopsy report were obtained. Our study population can be subdivided into four groups based on the kind of prostate disease clinically and histopathologically diagnosed. The first group consisted of 13 patients with biopsy-proven PrC who subsequently underwent radical retropubic prostatectomy performed within 3 weeks (mean = 9 ± 4 days). Histopathological examination of the radical prostatectomy specimen revealed a mean Gleason score of 7 ± 2 (range = 6–10) and the absence of BPH nodules. In the second group, we included 14 patients with biopsy-proven BPH who subsequently underwent trans-urethral resection. Pathologic assessment of the resected tissue confirmed TRUS results. The mean period between MR examinations and trans-urethral resection was 12 ± 7 days. The third group was composed of 13 patients with a biopsy-proven fibrosis performed for a TRUS suspicious prostate nodule. A control group of 11 patients with no pathological findings of PrC, BPH and fibrosis was also evaluated. In this group, a prostate TRUS-biopsy was performed following a suspicious clinical examination and/or rising serum PSA level (mean = 4.9 ± 7.0 ng/mL; median = 1.7 ng/mL; range = 0.5–18 ng/mL). Furthermore, digital rectal examination did not reveal any prostate abnormalities.

The following exclusion criteria were used in the selection of patient population: clinical diagnosis of acute prostatitis, coexistent clinically proven cancer, hormonal therapies (including five \(\alpha\) reductase inhibitors), radiotherapy, chemotherapy, previous prostate surgery.

The 51 included patients had a mean age of 65 (range 48–75 years); age was checked for its possible confounding effect on the MR-based parameters, without finding any significant correlation with the descriptors under study (data not shown).

This study was approved by the local ethics committee and signed informed consent was obtained from all patients.

MRI, \(^1\)H-MRSI and DCE-MRI data acquisition and processing

Acquisition of imaging data

All examinations were performed on a commercially available 1.5T scanner (Magneton Avanto, Siemens Medical Solutions, Erlangen, Germany), equipped with surface phased-array (Body Matrix, Siemens Medical Solutions) and endo-rectal coil (e-Coil, Medrad, Pittsburgh, PA, USA, combined with Endo-Interface, Siemens Medical Solutions). The balloon-mounted disposable endo-rectal coil was first lubricated with a local anaesthetic gel and then placed while the patient was in the left lateral decubitus position. Then the patient was turned supine and the balloon was inflated with up to 70 mL of room air based on the patient’s tolerance. Before scanning, 20 mg butyl scopalamine (Buscopan, Boehringer, Ingelheim, Germany) was injected to suppress peristalsis.

First, localizer images in the sagittal, axial and coronal planes were obtained to ensure endo-rectal coil position and to select locations for the transverse images. Following this, T2-weighted images in the three orthogonal planes were acquired providing coverage of the entire prostate using turbo spin-echo (TSE) sequences (TR = 5190 ms; TE = 95 ms; flip angle = 150°; average = 3; FOV read = 256 mm; FOV phase = 100; slice thickness = 3 mm; interslice gap = 0; matrix size = 512 × 512; phase resolution = 100%; bandwidth = 130 Hz; scan time = 3.40 min).

\(^1\)H-MRSI data were acquired by two skilled radiologists after a first review of morphological images to localize suspicious areas in the prostate, which were subsequently used to position the spectroscopic acquisition volume. In patients with no MR morphological evidence of changes, the volumes of interest (VOIs) were centred on each prostate gland emi-portion (left and right). The VOI to be studied with spectroscopy was selected in such a way, as to maximize the coverage of the prostate while minimizing the inclusion of surrounding structures (muscles, fat, rectal air and urine). \(^1\)H-MRSI was performed using a section-selected box drawn closely around the prostate fossa and a point-resolved spectroscopic sequence was obtained by using a 3D chemical shift imaging (CSI) sequence (FOV = 50 × 50 × 50 mm\(^3\); VOI = 30 × 30 × 30 mm\(^3\); TR = 700 ms; TE = 120 ms; ms; flip angle = 90°; interolation = 16; vector size = 512; TA = 11.50 min; delta frequency = −1.80 ppm; average = 6; filter = Hamming) (42).

