Chemometrics and Intelligent Laboratory Systems 235 (2023) 104777

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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

Partial Least Squares - Diffusion Tensor Imaging (PLS-DTI): A novel approach for biomarker imaging in breast cancer

E. Aguado-Sarrió [a,b,*], J.M. Prats-Montalbán [a,b], G. Robles-Lozano [b], J. Camps-Herrero [c], A. Ferrer [a,b]

[a] Kenko Imalytics S.L., Avda. Médico Vicente Torrent 8, 46015, Valencia, Spain
[b] Grupo de Ingeniería Estadística Multivariante (GIEM), Universitat Politécnica de València (UPV), Camino de Vera s/n, Edificio7A, 46022, Valencia, Spain
[c] Grupo Ribera Salud. Avda, Cortes Valencianas, 58, 46015, Valencia, Spain

ARTICLE INFO

Keywords:
Breast cancer
Early tumor detection
MRI
DTI
PLS
Imaging biomarkers

ABSTRACT

Currently, magnetic resonance imaging is the most sensitive imaging technique for detecting cancer processes in early stages. Regarding breast cancer, due to the characteristics of the tissue as it is formed by ducts (tubular structure), anisotropic diffusion should be considered instead of general isotropic Diffusion Weighted Imaging (DWI). Anisotropic diffusion is studied by applying a technique called Diffusion Tensor Imaging (DTI), where the diffusion gradient is applied with several different directions, calculated by Ordinary Least Squares (OLS) in clinical practice. In this paper, we propose a new DTI calculation method based on Partial Least Squares (PLS), which has some advantages over the traditional OLS calculation: i) the PLS model provides valid biomarkers (non-negative eigenvalues) in a larger percentage of pixels, improving the traditional OLS calculation and reducing the effect of noisier images; ii) OLS tensors are calculated pixel-by-pixel, whereas the PLS method calculates only one model taking advantage of the correlation structure between pixels with similar characteristics, obtaining more reliable estimations; iii) PLS performance is quite reliable when lowering the number of directions of the magnetic field, while this is not the case of OLS. PLS keeps providing a good solution even with low functional resolution equipment, reducing costs and acquisition times, which is an important advantage for its widespread use in value-based medicine-oriented clinical practice.

1. Introduction

Breast cancer is one of the most prevalent cancers at present, being the main cause of cancer mortality within the female population. It is estimated that about 2,088,849 new cases of breast cancer will be diagnosed this year worldwide [1,2]. Due to the prevalence of this disease, better diagnostic and treatment techniques need to be developed. The survival rate of this type of cancers is almost 100% if they are detected in very early stages, dropping dramatically to less than 15% when detected in advanced stages [3].

Among the wide variety of existing diagnostic methods, the most widespread currently is the screening of the population through the use of mammography, which is based on the use of X-rays in order to identify alterations that may indicate the presence of cancer. This technique presents several problems: i) the use of ionizing radiation always involves a risk; ii) the patient discomfort associated with the application of the technique; and iii) the early detection of all tumors (including incipient ones) or the correct measure of their size or shape present difficulties, due to the characteristics of the X-ray image, where only 2-dimensional projections of the densities of a 3-dimensional volume can be appreciated.

For this reason, other different imaging techniques such as magnetic resonance imaging (MRI) have been developed, in particular those using functional image acquisition techniques, due to its ability to early detection of angiogenesis (creation of new vascular tissue) and neovascularization (development of the existing one), as well as cell proliferation, main indicators of tumor processes. These techniques play a very important role in early diagnosis and, furthermore, in the evaluation of the tumor response, easing personal precision medicine applications [4]. Several papers in the literature have demonstrated a promising role for breast MRI techniques. Intravenous gadolinium-chelates based contrast agents (GBCA) are used to detect

Abbreviations: Breast Cancer, BC; Magnetic Resonance Imaging, MRI; Diffusion Tensor Imaging, DTI; Partial Least Squares, PLS.

* Corresponding author. Kenko Imalytics S.L., Avda. Médico Vicente Torrent 8, 46015, Valencia, Spain.
E-mail address: eric@kenkoimalytics.es (E. Aguado-Sarrió).

https://doi.org/10.1016/j.chemolab.2023.104777
Received 16 May 2022; Received in revised form 2 February 2023; Accepted 8 February 2023
Available online 14 February 2023

angiogenesis and neovascularization, and form the basis of dynamic-contrast enhanced (DCE)-MRI. A recent trend to utilize gadolinium-free MRI techniques in clinical scenarios where use of GBCA is a challenge, like pregnancy, severe contrast allergies, and lactational status rely on Diffusion Weighted (DW)-MRI technique, providing information on tissue cellularization, linked to the increase in cell density [5].

Diffusion process is related to the net movement of water molecules in tissues, in vivo and at a microscopic level. The tumors have a smaller interstitial space due to their greater cellularity, with traces of necrosis and fibrosis, thus decreasing the mobility of water molecules.

The diffusion process can be evaluated by means of DW-MR images. Depending on the configuration of the MR equipment and the intensity and shape of the gradient of the applied magnetic field, the acquisition of the images is associated with a parameter known as *b*-value [6]. The image signal decreases as this *b* value increases. This attenuation depends on the characteristics of the tissue, being greater if the tissue is vascularized and much more moderate if the tissue is highly cellularized.

The range of different signal attenuations between these two types of tissues at the same *b*-value is the basis to study the different behaviors in the diffusion process. In order to model the signal decay of the diffusion process signals, the curves can be fitted with different models. For those based on the isotropic decay of the signal as a function of the *b*-value (as happens in the prostatic tissue), the most widely used in clinical practice is the monoexponential diffusion model [6]:

$$I = I_0 e^{-bD} \quad \text{(eq. 1)}$$

where I is the intensity measured at a particular pixel after the application of the magnetic gradient with a particular *b*-value, I_0 is the intensity measured on a base image without applying a magnetic gradient ($b = 0$), and D is the Apparent Diffusion Coefficient (ADC) [7,8].

This model can be calculated for each pixel when the magnitude of the diffusion is similar in all directions and corresponds to a decreasing exponential model where D (ADC) expresses 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 two phenomena: (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.

However, in anisotropic tissues such as the breast (breast ducts), a single ADC may be not sufficient to properly characterize the mobility of water molecules in these tissues, since it depends on the direction the diffusion is calculated. Therefore, a measure of diffusion in the three spatial dimensions is required. For this, in order to characterize the magnitude of the diffusion according to the orientation of the applied gradients, it is necessary to use another type of diffusion sequences by modifying the directions in which the gradient of the magnetic field is applied with a certain *b*-value. Thus, the loss of signal will be greater in those directions in which the magnitude of diffusion is higher. The detection of the tumor is performed by calculating within each pixel what is known as diffusion tensor (Diffusion Tensor Imaging, DTI), a 3x3 matrix per pixel which contains information about the diffusion in the three spatial directions, and evaluates the water molecules restriction to movement in each direction of the tridimensional space [9–11]. In other words, it characterizes tissue microstructure and water diffusion directionality by performing in multiple orientations. This fact enables DTI to detect more information about this microstructure rather than DWI, since it can provide not only the average apparent diffusion coefficient (ADC, or mean diffusivity (MD) in the case of DTI), but also three-dimensional anisotropy diffusion parameters such as fractional anisotropy (FA) [12–14]. This way, it appears as a valuable tool to differentiate breast cancer from benign lesions with high sensitivity and specificity [15].

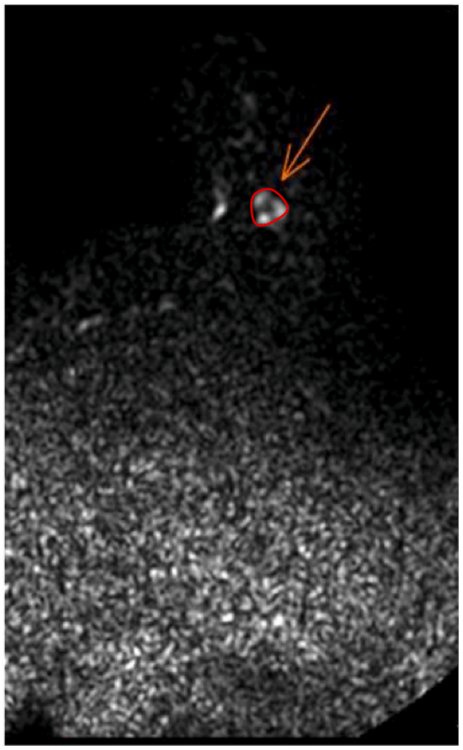

Fig. 1. Example of reference tumor region (ROI) including the identification and segmentation proportionated by the radiologist.

