Logistic regression to predict malignancy of breast tumors using IVIM parameters

M. Statache, B. Maris, Marco Barillari, Rossella Menghini, Adam Cybuliski, Giulia Zamboni
University of Verona, Computer Science, Verona, Italy

Keywords IVIM, segmentation, prediction, perfusion, diffusion, breast, tumour

Purpose

Breast tumour is the most frequent tumour affecting the female sex and represents 29% of all tumour cases. Multi-parametric MRI can provide give hints on the nature of a breast lesion, but its benign or malignant nature can only be determined for sure with a biopsy. But what if we could predict this (with a certain accuracy) only from a non-invasive imaging analysis?

One of the most advanced techniques is DW-MRI (diffusion-weighted MRI), a particular MRI sequence that does not require any type of contrast agent, thus resulting completely non-invasive and without contraindications. DW-MRI is based on the observation of the spin of water's molecules, after and before the application of a magnetic gradient, depending basically on the molecules' diffusion. The water molecules, naturally present in organs, act as a contrast agent.

The goal of this work is to predict the malignancy of a lesion from the analysis of DW-MRI in a retrospective study. The DW-MRI sequence is used to compute the intravoxel incoherent motion (IVIM) parameters that allow to divide the water movement into diffusion (due to the water present in the tissues) and perfusion (due to the water present in blood flowing in the capillaries). This second movement is not random, but oriented in the direction of the capillaries, but if we recall that capillaries are very short, randomly oriented and with a high density per volume, we can consider the perfusion as a "pseudo-diffusion". Knowing that benign and malign breast tumour have different perfusion characteristics, if we could identify and quantify this feature, we might be able to determine the type of the tumour.

In this work, we use state of the art algorithms to compute the IVIM parameters which are then plugged into a learning algorithm, based on retrospective data, that infer the malignancy of the lesion.

Method

The actual DW-MRI protocol allows to measure only ADC, i.e. Apparent Diffusion Coefficient, which contains both diffusion and perfusion terms. Following the work of Le Bihan [1], we compute the perfusion coefficient (called D*) separating it from the diffusion phenomenon (here the coefficient is called D).

The different gradient intensities in the DW-MRI data is given by a so-called b-value. Taking data with different b-values allows also a higher accuracy in estimating ADC. Anyway, the main concept is that with high b-values the principal phenomenon involved is diffusion, while with low b-values it is clearly appreciable also the perfusion event, related to microcirculation. With this consideration we make use of the bi-exponential model of IVIM [2]:

$$\frac{S_b}{S_0} = (1 - f_{IVIM}) * e^{-b*D} + f_{IVIM} * e^{-b*(D+D^*)} \quad (1)$$

Here, the f_{IVIM} coefficient gives the volume fraction of incoherently flowing blood in the tissue or the flowing vascular volume, S_b is the intensity of a voxel at the current b, while S_0 is the intensity of the same voxel at b=0.

In a typical DWI breast sequence, the following values of b are used: [0 10 20 30 50 80 100 200 400 1000 2000]. In a simplified approach and under some approximations, we calculate $ADC_{0\text{-}200}$ and $ADC_{200\text{-}2000}$ using

the mono-exponential model $\frac{S_b}{S_0} = e^{-b*ADC}$. We can write this formula also as $\ln\left(\frac{S_b}{S_0}\right) = -b * ADC$, where $ln\left(\frac{S_b}{S_0}\right)$ is ***y*** and b is ***x***, obtaining $y = -ADC * x$. Now, replacing ***x*** first with the b-values between 0 and 200 and their corresponding ***y*** values, and then with the b-values from 200 to 2000, we can fit two lines, whose slopes give the estimation of $ADC_{0\text{-}200}$ and $ADC_{200\text{-}2000}$. Then, we calculate f_{IVIM} with the relation

$$f_{IVIM} = b_{boundary}(ADC_{0-b_{boundary}} + ADC_{200-b_{boundary}})$$

where $b_{boundary}$ is the b-value from where perfusion phenomenon is no longer evident and the diffusion gets involved, which is 200 in our case

$$f_{IVIM} = 200 * (ADC_{0-200} + ADC_{200-2000}).$$

The final step is to compute D* by using the formula (1).