DCE-MRI images were acquired using 3D FLASH T1-weighted spoiled gradient-echo sequence (TR = 2.44 ms; TE = 0.9 ms; flip angle = 30°; average = 1; thickness = 4 mm; interslice gap = 0; slice number = 12; matrix size = 256 × 256; phase resolution = 100%; bandwidth = 120 Hz; TA = 4.40 min) performing 90 measurements in rapid succession, immediately after the completion of an intravenous bolus injection of 0.1 mmol of gadopentetate dimeglumine (Multihance, Bracco Spa, Milano, Italy). Contrast liquid was administered with a power injector (Spectris; Medrad) at 2.5 mL/s and was followed by a 15-ml saline flush. The 3D volume was acquired with the same positioning angle and centre as the transverse T2-weighted sequence covering the entire prostate fossa and the periurethral-perianastomotic region. Relative gadolinium chelate concentration curves were calculated in order to derive the three dynamic DCE-MRI parameters: onset time, time to peak and peak enhancement (PE).

Processing and analysis of imaging data

MR images were analysed in consensus by two radiologists with 5 and 9 years of experience in uro-genital MRI. They were unaware of serum PSA levels and TRUS-biopsy results. T2-weighted images were excluded from retrospective reviewing and the radiologists’ attempts were focused only on spectroscopy and DCE-MRI findings.

An operator-independent standard post-processing protocol was applied to the MR spectroscopic imaging data. These data were acquired as 16 × 8 × 8 phase-encoded spectral arrays (1024 voxels) with a nominal spatial resolution >0.3 cm\(^3\) before Fourier transformation in the spatial dimensions. After Fourier transformation, zero- and first-order phase correction and automated baseline correction (polynomial of 6th order), a frequency domain curve fitting was used subsequently for quantification with the assumption of Gaussian line shapes, by using the standard Syno Spectroscopic Evaluation software package (Siemens), provided with the MR imaging system (43). Goodness of fit of the obtained parameter by means of the classical
Gaussian distribution hypothesis was assessed by subtracting the processed spectrum from the fitted one, and checking that only signals indistinguishable from the baseline noise remained at the mentioned ppm in the residual curves. We used areas under the curve to compute Cho (3.2 ppm), Cr (3.0 ppm) and Cit (2.6 ppm) values. The average post-processing duration was 25–30 min for each data set. Cho, Cr and Cit peak areas were evaluated for all \(^1\)H-MRSI voxels not contaminated by inadequately suppressed water or lipids, and did not contain mixed tissues from the urethra, seminal vesicles, ejaculatory ducts, and bladder and rectal wall. In addition, only voxels in which the choline/creatinine/citrate peaks were detectable with a signal-to-noise ratio of 3:1 were assessed (44).

The dynamic MR post-processing procedure lasted 10 min for each patient. Functional dynamic imaging parameters were estimated from the gadolinium curve using the procedure of Fütterer et al. (19). The edge and the contour characteristics of the lesions were defined using the same sections on which the region-of-interest (ROI) analyses were performed.

A group of three ROIs were drawn independently by the radiologists, and differences in the measurements were assessed by consensus. The selected ROIs referred to three distinct areas: (1) pelvic muscle (acting as low baseline ROI), (2) most-enhancing areas within the main PrC foci, BPH nodules, fibrosis nodules or in these regions with suspect spectroscopic voxels and (3) iliac vessel (acting as high baseline ROI). In particular, suspect regions were identified based on higher enhancing values on DCE-MRI images (qualitative method). Correspondingly, normal tissue was identified as the one having homogenous enhancing regions.

When multiple suspicious areas were identified, the signal intensity–time (SI–T) records of the most enhancing lesion were considered as significant values for subsequent SI–T analyses. The following parameters were set to describe the SI time curve: onset time, time to peak and PE. We determined the enhancement onset time for the data sets by averaging (during 90 measurements) the intensity across the slices and using the last point before the averaged signal increased 2.5 standard deviations (SDs) above the running baseline average.