Although some of the first papers come from earlies 2010s [16–19], DTI is not still a high (or completely) developed technique. Even though, in a recent study, DTI has achieved higher accuracies than those reported by DWI-based techniques, such as ADC or the more complex Intra-Voxel Incoherent Motion (IVIM) [20].

DTI requires diffusion measurements in at least six noncollinear directions. In clinical practice, Ordinary Least Squares (OLS) is used for DTI calculation. This leads to matrix invertibility problems that cause instability in the eigenvalues computation, and afterwards in the biomarkers computed from them. In fact, a recent paper indicated that some of the most relevant DTI biomarkers, such as fractional anisotropy (FA), are sensitive to noise, and that *"The relative strength of tissue properties and noise amplification in anisotropy measurement is an important question to guide interpretation of results."* [21]. Other studies claim for the use *"of more advanced DTI techniques, with reduced distortion …"* [22] or for *"advances in DTI techniques that result in higher spatial resolution and reduced distortion may improve the ability to distinguish ductal carcinoma in situ (DCIS) from other lesion types."* [23]. Also, in Ref. [24], it is stated "*DTI measures a decrease in signal intensity and hence suffers from low* S/N …" and that "*… improvements are required in hardware, pulse sequence and software, in order to overcome the technical problems causing artifacts ….*"; or that "*there are also general technical limitations associated with the application of DTI, such as spin-echo echo-planar imaging (EPI)-related artifacts"* [25]. Therefore, there is still room for improving the technical aspects of DTI, related to the limitations of Ordinary Least Squares (OLS).

In order to improve the drawbacks of Ordinary Least Squares (OLS), this paper introduces a new approach for the calculation of the diffusion tensor at each pixel using Partial Least Squares (PLS).

In Section 2, the Materials and Methods used in this paper are presented. Specifically, DTI mathematical model is presented, as well as the adaptation of the PLS model to the data structure built from the MRI

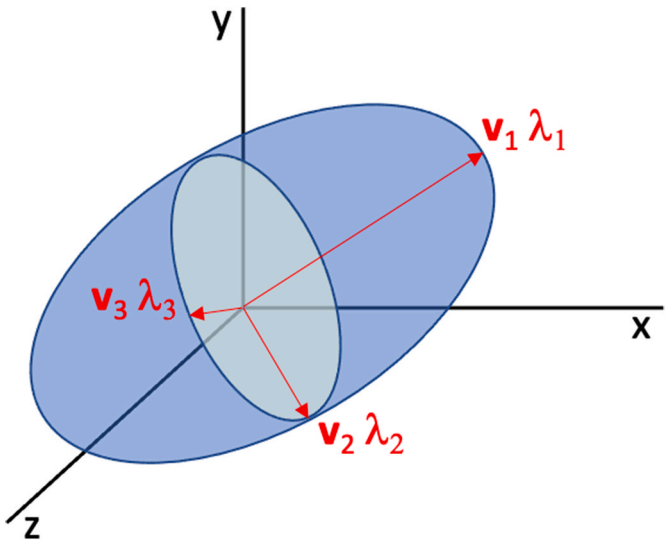

Fig. 2. Diffusion tensor with main axes ($\mathbf{v}_i$ eigenvectors) and eigenvalues λ_i (i = 1,2,3).

sequences.

In Section 3, the results and discussion of the PLS-DTI model are presented and compared to the classical OLS calculation.

Finally, in Section 4, conclusions are commented.

2. Materials and Methods

2.1. Patient database

The database consisted of 30 histologically-confirmed cases of breast tumors. DTI sequences were acquired in all cases, ensuring full breast coverage (81 slices, in-plane resolution of 336 × 336 voxels, each one measuring 1.0833 × 1.0833 × 2.5 mm^3). DTI sequences were acquired at 16 different diffusion gradients (different directions with b = 700 s/mm^2) plus one additional image acquired with b = 0.

Reference tumor regions of interest (ROIs) were identified and segmented by a radiologist with more than 20 years of experience in breast imaging, considering image findings and biopsy location. Fig. 1 shows a ROI example:

All patients gave consent for using their medical images, which were anonymized before post-processing. The local Ethics Committee approved the study protocol.

2.2. Anisotropic monoexponential diffusion model

As already commented, although in certain tissues (e.g. the prostate) movement restrictions can be considered the same in all directions and, consequently, isotropy can be assumed, in other organs (e.g. the breast or the brain) there exist diffusion preferred directions. Therefore, it is not possible to characterize this diffusion by means of a single parameter such as the ADC.

In these situations, it is necessary to define a 3x3 matrix in a three-dimensional environment, known as the diffusion tensor, which enables to model how diffusion occurs in each direction of space. Thus, diffusion behavior at a particular pixel can be modelled as an ellipsoid that is represented by the diffusion tensor **D,** which indicates the magnitude of diffusion in the 3 spatial axes (x, y and z) of the reference frame [26]:

$$\boldsymbol{D} = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{pmatrix} = \mathbf{V} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \mathbf{V}^T \quad \text{(eq. 2)}$$

where matrix **V** contains in columns the $\mathbf{v}_i$ eigenvectors of the diffusion tensor **D**, and λ_i (i = 1,2,3) are their corresponding eigenvalues, obtained after diagonalizing **D**.

Fig. 2 shows the diffusion tensor **D** represented by an ellipsoid in a three-dimensional space. Note that this ellipsoid has three main axes, defined by the eigenvectors $\mathbf{v}_a$, and their associated eigenvalues λ_i are proportional to the squares of the ellipsoid hemiaxis lengths.

Due to the physical constraints of diffusion, all three eigenvalues λ_1, λ_2, λ_3 must be positive so the diffusion tensor is positive-definite. This necessary constraint will be applied to validate the diffusion tensors calculated later at each pixel of the image.

Generalizing the exponential model of isotropic tissues (eq. (1)) to the anisotropic case, the mathematical exponential model that relates the intensity measured with the diffusion tensor as a function of the applied gradient in a particular direction is as follows:

$$I = I_0 e^{-\boldsymbol{B}:\boldsymbol{D}} \quad \text{(eq. 3)}$$

In this expression **D** corresponds to the diffusion tensor at a given pixel, I is the intensity measured at that pixel when applying the diffusion gradient, and I_0 is the intensity measured at that pixel without applying the gradient (b = 0). Finally, **B** is a matrix that contains the information of the magnitude and direction of the applied gradient, which in turn can be decomposed into the magnitude of the gradient b, and a matrix **G** with the information on the direction of the gradient vector **g**, formed by three components (g_x, g_y, g_z) in the 3 spatial axes (x, y and z) of the reference frame:

$$\boldsymbol{B} = b\boldsymbol{G} = b\mathbf{g}\mathbf{g}^T = b \begin{pmatrix} g_x{}^2 & g_x g_y & g_x g_z \\ g_x g_y & g_y{}^2 & g_y g_z \\ g_x g_z & g_y g_z & g_z{}^2 \end{pmatrix} \quad \text{(eq. 4)}$$

This way, the tensor product **B:D** is finally defined as:

$$\boldsymbol{B} : \boldsymbol{D} = b\boldsymbol{G}$$
$$: \boldsymbol{D} = b\left(g_x{}^2 D_{xx} + g_y{}^2 D_{yy} + g_z{}^2 D_{zz} + +2\, g_x g_y D_{xy} + 2\, g_x g_z D_{xz} + 2 g_y g_z D_{yz}\right) \quad \text{(eq. 5)}$$

2.3. Pixel-by-pixel diffusion tensor calculation by ordinary least squares (OLS)

In order to calculate the diffusion tensor from the information contained in the images, it is first necessary to linearize the model (eq. (3)) to simplify its estimation:

$$\frac{\ln \frac{I}{I_0}}{b} = -\mathbf{G} : \mathbf{D} \quad \text{(eq. 6)}$$

This equation corresponds to a single gradient direction. In order to build a system of equations that yields the diffusion tensor **D** as a solution, a minimum of 7 images need to be acquired (one for b = 0, and six in different directions of the gradient vector **g**). In the limit case of having only 6 directions, the solution is unique. However, due to the possible noise that the images may contain, the calculation with few directions is not robust, so it is recommended to have more gradient directions. In general, a total of $M > 6$ gradient directions are taken, thus building a system of M equations with six unknown values (the elements of the diffusion tensor **D**). The structure of the equation system can be consulted in Fig. S1 of the supplementary material.

y and **a** refer to the information of the signal intensity and the tensor elements respectively (eqs. (7) and (8)):

$$\mathbf{y} = \left(\frac{\ln \frac{I_1}{I_0}}{b}, \frac{\ln \frac{I_2}{I_0}}{b}, \ldots, \frac{\ln \frac{I_M}{I_0}}{b} \right)^T \quad \text{(eq. 7)}$$

$$\mathbf{a} = \left(D_{xx}, D_{yy}, D_{zz}, D_{xy}, D_{xz}, D_{yz}\right)^T \quad \text{(eq. 8)}$$

and **H** gathers, in rows, information of matrix **G** (eq. (4)) after applying the tensor product (eq. (5)) for *the M* gradient directions (eq. (9)):

$$\mathbf{H} = \begin{pmatrix} g_{1x}{}^2 & g_{1y}{}^2 & g_{1z}{}^2 & 2g_{1x}g_{1y} & 2g_{1x}g_{1z} & 2g_{1y}g_{1z} \\ g_{2x}{}^2 & g_{2y}{}^2 & g_{2z}{}^2 & 2g_{2x}g_{2y} & 2g_{2x}g_{2z} & 2g_{2y}g_{2z} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ g_{Mx}{}^2 & g_{My}{}^2 & g_{Mz}{}^2 & 2g_{Mx}g_{My} & 2g_{Mx}g_{Mz} & 2g_{My}g_{Mz} \end{pmatrix} \quad \text{(eq. 9)}$$

The solution of this equation system results in the 6 values of the diffusion tensor for a given pixel.

