Chemometrics and Intelligent Laboratory Systems 140 (2015) 43–48

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemolab

Prostate Diffusion Weighted-Magnetic Resonance Image analysis using Multivariate Curve Resolution methods

E. Aguado-Sarrió [a], J.M. Prats-Montalbán [a,*], R. Sanz-Requena [b,c], L. Martí-Bonmatí [c,d], A. Alberich-Bayarri [c], A. Ferrer [a]

[a] Multivariate Statistical Engineering Group, Universitat Politècnica de València, Valencia, Spain
[b] Biomedical Engineering, Hospital Quirón Valencia, Valencia, Spain
[c] Biomedical Imaging Research Group (GIBI230), Hospital Universitario y Politécnico La Fe, Valencia, Spain
[d] Radiology Department, Hospital Quirón Valencia, Valencia, Spain

ARTICLE INFO

Article history:
Received 27 May 2014
Received in revised form 3 November 2014
Accepted 6 November 2014
Available online 15 November 2014

Keywords:
MCR
Multivariate Image Analysis
Diffusion
Magnetic resonance
Prostate
Tumor

ABSTRACT

Multivariate Curve Resolution (MCR) has been applied on prostate Diffusion Weighted-Magnetic Resonance Images (DW-MRI). Different physiological-based modeling approaches of the diffusion process have been submitted to validation by sequentially incorporating prior knowledge on the MCR constraints. Results validate the biexponential diffusion modeling approach and show the capability of the MCR models to find, characterize and locate the behaviors related to the presence of an early prostate tumor.

1. Introduction

Two of the main indicators of a tumor process are the neovascularization and the increase in cellular density. When a group of growing cells presents abnormally high demands of oxygen and nutrients, the tissue responds by creating new vessels (angiogenesis) or developing existing ones (neovascularization). On the other hand, the biological process associated with higher cellular densities that leads to cell agglomeration in the tissue is called cellularization. The combination of both processes is what usually determines the presence of an early tumor as first steps in oncogenesis. One way to approach this combination is by studying the tissue local diffusion process [1], which is a physical process that occurs due to the thermal agitation of the water molecules inside the human body. These translational displacements depend, among other factors, on the tissue structure according to the cellular organization. When the tissue is highly cellularized, the molecules have more restrictions to movement due to a decreased interstitial space and higher cell membrane interfases. However, when the tissue is highly vascularized, molecules are in a non-restricted high velocity environment within the vessels, and the spatial movements are random with less restrictions in all spatial directions.

The diffusion process can be evaluated with a Diffusion-Weighted Magnetic Resonance Imaging (DW-MRI). This non-invasive technique provides high resolution images that are sensitive to water molecules movement inside the tissues. Depending on the configuration of the MR equipment and based on the duration and the amplitude of the applied magnetic field gradient, image acquisition is associated to a parameter known as b-value [2]. The signal of the image decreases with the increase in the b-value acquired. This attenuation depends on the characteristics of the tissue, being stronger if the tissue is vascularized and much more moderate if it is highly cellular. The range of different signal attenuations between these two types of tissue at the same b-value is the basics to study the different behaviors in the diffusion process.

The DW-MRI acquisition sequence is performed along the volume of the studied organ. Usually, images are acquired at spatial planes corresponding to different slices of the human body (the number of slices depends on the studied organ). For each slice, images are taken with different b-values, obtaining a 3D data structure. This way, a signal spectrum **s** is extracted from each pixel of the image, associated to the different b-values. The number of b-values varies among clinical studies, reaching up to 10 values for the clinical setting [3]. In our study, 6 b-values were used for prostate imaging based on previous experience.

* Corresponding author at: Departamento de Estadística e IO Aplicadas y Calidad. Universidad Politécnica de Valencia. Cno. De Vera s/n, Edificio 7A, 46022 Valencia, Spain. Tel.: +34 96 387 70 07x74949; fax: +34 96 387 74 99.
E-mail address: jopramon@eio.upv.es (J.M. Prats-Montalbán).

http://dx.doi.org/10.1016/j.chemolab.2014.11.002
0169-7439/

In order to model the signal decay of the diffusion process, spectra can be fitted with different expressions. The most widely used model in clinical routine is the monoexponential diffusion model [2] with the apparent diffusion coefficient (ADC) as its parameter:

$$\frac{s}{s_0} = e^{-b(ADC)} \quad (1)$$

where s_0 is the initial value of the signal when the b-value equals zero. The ADC values express the average distance that the water molecules cover within a voxel at a certain time. It is related with the cell density, the permeability of the membranes and the tortuosity of the intercellular interstitial space. It is called "apparent" because it reflects several different mechanisms, as it is a combination of the two phenomena expressed before: (i) the movement associated to the water molecules (Brownian movement), known as slow diffusion (cellular tissue), and (ii) the intravascular movement in the microcapillaries vessels, known as fast diffusion or perfusion (vascularized tissue). Fast diffusion may produce an overestimation of the real diffusion values if not properly considered.