**TRUS-biopsy evaluation**

TRUS guided biopsies were performed using a biplanar 7.5 MHz frequency probe according to a 12-core biopsy scheme (standard sextant scheme, plus laterally directed samples of the prostate apex (two cores), middle (two cores) and base (two cores)) (10). No samples in the transitional zone of the prostate were obtained. The operator was blinded from \(^1\)H-MRSI and DCE-MRI results. All biopsy specimens were obtained under TRUS guidance using an 18-gauge needle loaded in an automatic spring action biopsy device, and were fixed overnight in a solution of 10% neutral buffered formalin. The operator evaluated the distance from the prostate apex and basis and the distance from the urethra from which the biopsy specimen was drawn, in order to provide a method of comparison with T2-weighted MR images as reference. MRI films were interpreted independently by a third radiologist with 6 years of experience in uro-genital radiology, who had no knowledge of \(^1\)H-MRSI and DCE-MRI results and of the final diagnosis. Each axial T2-weighted image was branched in 12 radial triangles with apex orientation on the urethra in a clockwise order. Distance from the prostate apex and basis and the distance from the centre were noted by the radiologist. For each sample, all abnormalities were examined in consensus by the biopsy operator and radiologist.

**Statistical analysis**

**Raw data matrix structure**

Multivariate analysis was applied to the data set constituted by the \(^1\)H-MRSI spectral and dynamic DCE-MRI parameters measured on healthy prostatic tissues of 11 control (CO) as well as on the pathologic tissues of 40 patients who had one of the following pathologies: (1) chronic inflammation (CI, n = 14); (2) a fibrosis (FB, n = 13) or (3) adenocarcinoma (AC, n = 13). This produced a raw data set constituted by a matrix having as rows (statistical units) 51 patients and as column (variables) the 6 values relative to choline, creatine, citrate (all these three variables are expressed in terms of the area of the relative peak), onset time, time to peak (these dynamic descriptors are expressed in seconds) and PE parameters (\(c \times (\text{mmol/kg contrast agent})\) obtained from \(^1\)H-MRSI/DCE-MRI measurements.

The original 51 units in the population were separated into two sets: a training and a validation (test) set. The training set was made of 45 subjects, while 6 subjects (2 CI, 2FB, 2AC) composed the test set. The model was built upon the 45 subjects in the training set and then was checked on its ability to correctly classify the set of six patients. This procedure known as cross-validation allows for the testing of the generalization ability of the proposed model, outside the range of the specific data set it comes from. This eliminates overfitting and chance correlation problems, which is particularly important in this case, where the paucity of data greatly increases the risk of apparent correlations (41).

**Principal component analysis**

PCA is a projection method used for exploiting the information embedded in multidimensional data sets (40). The data are reduced to a few latent variables (LVs) (or principal components) collecting the information implicit in the original variables' correlation structure. The extracted components (PCs) are each orthogonal and ordered in terms of the percentage of explained variation, with the first components collecting the 'signal' (correlated) portion of information, while minor components can be considered as 'noise' components. From an algebraic point of view, each component is a weighted summation computed across the original variables in the form of:

\[
\text{PC} = aX_1 + bX_2 + cX_3
\]

where \(X_1, X_2 \) and \(X_3\) are the measured variables and \(a, b, c\) are the numerical constants. Each statistical unit is assigned a score relative to each extracted component, while the correlation coefficient between each original variable and extracted components (loading) allows us to give a meaning to the PCs.

**Partial least square discriminant analysis**

While PCA is an unsupervised technique in which each variable enters with the same role of description of the data set and the solution is driven by the maximal parsimony principle alone (maximal amount of explained variation with the minimum number of components); on the other hand, both PLS and DA are supervised techniques in which the analysed variables pertain to two classes: the ‘diagnosis’ (dependent, Y) and ‘symptoms’ (independent, X) variables. The goal of both PLS and DA is to find the linear combination of X variables that explains the Y

---

**www.interscience.wiley.com/journal/nbm**

variable(s) better. In the case of DA, this goal is achieved by the construction of a set of weights multiplying each X variable so as to build metrics in which the errors of assignment of each statistical unit to the correct Y class is minimized. PLS, on the other hand, works by the generation of mutually orthogonal linear combinations of X variables maximally correlated with Y counterparts. PLS-DA was used to build and test a supervised model that could predict the pathology of a patient based on its spectral and dynamic data. The 'leave-one-out' cross validation method was used to validate the model and to select the appropriate number of LV. The identification and the removal of outliers were performed by using the Q and T² statistics (45).