By solving the previous equation system for each one of the *N* pixels of an image, it is possible to combine the information of the different vectors **y** and **a** of each one of the *N* pixels in the matrices **Y** and **A**, respectively. A scheme of the OLS calculation for DTI can be consulted in Fig. S2 of the supplementary material [27]:where **Y** (eq. (10)) contains the linearized intensities of the *N* pixels of the *M* different gradients directions, **H** contains the information of the gradients applied to the different directions (eq. (9)), **A** contains in columns the components of the diffusion tensors for each one of the *N* pixels (eq. (11)), and **E** is a residual matrix. This matrix expression integrates *N* multiple regression models (one for each pixel) where **Y** contains the *N* response variables (one for each pixel), **H** the regression variables and **A** the regression coefficients.

$$\mathbf{Y} = \begin{pmatrix} \frac{ln\frac{I_{11}}{I_{01}}}{b} & \cdots & \frac{ln\frac{I_{1N}}{I_{0N}}}{b} \\ \vdots & \ddots & \vdots \\ \frac{ln\frac{I_{M1}}{I_{01}}}{b} & \cdots & \frac{ln\frac{I_{MN}}{I_{0N}}}{b} \end{pmatrix} \quad \text{(eq. 10)}$$

$$\mathbf{A} = \begin{pmatrix} D_{xx1} & D_{yy1} & D_{zz1} & D_{xy1} & D_{xz1} & D_{yz1} \\ & & \vdots & & & \\ D_{xxN} & D_{yyN} & D_{zzN} & D_{xyN} & D_{xzN} & D_{yzN} \end{pmatrix}^T \quad \text{(eq. 11)}$$

This model is fitted pixel by pixel by using ordinary least squares (OLS) to obtain the values D_{ij} ($i = x, y, z$), ($j = x, y, z$) of the diffusion tensor for each pixel (columns of matrix **A**) of the image (eq. (12)):

$$\mathbf{A} = (\mathbf{H}'\mathbf{H})^{-1}\mathbf{H}'\mathbf{Y} \quad \text{(eq. 12)}$$

However, OLS suffers from several limitations. One is that, when making an estimation by OLS, a large number gradient directions are required to obtain a robust estimate. If not enough images are taken or if they contain a high level of noise, incorrect estimations of the diffusion tensor are obtained (e.g. tensors having negative eigenvalues and thus not being positive-definite). This is commented in Ref. [28], when stating that *"a limitation of the use of the λ_1–λ_3 parameter, or of other absolute indices, such as λ_1/λ_3 or of λ_1-($\lambda_2 + \lambda_3$)/2, is the high sensitivity of λ_3 to noise, and to sorting bias of the principal diffusivities."* This problem may be alleviated by using more powerful equipment, using filters to denoise images, or increasing the number of gradients applied, but this is not usually possible due to time and budget constraints. Another important drawback is that the tensor is calculated pixel-by-pixel. Therefore, since the correlation structure between pixels with similar characteristics is not considered, less reliable estimations are obtained.

2.4. Diffusion tensor calculation by PLS

In order to overcome the limitations of OLS in the diffusion tensor calculation, in this work we propose the calculation of the diffusion tensor using Partial Least Squares (PLS) regression [29–32]. PLS is a well-known projection to latent structure model that explains the relationship between two sets of variables **X** and **Y**, as well as the variability in both **X** and **Y**, by maximizing the covariance between their inner latent variables. Instead of fitting one OLS model per pixel, PLS gets benefit of the internal correlation structure between pixels, fitting just one model with all pixels of the image, allowing a more reliable calculation in the presence of noise.

For the PLS algorithm, the same array structure is used as in the OLS model. However, since the solutions of the models are not the same. The model structure is presented with a different scheme in Fig. S3 of the supplementary material.where the only difference is the final matrix **A***(plus residual matrix **F**) that gathers the diffusion tensor elements, which can be obtained in one single step (avoiding the pixel-by-pixel calculation) for the *N* pixels in the following way:

$$\mathbf{A}^* = \mathbf{W}^*\mathbf{Q}^T = \mathbf{W}\left(\mathbf{P}^T\mathbf{W}\right)^{-1}\mathbf{Q}^T \quad \text{(eq. 13)}$$

where **A*** is the $\mathbf{B_{PLS}}$ matrix, and **W**, **P** and **Q** are the X-weights, X-loadings and Y-loadings, respectively (for more insight into PLS model, the reader is referred to, e.g. Ref. [29]).

2.5. Image date management

Before starting the tensor calculation, the images should be organized as a data structure that can be correctly read by the OLS/PLS algorithm. In this work, the images are acquired in DICOM format and the relevant information for the calculation is registered in the DICOM headers (such as slice location, gradient and *b*-value). First, the information of the gradient directions is extracted. For each slice, the information of *M* gradient directions is obtained with their respective components (g_x, g_y, g_z) creating a 2D matrix **G***. This matrix remains invariant when going through different slices (the acquisition is taken with the same characteristics for each patient). The structure of **G*** can be consulted in Fig. S4 of the supplementary material.

The information of **G*** is used to build **H** (eq. (9)) using the expression of the tensor product (eq. (5)).

The *b*-value information is also needed for calculation. As for the gradient directions, the *b*-value is extracted from the DICOM headers and it does not change when the slice or even the direction *m* is changed. Thus, in this paper, the *b*-value can only take 2 different values: 0 and 700 s/mm^2 700 is the scalar value used in eqs. (7) and (10) for the calculation of the diffusion tensor.

Finally, the information of the images is organized as follows:

For each of the *C* slices of a sequence, there are *M*+1 images of *I* x *J* pixels, associated to each of the *M* directions plus the image of $b = 0$. The information of each slice *c* is organized in a 3-way structure of *I* x *J* x (*M*+1) dimensions. A scheme is shown in Fig. S5 of the supplementary material.

Each one of these 3-way structures is unfolded pixelwise, obtaining a 2-way **Y*** matrix of *M* x (*I* x *J*) dimensions. A scheme is shown in Fig. S6 of the supplementary material.

After the 2-D unfolding, a basic mask is applied following the criteria of selecting the pixels that have higher intensity than 5 in the $b = 0$ image (b_0 is the image with higher signal intensity), removing the pixels with low/close to 0 intensity (considered noise). After this automatic pixel selection, the column dimension of **Y*** is reduced from *I* x *J* to *K* (**Y****).

In this new **Y**** matrix, each row starting from the 2nd to the last one is divided by the first row (the signal intensity when $b = 0$). Then, the logarithm of the division is calculated, afterwards divided by the scalar *b*-value associated to direction *m*, as shown in eq. (6) (the row associated to the direction $b = 0$ is not included in each final matrix **Y****). When every **Y**** matrix for each slice *c* is calculated, they can be stacked column-wise obtaining only one matrix **Y** that contains all the pixels from the sequence, as shown in eq. (10).

3. Results and discussion

In this section, the PLS method is compared with the traditional OLS

Table 1
Diffusion tensor imaging biomarkers based on clinical experience and bibliography. Source: Kingsley 2006a [12].

Parameter	Expression
Longitudinal diffusion (LD)	λ_1
Second longitudinal diffusion	λ_2
Third longitudinal diffusion	λ_3
Apparent diffusion coefficient (ADC)	$(\lambda_1+\lambda_2+\lambda_3)/3$
Maximum diffusion differences (dmax)	λ_1-λ_3
Maximum diffusion ratio (Rmax)	λ_3/λ_1

Table 2
Discrepancies between the tensor eigenvalues of the OLS model and the PLS models (from 1 to 6 latent variables) using eq. (14). The comparison is made with different number of latent variables. The table shows the results of the average for the 30 calculated cases.