The monoexponential model does not take into account the different mechanisms of the diffusion process. Currently, one way of dealing with this complexity is by using a biexponential model. This is a more complex model that considers two behaviors, slow and fast diffusion, weighted with a new parameter called vascular fraction (f), which relates to the proportion of vascular tissue in a voxel. This model is also known as intra-voxel incoherent motions (IVIM) [4] due to the two types of movements considered, related to cellularity (slow diffusion) and vascularization (fast diffusion). The expression of the IVIM model is shown below.

$$\frac{s}{s_0} = (1-f)e^{-bD} + fe^{-b(D+D^*)} \quad (2)$$

The spectra are normalized with s_0 as in the monoexponential approach. Three different parameters must be estimated: the diffusion coefficient (D), the pseudo-perfusion coefficient (D^*) and the vascular fraction (f). This way, the normalized signal s/s_0 is modeled as a weighted average of the slow diffusion (water movement inside the cellular tissue, characterized by D) and the fast diffusion (water movement inside the vascular tissue, characterized by $D + D^*$). The slow diffusion behavior is weighted by $(1 - f)$ and the fast diffusion term by f, since the major contribution of this phase (in the order of 10 times higher) is from D^* if compared to D; however this last parameter is not so low to be obviated. Despite the IVIM model is theoretically more appropriate according to physiological criteria, the monoexponential model is, nowadays, the most widely used in medical practice to model the diffusion process.

Besides, the IVIM model is not a classic biexponential model because the two exponential decays are not independent as they are complementary weighted by the vascular fraction, f. Furthermore, the normalization of the spectra causes a distortion, modifying the shape of the original curve (Fig. 1a) and decreasing the signal-to-noise ratio (as can be seen in Fig. 1b). Thus, this standardization reduces the variability range masking the different behaviors present in the spectra. All these concepts shown above and the difficulty in the interpretation of the results provided by these biomarkers (such as D, D^* and f parameters) have limited their applicability in clinical practice.

Furthermore, these biomarkers are obtained from a generally pixel-by-pixel modeling, which do not take advantage of the relation between pixels with the same behavior, increasing the uncertainty in their estimation; and degrading the corresponding imaging biomarkers (images built from each D, D^* and f parameters at each pixel location) used for clinical purposes.

One possible alternative to analyze these diffusion behaviors is by applying multivariate statistical models, so that it is possible to take advantage of the relation between pixels. When dealing with images, the application of these types of models is known as Multivariate Image Analysis (MIA) [5,6]. The main characteristic of MIA is the capability to study the whole set of pixels at the same time by extracting the sources of variation caused by the latent structures present in the images. In this way, MIA can help in providing new non-parametric models that can explain the principal diffusion behaviors extracted from DW-MRI. It may also be useful to check the appropriateness of the different modeling alternatives (e.g. monoexponential or biexponential) proposed in the literature.

The main and most widespread MIA tool is PCA (Principal Component Analysis) [7]. However, two problems arise when PCA is applied on DW-MRI data: (i) no prior information can be included in the model, and (ii) the orthogonality of the principal components is a limitation to model the different diffusion behaviors that are not necessarily orthogonal. In order to overcome these drawbacks, it is possible to use more flexible models, as is the case of Multivariate Curve Resolution (MCR), which has been already applied very recently to dynamic contrast-enhanced MRI data [8].

The goals of this work are: (i) to explore the capability of MCR methods to model the different behaviors associated to the diffusion process in DW-MRI helping specialists to detect and characterize early tumors in the prostate, (ii) to check the adequacy of the different theoretical models commonly applied in clinical practice, by sequentially incorporating constraints in the MCR algorithm using prior knowledge about the diffusion process, (iii) to provide new imaging biomarkers that may complement those commonly used for clinical diagnosis.

2. Materials and methods

The database consists of DW-MRI acquired from 10 patients with proven prostate carcinoma. The images for each patient were taken along 12 slices covering the whole prostate. For each slice, images with a resolution of 192 × 192 pixels were acquired with 6 different b-values (0, 50, 200, 400, 1000 and 2000 s/mm^2) and arranged in a 3D matrix (192 × 192 × 6) (see Fig. 2a). The output for each patient was a series of twelve 3D images. All images were anonymized and transferred to a dedicated workstation for post-processing.

In order to analyze the images by latent-based bilinear multivariate statistical models, the 3D matrix for each slice was unfolded keeping the b-values mode yielding a 2D matrix (36,864 × 6) that contains all the pixels for each slice in rows and the different b-values in columns (see Fig. 2 right). All the slices from the same patient were studied with the same model by stacking the unfolded 2D matrices of each slice one below the other obtaining a data matrix $\mathbf{S}$ (442,368 × 6). This way the fitted behaviors were forced to keep the same internal correlation structure along the whole prostate volume for a particular patient.

In order to focus the study in the prostate gland, local models were built for each of the 10 analyzed cases by removing the pixels that do not pertain to the prostate zone with manual masks provided by the doctors. This way, the interpretation of the results is improved and the computational time is hugely reduced.

As already commented, in the diffusion process, the studied phenomena are those related to slow diffusion, associated to cellularization, and fast diffusion, associated to vascularization. Assuming that the signal spectrum in a pixel j can be expressed as a weighted sum of different decreasing exponential functions modeling the different phenomena of the diffusion process, we propose the following model:

$$s_j = \sum_{i=1}^{I} c_{ij}\left(\alpha_i e^{-\beta_i b}\right); \quad \alpha_i,\ \beta_i,\ c_{ij} \geq 0 \quad (3)$$

where I stands for the number of exponential functions used. In this work, models using 1, 2 and 3 exponential functions are proposed. The triexponential approach is proposed in order to model a possible