Range restriction effect check

The presence of outliers is known to deeply influence the correlation analyses by means of the so called range restriction effect (46). This effect has to do with the fact that outliers (extreme values observations) acquire a disproportionate weight in the computation of any least-squares- based model. In fact, optimizing fit of an extreme point (outlier) has a much greater influence in the sum of errors (to be minimized by the system) than the best fitting of an average unit. This statistical effect could give rise to biased descriptions of the studied data sets. For this reason, we decided to perform two independent analyses: the first one on the complete data set (45 units) and the second one on the same data sets after removal of four possible outliers (41 units). We considered as outliers the units being located at more than 3 SD units from the centroid of the data cloud.

RESULTS

Multivariate analysis of ¹H-MRSI and DCE-MRI data

Individual and mean values of spectral ¹H-MRSI and dynamic DCE-MRI parameters measured on healthy prostatic tissues of control as well as on the pathologic tissues of patients with chronic inflammation, fibrosis or adenocarcinoma are shown in Fig. 1. The data indicate the presence of a substantial overlap among groups, even if we must take into account the fact that the relatively high variances of the considered variables is partly due to the coil sensitivity profiles. Nevertheless, correction factors are not routinely used in such kind of analyses. This figure allows to immediately perceive the huge advancement obtained by explicitly taking into account the fact that the same statistical unit (patient) is simultaneously defined over different variables. In fact, the possibility of distinction relies on the correlation between variables instead of considering each descriptor in isolation from all the others (47).

Before entering the actual statistical strategy with the explicit consideration of the correlation structure of the measured features, we adopted the classical approach applied in this category of problems by submitting the descriptors most widely used for differential diagnosis to DA. The discriminant function was computed over the (Cho + Cr)/Cit ratio and OT, TTP and PE. The specificity, sensitivity and accuracy are reported in Table 1 for differences between class comparisons. Although all the considered descriptors displayed a statistically significant discriminative power, there are some specific inter-group comparisons that were markedly sub-standard. In contrast, we will show how the multivariate approach not only reaches maximal accuracy on the entire set of comparisons, but how it is also able to correctly predict the test set statistical units, thus demonstrating a predictive ability outside the realm of the samples used to build the model.

Data analysis begun with the computation of PCA to identify any clustering of data related to the types of pathology in an unsupervised manner, while PLS-DA was subsequently applied to build a classification model to predict the clinical outcome of a patient based on its spectral and dynamic data and usable even for independently analysed patients (test set) (45).

PCA applied to the original 45 units/6 variables data set ((Cho + Cr)/Cit ratio is a derived variable and its information is implicit in the original descriptors), gave rise to a four component solution explaining about 87% of the total variability in the system. In Table 2 the variance explained by each component is reported. To compare the controls and patients (as a whole), a t-test was applied to the component scores, highlighting significant differences between the two groups on PC1 and PC3 (see Table 2). This can be appreciated in Fig. 2, where the component score plot is shown. A linear discriminant analysis applied to this space allowed for a clear separation of the two groups (Fisher's exact test, p < 0.0001 on the classification matrix).

The t-test on PCs is then repeated to compare all the specific pairs of control and pathology groups (see Table 2). As a result, CI and AC patients showed significant differences compared to controls on both PC1 and PC3 (Fisher's exact test p < 0.0001), whereas the FB group differed from controls as for PC2 and PC3 (Fisher's exact test p < 0.0001). Figures 2 and 3 display the above-mentioned differences in various component planes.

We subsequently analysed the data by performing a t-test on PCs to compare a single pathology (see Table 2). PC1 was responsible for the discrimination among pathologies and pairwise discriminations are also observed on PC2 and PC4 (Fisher's exact test p < 0.0001) for CI versus FB as well as CI versus AC groups (see Fig. 4). The above results indicate that even by using this unsupervised method of analysis, with only a posteriori computation of the statistical significance over the extracted components, there is a clear separation between controls and pathologic groups, controls and single-pathologic groups and among patients with different types of pathology.

After having checked for the general robustness of the correlation structure, we considered the possibility of using multivariate methods to classify patients not used for model building. This means shifting from a purely descriptive to a practical diagnostic use of the technique that by definition must be effective in predicting the diagnosis of samples not explicitly taken into account for statistical model building. For this goal, we built a PLS-DA model based on 45 patients data set. This method, at odds with PCA, is a supervised procedure explicitly driven by the optimization of discrimination power of the model. The feasibility of the proposed method as a routine diagnostic procedure depends on the successful classification of the test set.