		# Latent Variables					
		1	2	3	4	5	6 (OLS)
Discrepancy (%)	λ_1	23.27	19.09	13.82	10.25	3.51	0.00
	λ_2	5.36	5.31	5.62	5.23	5.55	0.00
	λ_3	41.51	35.47	26.16	20.04	7.30	0.00

calculation. First, some considerations have to be taken into account: i) to be considered valid, the diffusion tensor must be definite positive, so all its eigenvalues λ_i are positive; ii) the number of components in PLS should be$\leq$6 (6 is the maximum number of different elements in the tensor since it is symmetric); and iii) taking 6 components (maximum number of components) in PLS is equivalent to OLS method.

In order to compare the results of OLS and PLS models, 30 cases of breast cancer where the tumor is confirmed will be calculated. Once the tensor is calculated and diagonalized resulting in the 3 main eigenvalues (λ_1, λ_2 and λ_3) for each pixel of the image, the next step is to calculate the main imaging biomarkers that can be obtained from these eigenvalues. The most widely used biomarkers are shown in Table 1:

The longitudinal diffusion (LD) represents the highest possible diffusion in that pixel, known as the diffusion coefficient (λ_1) of the preferential diffusion direction (i.e., the longitudinal axis of the ducts, $\mathbf{v}_1$). When LD is high, the water molecules can move freely inside the tissue, but when LD is low, this movement is restricted and the duct is suspicious of being blocked by a tumoral presence. The second and third longitudinal diffusions represents the diffusion coefficient as well, but measured in the orthogonal directions ($\mathbf{v}_2$ and $\mathbf{v}_3$). In a healthy region, λ_1 is always higher than λ_2 and λ_3, geometrically represented as an ellipsoid (Fig. 2). However, in a tumoral region, λ_1 is similar to λ_2 and λ_3. In this case, the tensor is geometrically interpreted as a sphere.

The ADC is the mean value of the 3 principal diffusion coefficients while dmax and Rmax represents, respectively, the differences between the maximum and the minimum diffusion in terms of differences or ratio. When dmax is high, the pixel is interpreted as a sane region (ellipsoid) but when it is low (close to 0) the pixel is interpreted as tumoral (sphere); this biomarker cannot take negative values. Rmax takes values between 0 and 1, when it is close to 1 represents tumoral tissue (sphere) and when it is close to 0 represents sane tissue (ellipsoid).

3.1. Differences in the value of the tensor eigenvalues

In an attempt to measure the discrepancy between the eigenvalues (λ_1, λ_2, λ_3) obtained from the models, the following discrepancy index is defined.

$$Discrepancy_i(\%) = \frac{\sum_{n=1}^{n=npix} (\lambda_{iOLS} - \lambda_{iPLS})^2}{\sum_{n=1}^{n=npix} \lambda_{iOLS}{}^2} \cdot 100 i = 1,2,3 \qquad \text{(eq. 14)}$$

Table 2 shows the discrepancies between the tensor eigenvalues of

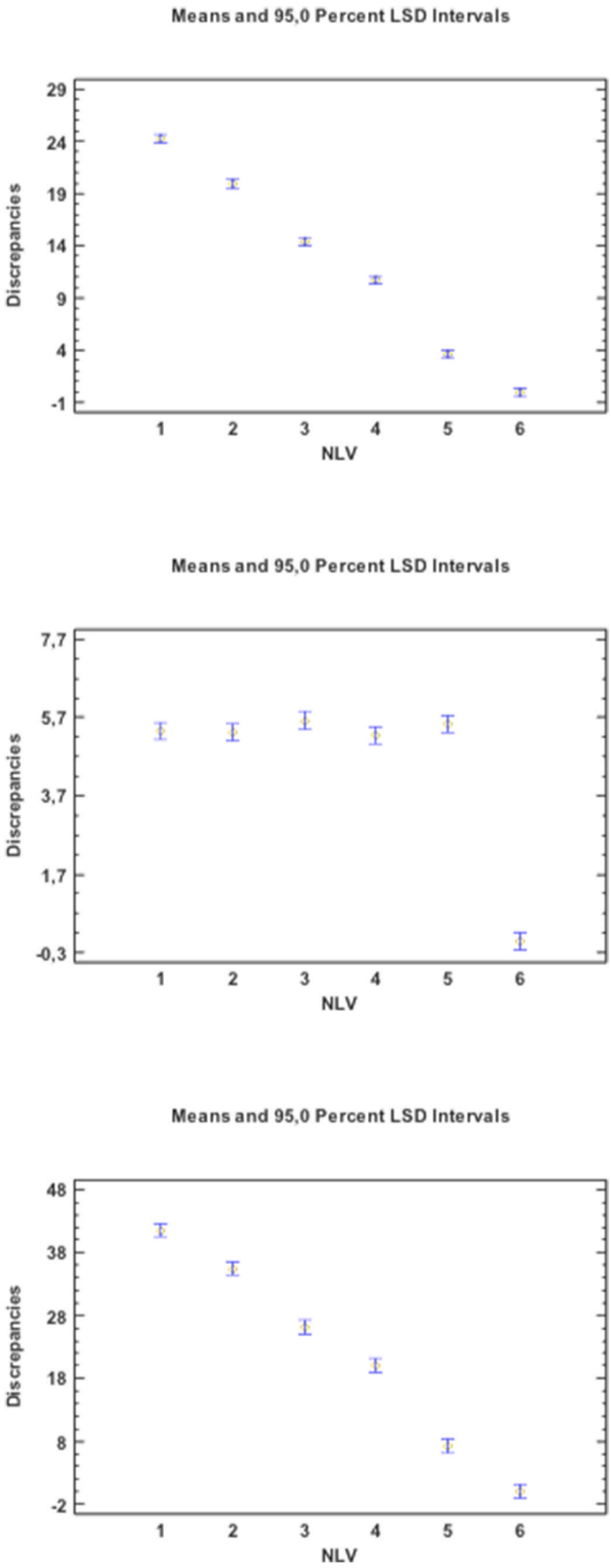

Fig. 3. 95% LSD interval plots for mean discrepancies for λ_1 (top), λ_2 (middle) and λ_3 (bottom) vs number of latent variables (NLV).

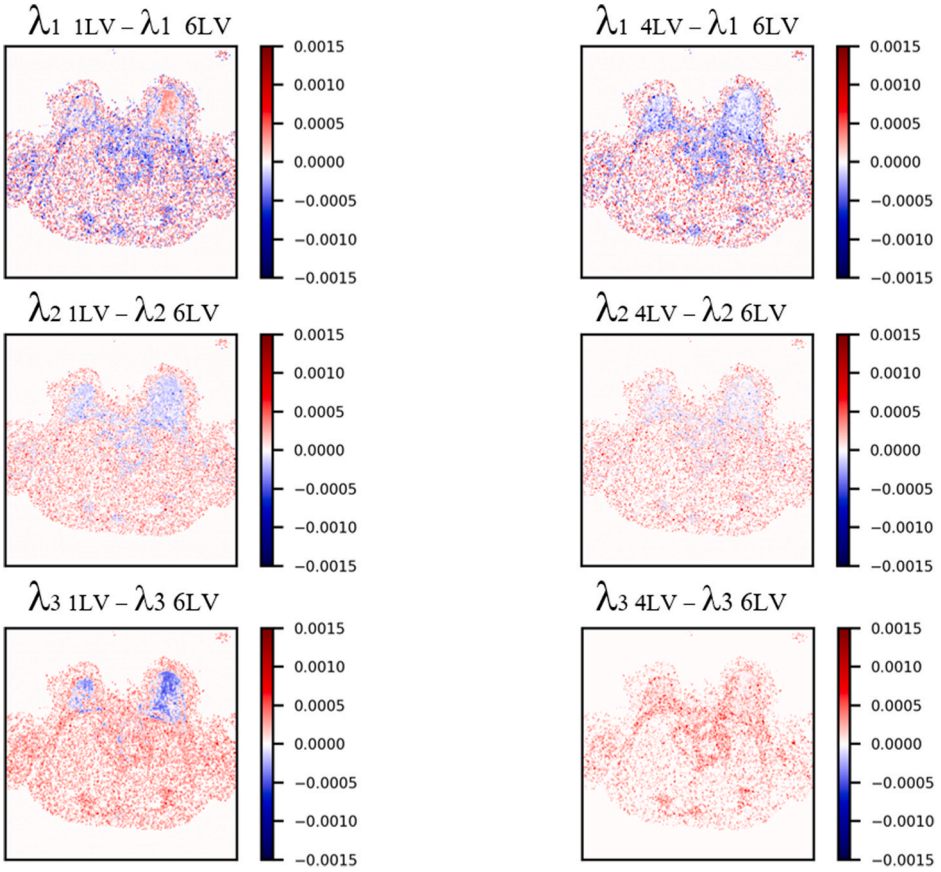

Fig. 4. Differences in parametric map between PLS models: from left to right (1LV vs 6LV; 4LV vs 6LV) and top to bottom (λ_1, λ_2, λ_3): negative values (blue) indicate lower values of eigenvalues in PLS wrt OLS model, while positive values (red) indicate higher values in PLS wrt OLS model. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

the OLS model and the PLS model using eq. (14). The comparison is made with different number of latent variables. The table shows the results of the mean for the 30 calculated cases:

As observed in Table 2, the differences between both models decrease as the number of latent variables increase. In the limit case of 6 components, both models are identical as OLS is a particular case of PLS when extracting the maximum number of latent variables (in this case, 6).

