*Article*

# Differentiating False Positive Lesions from Clinically Significant Cancer and Normal Prostate Tissue using VERDICT MRI and Other Diffusion Models

**Snigdha Sen [1], Vanya Valindria [1], Paddy J. Slator [1], Hayley Pye [2], Alistair Grey [3], Alex Freeman [4], Caroline Moore [3], Hayley Whitaker [2], Shonit Punwani [5], Saurabh Singh [5,†] and Eleftheria Panagiotaki [1,*,†]**

[1] Centre for Medical Image Computing, Department of Computer Science, University College London, London, United Kingdom; snigdha.sen.20@ucl.ac.uk
[2] Molecular Diagnostics and Therapeutics Group, University College London, London, United Kingdom; h.pye@ucl.ac.uk
[3] Department of Urology, University College London Hospitals NHS Foundations Trust, London, United Kingdom; alistair.grey@ucl.ac.uk
[4] Department of Pathology, University College London Hospitals NHS Foundations Trust, London, United Kingdom; arya.freeman@ucl.ac.uk
[5] Centre for Medical Imaging, University College London, London, United Kingdom; s.punwani@ucl.ac.uk
* Correspondence: e.panagiotaki@ucl.ac.uk
† These authors contributed equally to this work.

**Abstract:** False positives on multiparametric (mp)-MRI result in a large number of unnecessary biopsies in men with clinically insignificant diseases. This study investigates whether quantitative diffusion MRI can improve differentiation between false positives, true positives and normal tissue. Twenty-three patients underwent mp-MRI and Vascular, Extracellular and Restricted Diffusion for Cytometry in Tumours (VERDICT)-MRI, followed by transperineal biopsy. The patients were categorised into two groups following biopsy: (1) significant cancer - true positive (2) atrophy/inflammation/high-grade prostatic intraepithelial neoplasia (PIN) - false positive. The clinical apparent diffusion coefficient (ADC) values of the lesions were obtained, and the intravoxel incoherent motion (IVIM), diffusion kurtosis imaging (DKI) and VERDICT models were fitted using a deep learning approach. Significant differences ($p < 0.05$) between true positive and false positive lesions were found in ADC, IVIM perfusion fraction ($f$) and diffusivity ($D$), DKI diffusivity ($D_K$) and kurtosis ($K$) and VERDICT intracellular volume fraction ($f_{IC}$), extracellular-extravascular vol-ume fraction ($f_{EES}$) and diffusivity ($d_{EES}$) values. Significant differences between false positives and normal tissue were only found for the VERDICT $f_{IC}$. These results demonstrate that model-based diffusion MRI could reduce the number of unnecessary biopsies due to false positive prostate lesions and shows promising sensitivity to benign diseases that mimic cancer.

**Keywords:** prostate cancer; diffusion MRI; false positives; biophysical modelling; deep learning

## 1. Introduction

Multiparametric MRI (mp-MRI) has recently been introduced as a standard part of the prostate cancer (PCa) clinical diagnosis pathway [1] – it consists of T1 and T2-weighted images, diffusion-weighted (DW) images and dynamic contrast-enhanced (DCE) imaging. This technique has high sensitivity (90%) but moderate specificity (50%), translating to a high rate of false positive cases [2]. This results in 1 in 2 men undergoing mp-MRI having unnecessary uncomfortable biopsies and risking the associated side effects for benign conditions or clinically insignificant cancer [3]. This is a significant issue as 75% of suspected cancer patients have abnormal mp-MRI findings and the number of people considered for MRI and biopsy each year is set to increase, therefore reducing the number of unnecessary negative biopsies is an important clinical problem [4]. Benign pathologies such as atrophy, inflammation, and high-grade prostatic intraepithelial neoplasia (PIN) are examples of diseases that cause these false positive results [5]. This is due to the changes these diseases cause to the microstructure, similar to cancer. For example, atrophy is characterised by shrinkage of prostate tissue due to the reduction in cytoplasm prostatic acinar cells and has been associated with prostatic inflammation (swelling of the prostate gland) [6]. High-grade PIN represents the pre-invasive end of the range of cellular proliferations within the lining of prostatic ducts and acini and is considered the most likely precursor of prostatic carcinoma [7]. It is critical to discriminate these cases from cancer to avoid unnecessary procedures, however it is also important to distinguish these diseases from normal tissue and identify correctly the type of the disease [8]. Some of the benign diseases can present with symptoms similar to PCa such as difficult or frequent urination and pain, requiring treatment of their own [9]. Identifying a unique noninvasive signature for such diseases can lead to early and informed treatments.