The PLS-DA model generated three LVs which explain 70% of the X-variance (spectral and dynamic data) and 50% of the Y-variance (which represent the membership class). Table 3 summarizes the features of this model. A clear separation among the classes was found in the first and second components (Fig. 5) thus confirming the unsupervised approach (PCA). This model was used to assess the predictive capabilities for six other patients with unknown pathology. The model predicted that two of the unknown samples were from the CI group, two from the FB and two from the AC group. The actual identities of the samples coincided with the predicted ones (i.e. the sensitivity and...
specificity for the PLS model based on pathology are both 100%), thus providing a validation of the PLS model as a diagnostic tool (see Table 4).

Having proven the efficacy of the model as a predictor, we decided to go more in depth into the nature of the metabolic and dynamic biomarkers which permit discrimination. To this aim, we derived a PLS model using a reduced data set by eliminating outliers that could bias the solution by their excessive weight in computing the correlation. The outliers were removed by means of the procedure previously outlined in the Method section. The removal of outliers provided a final 41 statistical units and a raw data matrix of six variables. The loading plot relative to the discriminating variables is shown superimposed over a score plot in Fig. 6. This representation allows us to simultaneously appreciate the discrimination power (position of the patients in the plane) and the functional meaning (the loadings correspond to the correlation coefficients of the original variables with the axes) of the proposed solution.

From Fig. 6, it is evident that choline concentration and PE are crucial parameters to discriminate between the benign and malignant diseases. Choline has a high negative correlation coefficient with LV1, while PE shows a strong negative correlation with LV1 and LV2. Since the malignant group displays lower scores in both LV1 and LV2 compared to benign groups, this corresponds to the fact that patients with adenocarcinoma have higher values of choline and PE compared to patients with benign pathologies. Concerning the differential diagnosis of benign pathologies, the descriptors endowed with the highest discrimination power for FB patients are the scoring of a high level of citrate (loadings = 0.460 and −0.300 on LV1 and LV2,

**Figure 1.** Plot of individual (open circle) and mean (filled circle) values of spectral $^1$H-MRSI (area) and dynamic DCE-MRI (onset time and time to peak in seconds; peak enhancement in $c \times$ (mmol/kg contrast agent) parameters measured on healthy prostatic tissue of normal control as well as on the pathologic tissue of patients with chronic inflammation, fibrosis or adenocarcinoma.
respectively) and the presence of a comparatively high value of time-to-peak (loadings $= 0.308$ and $-0.147$ on LV1 and LV2, respectively). It is worth noting that some patients with fibrosis are characterized by high values of onset time (loadings $= 0.630$ and $0.133$ on LV1 and LV2, respectively). Furthermore, chronic inflammation provokes an increase of creatine level (high positive correlation with LV2, loading $= 0.605$) with respect to other benign pathologies.

**DISCUSSION**

Due to the increased use of both serum PSA screening and TRUS-guided biopsies, prostate cancer is being identified at an earlier and more treatable stage (48). Therefore there is an increased interest in routine check-ups, but clinical parameters alone are not sufficient to predict the course of a benign disease. In fact, the risk of over-detection has been estimated to vary between 15 and 84% (49,50). Current classification systems are able to predict only a binary outcome, i.e. benign or malignant, with sensitivities of 95 and 73% as well as specificities of 91 and 81% for $^1$H-MRSI and DCE-MRI, respectively (17,51). The most widely used metabolic classification method is based on previously reported differences between cancer and normal prostate tissue; voxels are considered suspicious for cancer if the $\frac{[\text{Cho} + \text{Cr}]}{\text{Cit}}$ ratio is at least 2 SDs above the average ratio for the normal peripheral zone, and voxels are considered very suspicious for cancer if $\frac{[\text{Cho} + \text{Cr}]}{\text{Cit}}$ ratio is more than 3 SDs above the average ratio (21,22). However, other conditions such as prostatitis or post biopsy haemorrhage might increase the $\frac{[\text{Cho} + \text{Cr}]}{\text{Cit}}$ ratio, and also normal prostatic tissue may show higher $\frac{[\text{Cho} + \text{Cr}]}{\text{Cit}}$ ratios within the transitional and periurethral tissue (52).