The ANOVA carried out on the database, using the cases of breast tumor as a blocking factor and the number of latent variables (NLV) as the factor of interest, shows statistically significant differences between the mean discrepancies values for all three eigenvalues. The corresponding 95% least significance difference (LSD) interval plots of the mean discrepancies vs NLV are shown in Fig. 3.

For λ_1 and λ_3 there is a clear linear descendant evolution of the average discrepancy as the number of latent variables increases. However, for λ_2, the higher statistical difference appears between the PLS models and the equivalent OLS (classical) model (i.e. PLS with 6 latent variables). Regarding to the differences between eigenvalues, the discrepancies in λ_3 are higher than the ones found in λ_1. Besides, the differences in λ_2 are not as high as the other eigenvalues. In order to visually study the discrepancies, Fig. 4 shows the parametric maps of the differences between the 6LV model and the 1 and 4 LV models for λ_1, λ_2 and λ_3.

As can be seen in Fig. 4, the differences are higher when comparing 1LV vs 6LV (OLS). But the solution comes closer when the number of LV is increased (4LV vs 6LV). When observing λ_1 and λ_3 maps, the 1LV model tends to underestimate (blue color) the values of λ_3 in the breast region at the expense of overestimating (red color) the values of λ_1 at the same region. In general (out of the breast region), λ_1 is underestimated (in blue) and λ_3 overestimated (in red) in the 1LV model. This makes the parametric maps of the biomarkers based on eigenvalues differences (described in Table 1) to be affected when selecting a different number of LV. Besides, as also shown in Table 2 and Fig. 3, λ_2 is the less affected eigenvalue showing lower map differences.

Table 3
Percentage of valid pixels considering different PLS models (from 1 to 6 latent variables). Mean value and standard deviation (std) calculated for 30 cases of breast cancer.

	Valid pixels (%)					
# Latent Variables	1	2	3	4	5	6 (OLS)
Mean	64.59	54.51	49.17	45.06	35.78	28.83
std.	1.88	2.19	2.36	2.43	2.52	2.64

3.2. Differences in the number of valid pixels

As described before, a valid pixel can be defined as a tensor with a physical meaning. In this case, this implies that the tensor for one pixel is positive-definite, i.e., all its eigenvalues are positive. Since they are ordered from higher to lower, it is enough to check if λ_3 is positive.

The applied mask is defined in section 2 as the body region of the image after removing the background where the intensity is lower than 5 (noise). Table 3 shows the mean and the standard deviation of the percentage of valid pixels for the 30 calculated cases:

As can be observed, the percentage of valid pixels decreases as the number of latent variables increases. For 6 components (OLS model), the number of valid pixels gets its minimum. Therefore, the PLS method has the flexibility to adapt the model in order to obtain a larger number of

Table 4
Percentage of valid pixels in ROI's considering different PLS models (from 1 to 6 latent variables). Mean value calculated for 30 cases of breast cancer.

	# Latent Variables					
Valid pixels (%)(ROI pixels)	1	2	3	4	5	6 (OLS)
	96.93	95.98	94.66	94.10	92.44	91.40

valid pixels (i.e., pixels with positive-definite diffusion tensors).

This study is repeated only for the pixels contained in the ROI's defined in section 2 (cases where the tumor has been confirmed, identified and segmented by an experienced radiologist). The results are presented in Table 4:

As can be observed, the percentage of valid pixels increases in the breast region (ROI). This happens because lesion pixels, generally, present higher intensity in the acquisition, which results in a higher probability of obtaining a valid tensor. However, the conclusions obtained from Table 3 can be also extended to the breast region: the number of valid pixels increases as the number of components decreases. Although the differences are not as big as in Table 3, it is important to highlight that an invalid pixel means that there is no biomarker calculation in that pixel, so there is no information on the presence/absence of tumor. Selecting the number of components for the calculation will depend on the signal-to-noise ratio of the images. In images with a high noise level, it will be interesting to take a low number of components in order to avoid loss of information.

In order to make a proper statistical comparison of these values, an ANOVA has been performed in section 3.4, based not only on the number of latent variables, but also on the number of directions used for building the DTI models.

3.3. Biomarkers parametric maps

Fig. 5 shows the eigenvalues parametric maps for the breast region of one of the 30 cases. There is one characteristic regarding the map representation that has to be taken into consideration. The colormap is inverted and the background is established black. This is done because high diffusion represents less cellularization (free environment) and thus, lower tumor risk (represented in blue in the colormap). Low diffusion represents high cellularization, instead, and therefore, high tumor risk (in red). The black pixels include the out-of-the-mask pixels and the ones with non-valid tensors.

As can be seen in Fig. 5, there are no major visual differences for the same eigenvalues from different LV models in the breast region, although there are differences in the values of the eigenvalues as shown in Table 2. The main advantage of PLS with lower number of latent variables in map visualization is the higher number of pixels with valid tensors (Fig. 5). Regarding to map interpretation, in λ_3 map it is easier to distinguish the tumor region (marked in black line) than in λ_1 map.

The next step is to check how the biomarkers calculated from the eigenvalues are affected by the number of LV. In order to ease the representation, only dmax and 1-Rmax are considered. 1-Rmax is represented instead of Rmax in order to keep the coherence between the maps: pixels with similar values for λ_1 and λ_3 (high cancer risk) will yield small values both in dmax and 1-Rmax and will be colored in red (inverted colormap). On the other hand, pixels with different values for λ_1 and λ_3 (low cancer risk) will yield high values both in dmax and 1-Rmax and will be colored in blue. The point of defining Rmax $= \lambda_3/\lambda_1$ instead of λ_1/λ_3 is to avoid extreme values or even indeterminations provided by the ratio (since λ_3 will be always lower than λ_1).

A visual comparison between models can be consulted in Fig. S7 of the supplementary material. Additionally, some examples of the biomarkers obtained from a 4 LV PLS model can be seen in Fig. S8 of the supplementary material.

dmax and 1-Rmax (Fig. S8) resulted useful to identify tumors in combination with individual λ_i (especially λ_3) (Fig. 5). This can be explained as a consequence of the physical phenomena of cellularization and studying the geometric representation of the diffusion ellipsoid. As already commented, when the tissue is sane, there is one preferential diffusion direction (the ductal axis) and the differences between λ_1 and λ_3 are big, yielding high values for dmax and 1-Rmax (in blue). But, when the duct is collapsed (cellularized), there is no preferential

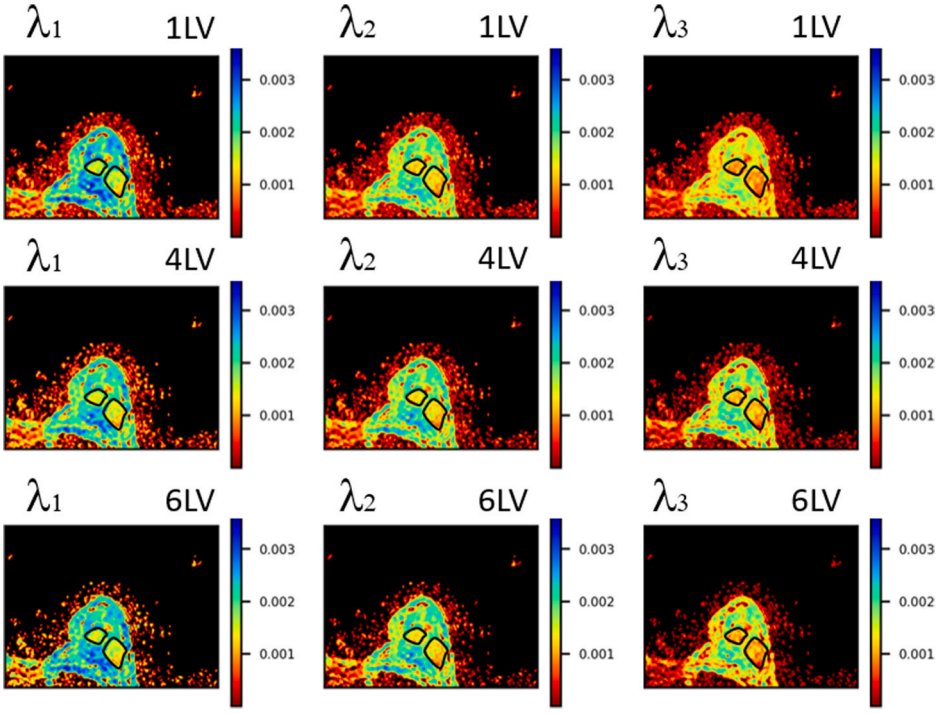

Fig. 5. Eigenvalues parametric map comparison between PLS models for one specific slice of the same patient: from left to right (λ_1, λ_2, λ_3), and top to bottom (1LV, 4LV, 6LV). Black lines indicate the tumor regions confirmed, identified and segmented by an experienced radiologist.