DW-MRI is an integral component of mp-MRI due to the unique insight it provides into tissue microstructure. Changes in histological features such as the cellular density, size, shape and arrangement produce contrast in DW-MR images as they all affect tissue-water mobility. Most studies using DW-MRI have focused on calculating the apparent diffusion coefficient (ADC) to distinguish between tumour regions and healthy tissue [10,11]. Typically, ADC values are lower in prostate tumours than in the surrounding tissue, reflecting the highly cellular environment constraining the water mobility. However, the simultaneous dependence of the ADC on a multitude of histological features limits its biological specificity [12], reducing its ability to distinguish cancer from similar diseases such as high-grade PIN and hyperplasia, which often appear as false positive cases [13,14]. More sophisticated models have been proposed to improve the sensitivity and specificity of DW-MRI for cancer diagnosis such as the intravoxel incoherent motion (IVIM) model that separates the pure water diffusion in tissue from the microcirculation of water in capillaries [15]. It has been used to study various cancer types such as breast [16], prostate [17] and pancreatic [18] tumours, showing improvement in data description over ADC. Another method that has shown greater sensitivity for discrimination of benign and cancerous prostate tissue in comparison to ADC is using diffusion kurtosis imaging (DKI) [19,20]; this technique quantifies the Gaussian and non-Gaussian components of water diffusion in biological tissues [21].

In an attempt to increase biological specificity, multicompartment microstructure models have also been proposed for imaging the prostate. One of the first multicompartment methods for cancer imaging is the Vascular, Extracellular and Restricted Diffusion for Cytometry in Tumours (VERDICT) MRI, which is a non-invasive imaging technique for quantifying microstructural features of tumours in vivo. It consists of a specific imaging protocol, as well as a model for the DW-MRI signal [22]. VERDICT allows for estimation of specific tissue properties such as cell size and packing density. It has been successful in delineating benign from cancerous lesions [23], and preliminary results from the clinical trial INNOVATE [24] reveal that the VERDICT intracellular volume fraction can discriminate between Gleason 3+3 and 3+4 lesions, unlike ADC [25].

This study uses different DW-MRI techniques to investigate quantitative differences between clinically significant cancer, false positive biopsy results and healthy tissue. The aims are i) to discriminate false positives from cancer and ii) to discriminate false positives from normal tissue in an attempt to identify potential diffusion signatures of benign diseases that mimic cancer. We analyse a total of 23 patients that underwent mp-MRI followed by VERDICT-MRI. We fit the diffusion models to the VERDICT-MRI data and obtain the ADC from the mp-MRI. The model fitting procedure uses deep neural networks (DNNs) for ultra-fast and robust parameter estimation [26]. We compare parameter estimates between both false positives and clinically significant cancer and normal tissue and false positives using statistical tests. We investigate the diagnostic accuracy of different parameters using receiver operating characteristic (ROC) curves and analyse the correlation between VERDICT parameters and those from the simpler models. This work demonstrates that diffusion MRI models can differentiate between false positives and true cancer, and that VERDICT has the further sensitivity to discriminate false positives from normal tissue.

## 2. Materials and Methods

### *Patient Cohort*

The study was performed with local ethics committee approval embedded within the INNOVATE clinical trial [24]. The trial is registered with ClinicalTrials.gov identifier NCT02689271. Twenty-three men (median age: 68 years, age range: 57-74 years) were recruited and provided informed written consent. The inclusion criteria were as follows:

- suspected PCa

undergoing active surveillance for known PCa.

Exclusion criteria included:

- inability to have a MRI scan, or presence of an artefact that would reduce quality of MRI
- previous hormonal/radiation therapy or surgical treatment for PCa
- biopsy within 6 months prior to the scan.

All patients underwent mp-MRI in line with international guidelines [27] on a 3T scanner, supplemented by VERDICT-MRI. The clinical dynamic contrast enhanced (DCE) part of mp-MRI was performed last after VERDICT DW-MRI. After clinical mp-MRI and VERDICT MRI indicated suspected PCa, all patients underwent targeted transperineal template biopsy of their index lesion, defined as the highest scoring lesion identified on mp-MRI with Likert scores (3-5) [28]. Specialist genitourinary pathologists evaluated histological specimens stained with haematoxylin and eosin from the biopsy cores. Patients who had a biopsy with diagnoses of atrophy, inflammation, and high-grade PIN (or combinations of these) or clinically significant PCa were retrospectively selected. Table 1 presents the clinical and pathological information of the patients.