### Table 1. Discriminant function analysis comparing controls versus single pathology and single pathology versus single pathology with regard to spectroscopic $^1$H-MRSI or dynamic DCE-MRI or both spectroscopic and dynamic parameters

<table>
<thead>
<tr>
<th></th>
<th>Control vs. chronic inflammation</th>
<th>Control vs. fibrosis</th>
<th>Control vs. adenocarcinoma</th>
<th>Chronic inflammation vs. fibrosis</th>
<th>Chronic inflammation vs. adenocarcinoma</th>
<th>Fibrosis vs. adenocarcinoma</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-value</td>
<td>34.45</td>
<td>1.34</td>
<td>7.03</td>
<td>46.82</td>
<td>5.54</td>
<td>8.49</td>
</tr>
<tr>
<td>Specificity</td>
<td>100</td>
<td>27</td>
<td>46</td>
<td>100</td>
<td>31</td>
<td>46</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>100</td>
<td>77</td>
<td>100</td>
<td>93</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Accuracy</td>
<td>100</td>
<td>54</td>
<td>71</td>
<td>96</td>
<td>67</td>
<td>73</td>
</tr>
<tr>
<td>(Cho + Cre)/Cit ratio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Onset time, time to peak and peak enhancement parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-value</td>
<td>2.31</td>
<td>8.003</td>
<td>53.81</td>
<td>9.22</td>
<td>31.43</td>
<td>49.53</td>
</tr>
<tr>
<td>Specificity</td>
<td>64</td>
<td>100</td>
<td>100</td>
<td>85</td>
<td>92</td>
<td>100</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>79</td>
<td>85</td>
<td>100</td>
<td>93</td>
<td>93</td>
<td>100</td>
</tr>
<tr>
<td>Accuracy</td>
<td>72</td>
<td>92</td>
<td>100</td>
<td>89</td>
<td>93</td>
<td>100</td>
</tr>
<tr>
<td>(Cho + Cre)/Cit ratio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Onset time, time to peak and peak enhancement parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-value</td>
<td>12.19</td>
<td>6.05</td>
<td>39.12</td>
<td>15.39</td>
<td>22.90</td>
<td>38.24</td>
</tr>
<tr>
<td>Specificity</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>92</td>
<td>100</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>93</td>
<td>85</td>
<td>100</td>
<td>93</td>
<td>93</td>
<td>100</td>
</tr>
<tr>
<td>Accuracy</td>
<td>96</td>
<td>92</td>
<td>100</td>
<td>96</td>
<td>93</td>
<td>100</td>
</tr>
</tbody>
</table>

For each discriminant function the specificity, sensitivity and accuracy are reported.

### Table 2. t-test comparing pathologic patients on the whole versus controls, controls versus single pathology and single pathology versus single pathology

<table>
<thead>
<tr>
<th>Patient group</th>
<th>Statistical significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PC1 (35.43)</td>
</tr>
<tr>
<td>Pathologic patients on the whole vs. controls</td>
<td>0.004*</td>
</tr>
<tr>
<td>Controls vs. chronic inflammation</td>
<td>0.043*</td>
</tr>
<tr>
<td>Controls vs. fibrosis</td>
<td>0.144</td>
</tr>
<tr>
<td>Controls vs. adenocarcinoma</td>
<td>0.0001*</td>
</tr>
<tr>
<td>Chronic inflammation vs. fibrosis</td>
<td>0.003*</td>
</tr>
<tr>
<td>Chronic inflammation vs. adenocarcinoma</td>
<td>0.0001*</td>
</tr>
<tr>
<td>Fibrosis vs. adenocarcinoma</td>
<td>0.0001*</td>
</tr>
</tbody>
</table>

*Threshold $p < 0.05$.

In parentheses the percent of variance explained by each principal component is reported.
although average values and SDs for [Cho + Cr]/Cit in healthy prostate tissue have been available, to date, it remains unclear if these values are independent of the respective MR scanner and/or MRS sequence applied. In the Shukla-Dave et al. (53) study, Cho was elevated in 9 out of 12 patients with histopathologically confirmed chronic prostatitis, and 86% of the voxels indicated intermediate or high-grade diseases.