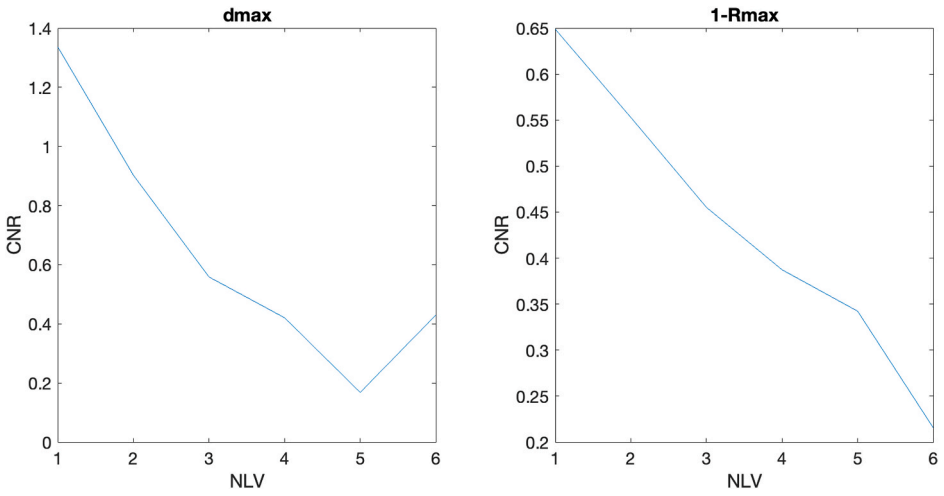

Fig. 6. Evolution of CNR with NLV for dmax (left) and 1-Rmax (right).

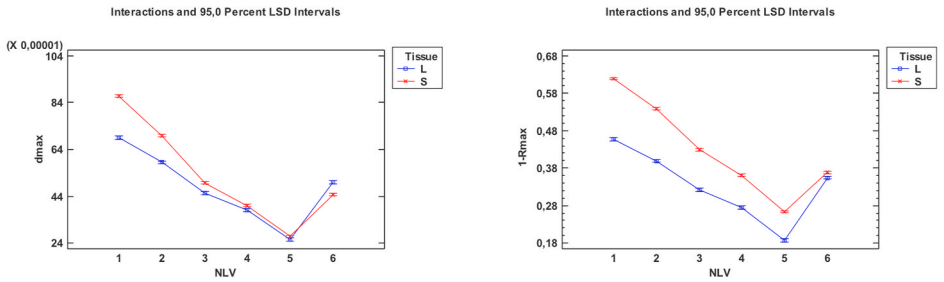

Fig. 7. 95% LSD intervals plot of average dmax (left) and average 1-Rmax (right) for the NLV × tissue interaction.

diffusion direction and the diffusion ellipsoid becomes kind of a sphere, thus the differences between λ_1 and λ_3 are small yielding low values for dmax and 1-Rmax (in red).

In order to assess for the tumor conspicuity of the images, a contrast to noise ratio (CNR) analysis has been performed [33], which in turn tries to perform a comparison of the statistical mean differences between the lesion tissue and sane tissue. Fig. 6 shows the corresponding results, both for dmax and 1-Rmax. An alternative ANOVA analysis is also provided, which gives richer information, since it assesses for these statistical differences between these mean values.

In both cases, there is a clear descendant trend of the CNR, which means that, as the number of latent variables increases, the differences in the means of the lesion and sane tissues decrease, so in principle also the conspicuity. Therefore, again, PLS shows up as a better option when compared to OLS (equivalent to PLS with 6 LV). One could maybe argue that, for dmax, CNR increases a little bit after 5 LV, but just reaching the same value as the CNR corresponding to a 4 NLV PLS model.

In order to study the statistical significance of these differences and the potential factors affecting them, an ANOVA is carried out, using the cases as a blocking factor, and the number of latent variables (NLV), number of directions, and type of lesion (L, lesion vs S, sane) as factors under study.

For dmax, the ANOVA table (not shown) shows that all effects and interactions are statistically significant (P-value$<$0.05).

The effect of the interaction between NLV and the Type of tissue, which means that the evolution of the means of dmax with NLV depends on the type of tissue (different slopes or trends) is shown in Fig. 7. Or in a different way, that the differences between the dmax means of both tissues is different as the number of latent variables evolves, which is in essence what CNR tries to measure.

When performing the ANOVA on 1-Rmax, similar conclusions can be extended to those from CNR, but also in this case with a higher level of information related to the blocking factor (there is a quite obvious difference between the cases), as well as to the main effects and interactions, always with statistical support, based on the p-values$<$0.05 and the LSD intervals (ANOVA table not shown).

From Fig. 7, we can also see the clear descendant evolution of the differences between the means of 1-Rmax already shown in Fig. 6 from the CNR analysis, when analyzing the interaction between NLV and Type of tissue. But in this case, we can appreciate that 1-Rmax mean values increase for 6 latent variables, although the difference between their means are still the smallest, yielding a very low CNR (as shown in Fig. 6).

3.4. *Differences when reducing the number of gradient directions*

This last study consists in calculating the diffusion tensor with less gradient directions than the original acquisition (i.e., simulating a lower functional resolution equipment) and how it affects the percentage of valid pixels. Only one consideration has to been taken into account: at least 6 directions are needed in order to estimate the 6 values of the diffusion tensor; otherwise, less than 6 directions makes the system of equations unsolvable.

Unlike the OLS method (pixel-by-pixel calculation), PLS takes advantage of the information of the covariance with the other pixels in the calculation of the tensor. This way, it is expected that the PLS results will be less affected by the reduction of the number of directions than the OLS method.

To prove this, the tensor is calculated with a restricted number of directions: 10 and 6. The directions were selected from the original acquisition trying to assure that distance between selected gradients is maximum. With this subset of directions, the tensor calculation is performed in the same way as the previous section. For both all-pixels and ROIs pixels datasets, the ANOVAs performed (results not shown) showed

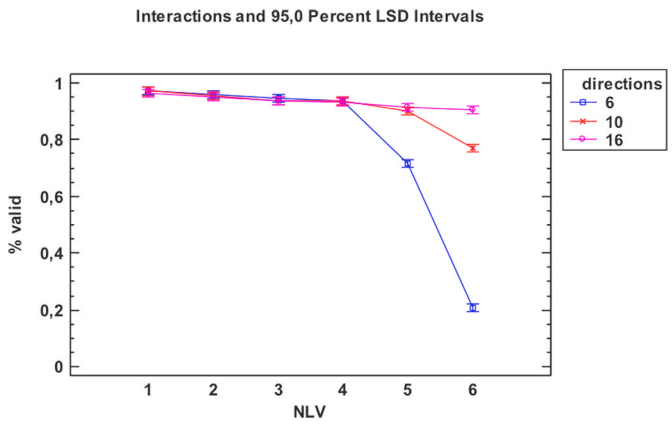

Fig. 8. 95% LSD intervals plot of average percentage of valid pixels in the ROIs for the NLV x number of directions used for calculating the DTI interaction.

statistically significant effects of both factors under study (NLV and number of directions) on the average percentage of valid pixels.

For the pixels contained in the ROI's defined in section 2, the 95% LSD intervals for the interaction plot is shown in Fig. 8.