The computation of these parameters is done by using a ROI (region of interest) containing the full DW-MRI sequence of the suspected lesion. On each ROI we then calculate the mean of the IVIM parameters: D, D* and f_{IVIM}.

The analysis is performed on a cohort of 25 patients of which 11 with benign tumours and 14 with malignant tumours. All patients underwent biopsy after the MRI scanning and prior to the study.

A logistic regression algorithm based on non-linear decision boundaries was employed to compute a prediction model based on these training data.

The model was then tested on 5 patient that were not included in the training data.

All the software was implemented in Matlab (MathWorks, Inc., Natick, Massachusetts, United States).

Results

Fig.1 shows the plots of the two classes of lesions divided by the decision boundary computed through the logistic regression algorithm.

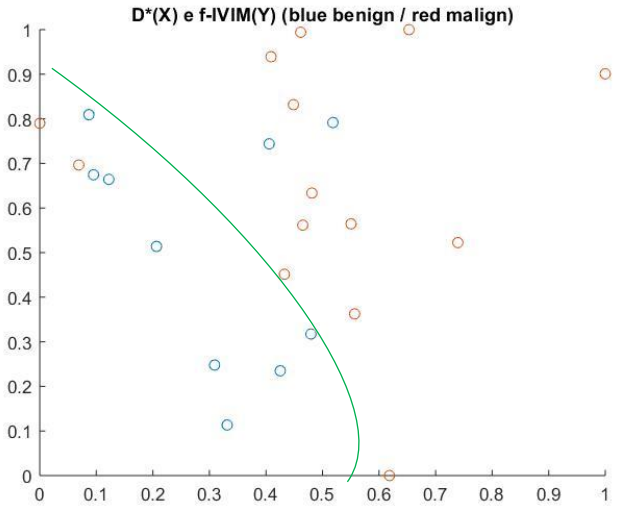


Fig.1 Plot of the benign and malign breast tumour.

We have a few outliers, due to some errors and/or noise that we have not considered. The overall accuracy of the prediction is 85% for now. We are introducing more patient data in the study so that the model will improve and we expect to reach an accuracy of the prediction of more than 90%.

We have also calculated a colormap for each parameter, which can be useful for a medical direct observation. An example is shown in *Fig.2*.

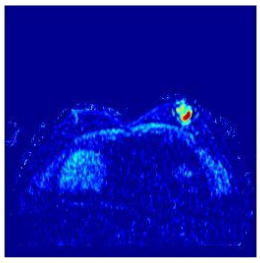
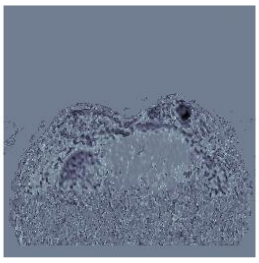
(a)  (b) 

Fig.2 (a) Colormap of a malign breast tumour. (b) 'Bone-grey' map of the same section.

Conclusion

A method to predict malignancy of a breast tumour based on DW-MRI data was presented. The method uses the IVIM analysis to compute the perfusion and diffusion coefficients in a ROI representing the suspicious area. The coefficients are then used to compute the boundary of a logistic regression model that allows to predict the malignancy and keep learning from new results.

References

[1] Le Bihan D, Breton E, Lallemand D, et al. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. Radiology 1988; 168:497-505.

[2] Du J, Li K, Zhang W, et al. Intravoxel incoherent motion MR imaging: comparison of diffusion and perfusion characteristics for differential diagnosis of soft tissue tumors. Medicine 2015; 94:1-8.