The diagnostic value of DCE-MRI in histologically proven benign and malignant prostate tissues has been evaluated by several studies (54,55), which postulated that prostate cancer showed earlier and stronger enhancement than normal tissue. In particular, Ren et al. (39) demonstrated the potential of DCE-MRI to distinguish between BPH nodules and PrC foci; the time-to-peak of PrC lesion occurred earlier than the BPH peak time and the enhancement degree and rate of PrC were higher than those of BPH. On the other hand, limitations of the technique including inadequate lesion characterization, particularly in the differentiation of prostatitis from cancer in the peripheral gland and in the discrimination between BPH and central gland tumours has been established (34).

The combined use of ¹H-MRSI and DCE-MRI techniques could be able to address the limitations found in the two techniques when used independently, improving the prediction accuracy (19). van Dorsten et al. (56) showed that the addition of ¹H-MRSI and DCE-MRI to the conventional MRI protocol has great potential for improved localization and characterization of prostate cancer in a clinical setting.

Our results represent a further improvement along this line. Multivariate analysis is much more efficient in discrimination than in the use of original variables as they do not allow for an equally precise discrimination: the ([Cho + Cr]/Cit) ratio, the most discriminant ¹H-MRSI index, fails to separate between control and FB groups, whereas dynamic parameters are not able to separate the CI/CO groups. Furthermore, the classical, combined

Table 3. PLS-DA model summary for discriminating ¹H-MRSI/DCE-MRI data from patients with both benign and malignant prostatic diseases

<table>
<thead>
<tr>
<th>Model</th>
<th>LV</th>
<th>R²X</th>
<th>R²Y</th>
<th>Q²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>45 units</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LV1</td>
<td>0.353</td>
<td>0.273</td>
<td>0.249</td>
</tr>
<tr>
<td></td>
<td>LV2</td>
<td>0.187</td>
<td>0.172</td>
<td>0.164</td>
</tr>
<tr>
<td></td>
<td>LV3</td>
<td>0.158</td>
<td>0.099</td>
<td>0.078</td>
</tr>
<tr>
<td></td>
<td>41 units</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LV1</td>
<td>0.364</td>
<td>0.283</td>
<td>0.274</td>
</tr>
<tr>
<td></td>
<td>LV2</td>
<td>0.155</td>
<td>0.185</td>
<td>0.159</td>
</tr>
<tr>
<td></td>
<td>LV3</td>
<td>0.151</td>
<td>0.102</td>
<td>0.069</td>
</tr>
</tbody>
</table>

R²X, cumulative fraction of the variation of the X variable explained per component; R²Y, cumulative fraction of the variation of the Y variable explained per component; Q², the cumulative predicted fraction (cross-validation) of the variation joint X and Y.
1H-MRSI and DCE-MRI analysis showed a lower specificity, sensitivity and accuracy compared to those obtained for the PLS model.

The first multivariate data analysis applied was PCA. The application of this method embodies a sort of ‘natural normalization’ of the studied data set, given that principal components correspond to the eigenvectors of the correlation matrix that in turn corresponds by definition to the covariance matrix of the standardized variables. This is particularly convenient when dealing with heterogeneous variables defined by completely different measurement units (57), ruling out all questionable a priori defined standardization processes.

PCA showed a natural trend to clustering of the prostatic diseases in the MR space. This is a proof-of-concept of the possibility of obtaining metabolic and morphological fingerprints useful for the differential diagnosis of prostatic diseases, even without specifically imposing the discrimination task to the model. Principal components are orthogonally constructed, thus the different relevance of the components for the discrimination of diverse prostatic diseases is the image in light of the biological differences between pathologies (58).

This clustering tendency of the prostatic diseases in the MR space was confirmed in terms of diagnostic accuracy, by a PLS methodology that highlighted choline, creatine and citrate as the main discriminant metabolites among different prostate diseases. The dynamic parameters endowed with the highest clinical significance were onset time, time-to-peak and PE. Taken altogether, metabolic and dynamic descriptors allowed us to obtain a correct reclassification of an independent test set in addition to a complete classification of the training set. The prediction of the six patients in the test set confirms the exact discrimination already found in the general (45 patients) data set, adding the dimension of the generalization ability to the pure internal consistency of the model.