Fig. 8 confirms what was suspected for the previous analyses: for the ROI's, the percentage of valid pixels remains more or less constant from 1 to 4 latent variables (no statistically significant differences, P-values>0.05), no matter the number of directions used for building the model. However, when using 5 o 6 (OLS equivalent model) latent variables, the number of directions used has a clear impact on the percentage of valid ROIs pixels, this impact being especially severe in the case of OLS. This means that PLS with less than 5 LV appears as a much more robust model for computing the DTI, letting the number of directions used (and then the time required and cost) to be lesser.

Therefore, in clinical practice, the OLS model is not suitable when the number of directions in the acquisition is low. Conversely, the PLS model with less than 5 LV is barely affected when lowering the image functional resolution. This is illustrated in Fig. 9 that represents the value of λ_1 for the different models from the same case example of Fig. 5:

As observed, the PLS model with 1 LV remains consistent when the number of directions decreases, while the OLS method (6 LV) losses interpretation power and in the limiting case of 6 directions is not able to represent the appropriate parametric map due to the lack of valid pixels.

Finally, this effect has also been analyzed with a ROC curves analysis performed on the two biomarkers presented, dmax and 1-Rmax (Fig. 10).

For both biomarkers, there is a clear degradation of OLS (PLS with 6 LV) as the number of directions used is reduced and, although it presents the best ROC curves for 16 directions, they are really close to the 5 LV PLS models. In fact, for 6 directions, the best NLV is 4. This strengthens the idea that PLS is more robust, and at the same time assures a higher number of valid pixels; allowing for less time and money consuming acquisitions, thus allowing DTI to be affordable for less well-resourced health systems.

4. Conclusions

A new method is proposed for diffusion tensor calculation in breast cancer diagnosis based on PLS. The PLS technique has some advantages over the traditional calculation by OLS: i) the ability to provide higher percentage of valid pixels (pixels where all the elements of the tensor are positive), and ii) the flexibility of adapting the number of latent variables depending on the quality of the image, acting as a filter when the noise of the images is high. A compromise solution between the percentage of valid pixels and the interpretability of the biomarker maps is required to provide the appropriate number of latent variables for the PLS model.

New imaging biomarkers (dmax and 1-Rmax) have been proposed for identifying tumors showing good performance (in combination with the eigenvalues) for tumor detection.

Finally, the clinical added value of the PLS method is supported by the lower percentage of valid pixels available as the number of latent variables (NLV) increase; and more important, as the number of

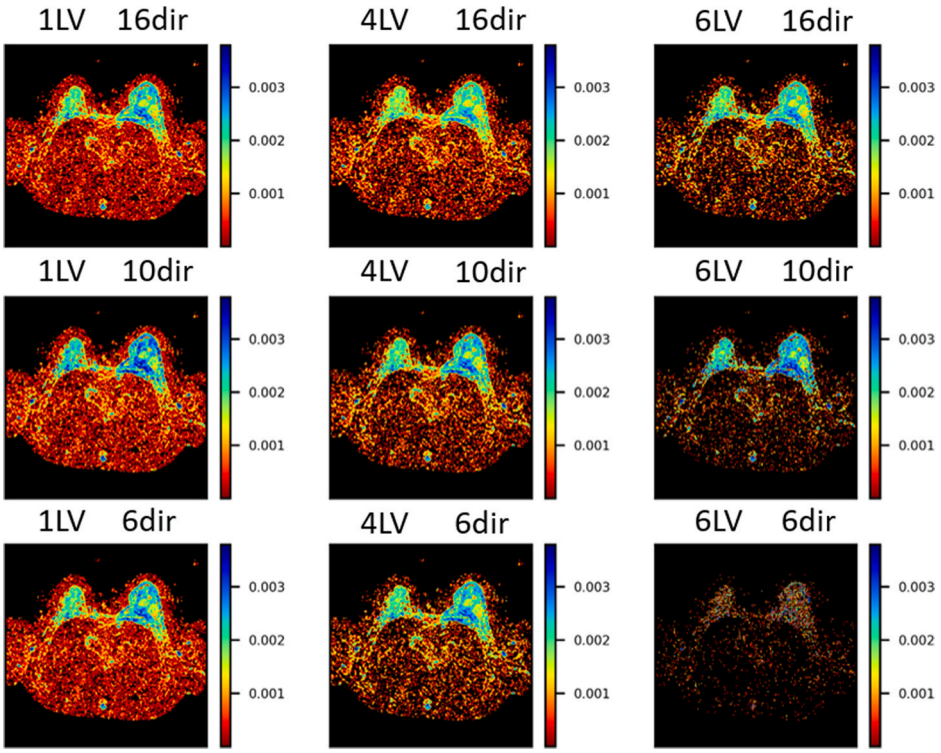

Fig. 9. λ_1 parametric map comparison between PLS models: from left to right (1LV, 4LV, 6LV) and top to bottom (16dir, 10dir, 6dir).

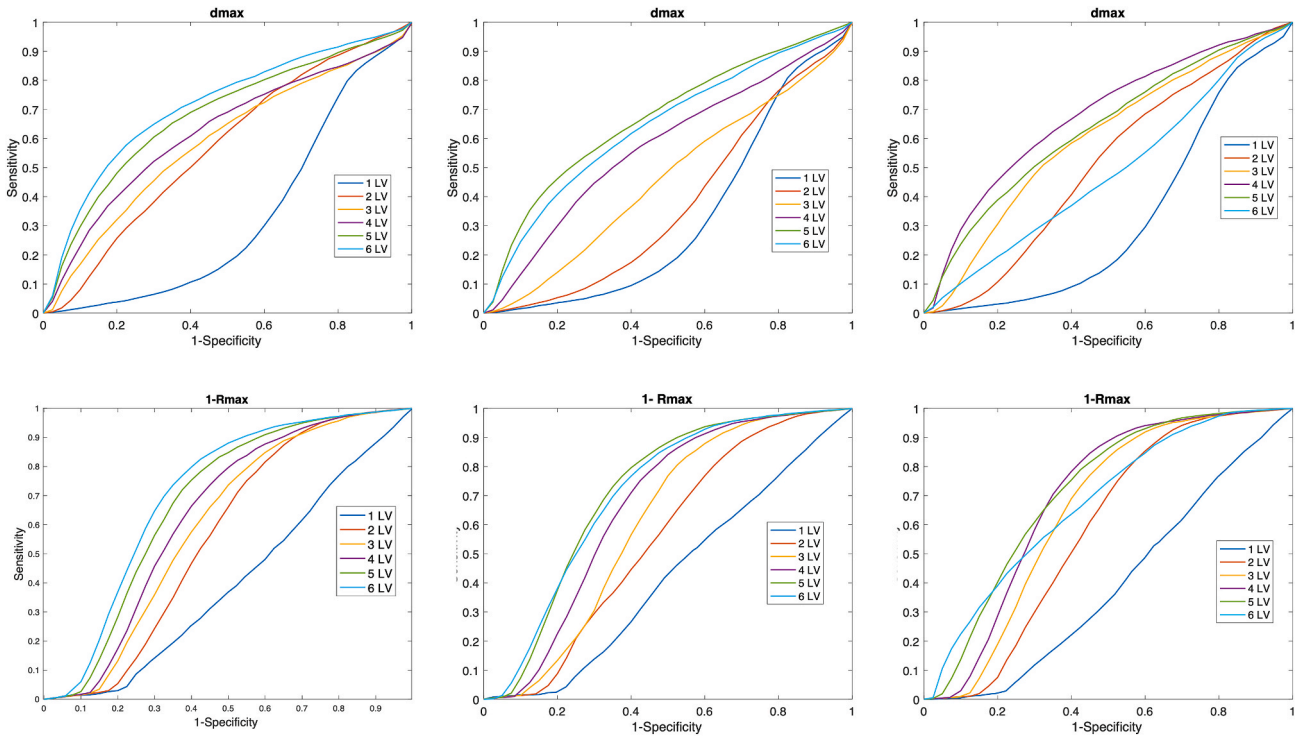

Fig. 10. Evolution of ROC curves from 16 (left), 10 (middle) and 6 (right) directions, and with NLV (legend) for dmax (top) and 1-Rmax (bottom).

directions used for the DTI calculation decreases. PLS models with low number of latent variables (less than 5) remains almost invariant, while the OLS model loses interpretability, and in the limiting case of 6 directions does not provide a valid model. Therefore, PLS is able to provide a valid model even when the equipment has low functional resolution (limited number of directions), reducing cost and acquisition time, in contrast with OLS, which cannot perform properly in these conditions. This has a tremendous impact on the scalability of the proposed approach as it increases the applicability of the DTI to many more radiology centers and hospitals consequent impact on early stage cancer detection.