To investigate the metabolic and dynamic biomarkers, we obtained a PLS model using a reduced data set by removing the outlier (41 patients). The reproduction of the same correlation structure, by means of a data set depurated by the most extreme statistical units, is proof of the fact that we can safely rule out a ‘range restriction effect’ (46) as a possible source of confounding for our results. The fact that both the PLS analyses gave rise to the same result is an important proof of the robustness of the classification.

As we mentioned in the Results section, the two latent variables (LV1 and LV2) endowed with the highest discrimination power were mainly related to the opposition between choline and onset time for LV1 (these two variables are at the opposite poles of the LV1 axis in Fig. 6) and PE and creatine for LV2 (extreme opposite poles of the LV2 axis in Fig. 6). As for the LV1 axis, the high choline pole matches to adenocarcinoma patients (black squares in Fig. 6), while the opposite LV pole corresponding to high values of onset time is the preferred location of fibrosis patients (black circles in Fig. 6). As for the LV2 axis, near the creatine (Cr) pole in the LV space we observe the chronic inflammation patients (black triangles in Fig. 6), whereas the adenocarcinoma patients shifted toward the direction of the high PE pole. The increased efficiency of both PCA and PLS-DA analyses is a natural consequence of the fact that both

Table 4. Membership score (correlation coefficient) for each of the six samples derived from the PLS model built using the whole data set (45 units/6 variables)

<table>
<thead>
<tr>
<th>Sample</th>
<th>Control</th>
<th>Chronic inflammation</th>
<th>Fibrosis</th>
<th>Adenocarcinoma</th>
</tr>
</thead>
<tbody>
<tr>
<td>CI</td>
<td>0.233</td>
<td>0.428</td>
<td>0.126</td>
<td>0.212</td>
</tr>
<tr>
<td>CI</td>
<td>0.340</td>
<td>0.501</td>
<td>0.041</td>
<td>0.118</td>
</tr>
<tr>
<td>FB</td>
<td>0.323</td>
<td>−0.008</td>
<td>0.852</td>
<td>−0.168</td>
</tr>
<tr>
<td>AC</td>
<td>0.048</td>
<td>0.107</td>
<td>−0.028</td>
<td>0.873</td>
</tr>
<tr>
<td>AC</td>
<td>0.011</td>
<td>0.355</td>
<td>−0.058</td>
<td>0.693</td>
</tr>
<tr>
<td>FB</td>
<td>0.175</td>
<td>0.069</td>
<td>0.813</td>
<td>−0.057</td>
</tr>
</tbody>
</table>

The sample column reports the effective clinical status of the patient, the other columns the allocation coefficient made by the system on the sole basis of MR information.
techniques filter out the noisy portion of information into minor components while concentrating on the most informative portion of information retained by the major axes (40,45).

Our data are consistent with the well-known increased degree of vascularization in tumour pathologies that scale with PE, while the discriminant ability of choline can be related to the changes in cell membrane synthesis and degradation of tumour tissues that go hand in hand with an increased choline concentration (19–26,39).

Beside the mechanistic interpretation of the results, we can safely affirm that the non-invasive acquisition of $^1$H-MRSI/ DCE-MRI data is a potentially valid approach in both the differential diagnosis and treatment evaluation of prostatic diseases.

Our study can be considered as a pilot study: the paucity of the considered data set together with the limitation of ‘pure’ diseases (e.g. patients with mixed syndromes where BPH nodules go hand in hand with cancer are excluded by the analysis) are strong case against the generalization of our findings. A particularly hard constraint is the limitation to pure diseases given the high prevalence of mixed syndromes in nature. Nevertheless, in order to validate the method, at first, we preferred to rely upon a more near case. Further developments of metabolomic research in the future will need to address this very important point (59).

Acknowledgements

While this paper was in the final stage of preparation, our co-author Prof. Filippo Conti passed away on 30 May 2009. He was our principal source of inspiration in the search for a metabolism-based holistic perspective in medical diagnosis. His death was both a great sorrow and a potent drive for all of us for pursuing his dream of an integrated scientific culture offering a concrete hope of real scientific advancement.

REFERENCES


METABOLIC ANALYSIS OF PROSTATIC DISEASES