CRediT Author statement

Eric Aguado-Sarrió: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing- Original Draft, Writing – Review & Editing, Visualization, **José Manuel Prats-Montalbán:** Conceptualization, Validation, Writing – Review & Editing, Visualization, Funding acquisition, **Guillermo Robles-Lozano:** Conceptualization, Methodology, Software, Writing- Original Draft, **Julia Camps-Herrero:** Validation, Investigation, Resources, Writing – Review & Editing, Supervision, **Alberto Ferrer:** Conceptualization, Validation, Writing – Review & Editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgments

This research was supported by the Spanish Government (Science and Innovation Ministry) under the project PID2020-119262RB-I00.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chemolab.2023.104777.

References

[1] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, Global cancer statistics: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 Countries, CA A Cancer J. Clin. 68 (6) (2018) 394–424.
[2] Cancer Today. International Agency for Research on Cancer. World Health Organization (https://gco.iarc.fr/today/home).
[3] W. Yue, Z. Wang, H. Chen, A. Payne, X. Liu, Machine learning with applications in breast cancer diagnosis and prognosis, Designs 2 (2) (2018) 13.
[4] A. Sheye, Diffusion MRI Methods for Improved Treatment Monitoring in Breast Cancer, University of California, Berkeley, 2009.
[5] A.M. Scaranelo, What's hot in breast MRI, Can. Assoc. Radiol. J. 73 (1) (2022) 125–140.
[6] D. Le Bihan, Molecular diffusion nuclear magnetic resonance imaging, Magn. Reson. Q. 7 (1991) 1–30.
[7] J.M. Prats-Montalbán, E. Aguado, A. Ferrer, Chapter 16: multivariate curve resolution for magnetic resonance image analysis: applications in prostate cancer biomarkers development, in: Resolving Spectral Mixtures, with Application from Ultrafast Spectroscopy to Super-resolution Imaging", Data Handling in Science and Technology, vol. 30, Elsevier, 2016, pp. 519–550.
[8] E. Aguado-Sarrió, J.M. Prats-Montalbán, R. Sanz-Requena, A. Alberich-Bayarri, L. Martí-Bonmatí, Ferrer A. Prostate, Diffusion weighted-magnetic resonance image analysis using multivariate curve resolution methods, Chemometr. Intell. Lab. Syst. 140 (2015) 43–48.
[9] D. Le Bihan, J.-F. Mangin, C. Poupon, C.A. Clark, S. Pappata, N. Molko, H. Chabriat, Diffusion tensor imaging: Concepts and applications, J. Magn. Reson. Imag. 13 (4) (2001) 534–546.
[10] P.J. Basser, D.K. Jones, Diffusion-tensor MRI: theory, experimental design and data analysis – a technical review, NMR Biomed. 15 (7–8) (2002) 456–467.
[11] E.O. Stejskal, J.E. Tanner, Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient, J. Chem. Phys. 42 (1965) 288–292.
[12] R. Jiang, Z. Ma, H. Dong, S. Sun, X. Zeng, X. Li, Diffusion tensor imaging of breast lesions: evaluation of apparent diffusion coefficient and fractional anisotropy and tissue cellularity, Br. J. Radiol. 89 (2016), 20160076.

[13] S.C. Partridge, N. Nissan, H. Rahbar, A.E. Kitsch, E.E. Sigmund, Diffusion-Weighted breast MRI: clinical applications and emerging techniques, J. Magn. Reson. Imag. 45 (2017) 337–355.
[14] A.A. Khalek, A. Razek, M. Zaky, D. Bayoumi, S. Taman, K. Abdelwahab, R. Alghandour, Diffusion tensor imaging parameters in differentiation recurrent breast cancer from post-operative changes in patients with breast-conserving surgery, Eur. J. Radiol. 111 (2019) 76–80.
[15] K. Wang, Z. Li, Z. Wu, Y. Zheng, S. Zeng, L. E, J. Liang, Diagnostic performance of diffusion tensor imaging for characterizing breast tumors: a Comprehensive meta-analysis, Front. Oncol. 9 (2019), 1229.
[16] S.C. Partridge, R.S. Murthy, A. Ziadloo, S.W. White, K.H. Allison, C.D. Lehman, Diffusion tensor magnetic resonance imaging of the normal breast, Magn. Reson. Imaging 28 (2010) 320–328.
[17] S.C. Partridge, A. Ziadloo, R. Murthy, S.W. White, S. Peacock, P.R. Eby, W. B. DeMartini, C.D. Lehman, Diffusion tensor MRI: preliminary anisotropy measures and mapping of breast tumors, J. Magn. Reson. Imag. 3 (2010) 339–347.
[18] P.A.T. Baltzer, A. Schäfer, M. Dietzel, D. Grässel, M. Gajda, O. Camara, W.A. Kaiser, Diffusion tensor magnetic resonance imaging of the breast: a pilot study, Eur. Radiol. 21 (2011) 1–10.
[19] O. Cakir, A. Arslan, N. Inan, Y. Anık, T. Sarısoy, S. Gumustas, G. Akansel, Comparison of the diagnostic performances of diffusion parameters in diffusion weighted imaging and diffusion tensor imaging of breast lesions, Eur. J. Radiol. 82 (2013) e801–e806.
[20] G.C. Baxter, M.J. Graves, F.J. Gilbert, A.J. Patterson, A meta-analysis of the diagnostic performance of diffusion MRI for breast lesion Characterization, Radiology (2019) 1–11, 00.
[21] M. Iima, M. Honda, E.E. Sigmund, A.O. Kishimoto, M. Kataoka, K. Togashi, Diffusion MRI of the breast: Current status and future directions, J. Magn. Reson. Imag. 52 (2020) 70–90.
[22] L.J. Wilmes, W. Li, H.J. Shin, D.C. Newitt, E. Proctor, R. Harnish, N.M. Hylton, Diffusion tensor imaging for assessment of response to neoadjuvant Chemotherapy in patients with breast cancer, Tomography 2 (4) (2016) 438–447.
[23] H.I. Greenwood, L.J. Wilmes, T. Kelil, B.N. Joe, Role of breast MRI in the evaluation and detection of DCIS: opportunities and challenges, J. Magn. Reson. Imag. 52 (2020) 697–709.
[24] E. Furman-Harana, E. Eyala, M. Shapiro-Feinbergb, N. Nissan, D. Grobgelda, N. Weisenberg, H. Degani, Advantages and drawbacks of breast DTI, Eur. J. Radiol. 81S1 (2012) S45–S47.
[25] E. Furman-Haran, N. Nissan, V. Ricart-Selma, C. Martinez-Rubio, H. Degani, J. Camps-Herrero, Quantitative evaluation of breast cancer response to neoadjuvant Chemotherapy by diffusion tensor imaging: initial results, J. Magn. Reson. Imag. 47 (2018) 1080–1090.
[26] B. Peter, Kingsley. Introduction to diffusion tensor imaging mathematics: Part I. Tensors, rotations, and eigenvectors in, Concepts Magn. Reson. 28A (2) (2006) 101–122.
[27] B. Peter, Kingsley. Introduction to Diffusion Tensor Imaging Mathematics: Part III. Tensor Calculation, Noise, Simulations, and Optimization. Concepts Magntetic Resonance Part A, Wiley Periodicals, New York, 2006, pp. 155–179.
[28] E. Furman-Haran, D. Grobgeld, N. Nissan, M. Shapiro-Feinberg, H. Degani, Can diffusion tensor anisotropy indices assist in breast cancer detection? J. Magn. Reson. Imag. 44 (2016) 1624–1632.
[29] P. Geladi, B.R. Kowalski, Partial least-squares regression: a tutorial, Anal. Chim. Acta 185 (1986) 1–17.
[30] S.B. Gurden, J.A. Westerhuis, R. Bro, A.K. Smilde, A comparison of multiway regression and scaling methods, Chemometr. Intell. Lab 59 (1-2) (2001) 121–136.
[31] A. Hoskuldsson, PLS regression methods, J. Chemometr. 2 (1998) 221–228.
[32] J.V. Kresta, T.E. Marlin, J.F. MacGregor, Development of inferential process models using PLS, Computers chem. Eng. 18 (7) (1994) 597–611.
[33] N. Nissan, T. Allweis, T. Menes, A. Brodsky, S. Paluch-Shimon, I. Haas, O. Golan, Y. Miller, H. Barlev, E. Carmon, M. Brodsky, D. Anaby, P. Lawson, O. Halshtok-Neiman, A. Shalmon, M. Gotlieb, R. Faermann, E. Konen, M. Sklair-Levy, Breast MRI during lactation: effects on tumor conspicuity using dynamic contrast-enhanced (DCE) in comparison with diffusion tensor imaging (DTI) parametric maps, Eur. Radiol. 30 (2020) 767–777.