**Table 1.** Clinical and pathological information of 23 patients included in the study (12 with no/clinically insignificant cancer and 11 with clinically significant cancer). The median age and PSA/PSAD results for each cohort are presented, along with the biopsy results/Gleason scores. The false positive patient cohort had combinations of the three disease types considered.

| | **No/clinically insignificant cancer** | **Clinically significant cancer** |
|---|---|---|
| Age (median) | 66 | 69 |
| PSA (median) | 6.61 | 10.92 |
| PSAD (median) | 0.112 | 0.226 |
| Biopsy Result | Atrophy: 9 | 3+3: 1 |
| | Inflammation: 10 | 3+4: 4 |
| | High-grade PIN: 4 | 4+3: 4 |
| | | 4+4: 1 |
| | | 4+5: 1 |

### *Image Acquisition*

Mp-MRI

All participants underwent mp-MRI with a 3T MRI system (Achieva; Philips, Best, the Netherlands) as part of their standard clinical care. A spasmolytic agent (Buscopan, Boehringer Ingelheim, Ingelheim am Rhein, Germany; 0.2 mg/kg, up to 20 mg) was administered intravenously before imaging to reduce bowel peristalsis. Imaging parameters for the diffusion-weighted echo-planar imaging sequences were as follows: repetition time 2753 ms, echo time 80 ms; field of view, 220 x 220 mm; section thickness, 5 mm; no intersection gap; acquisition matrix, 168 x 169; b-values, 0, 150, 500, and 1000 s/mm$^2$; and six signals acquired per b-value for signal averaging. The total imaging time for the clinical diffusion-weighted sequences was 5 minutes 16 seconds. ADC maps were calculated by using all b-values except b = 0 to reduce perfusion effects [29] and were calculated with the Camino Diffusion MRI toolkit [30].

### VERDICT-MRI

VERDICT-MRI was performed before dynamic contrast material–enhanced imaging on the same 3T unit as the clinical mp-MRI acquisition. A PGSE sequence was used at five combinations of b-value (in s/mm$^2$), gradient duration, separation, echo time TE and repetition time TR (in ms), in three orthogonal directions using a cardiac coil. For each combination, a separate b = 0 image was acquired. Sequences used an echo-planar readout, and imaging parameters were as follows: repetition time 2482–3945 ms/echo time 50–90 ms; field of view, 220 x 220 mm; voxel size, 1.3 x 1.3 x 5 mm; no intersection gap; acquisition matrix, 176 x 176; b-values, 90, 500, 1500, 2000, and 3000 s/mm$^2$; and six signals acquired per b-value (except for b = 90 s/mm$^2$, for which four signals were acquired) for signal averaging. The total imaging time was 12 minutes 25 seconds [31].

## *Image Analysis*

### ROIs

Patients were biopsied depending on their mp-MRI score as reported by two board-certified experienced uroradiologists (reporting more than 2,000 prostate MR scans per year). The regions of interest (ROI) were targeted drawn by a board-certified study radiologist (S.Singh) using a pictorial report made by the uroradiologist and confirmed as cancerous or non-cancerous retrospectively by transperineal biopsy. The ROIs were chosen to be as large as possible while having minimal contamination from surrounding tissue. It was concluded that 12 of the patients had benign pathologies such as atrophy, inflammation or high-grade PIN, whilst the remaining 11 had cancerous prostate lesions. After a review of the biopsy result confirmed the absence of tumour on the contralateral side of the peripheral zone for the 11 patients with PCa, ROIs were located for each patient in an area of benign tissue to be used for comparison.

### DW-MRI Data Pre-processing

The pre-processing pipeline included denoising of the raw DW-MRI using MP-PCA [32] as implemented within MrTrix3 [33] 'dwidenoise', and then correction for Gibbs ringing [34] with custom code in MATLAB (The Mathworks Inc., Natick, Massachusetts, USA). In an effort to reduce possible artefacts caused by patient movement during scanning and eddy current distortions, we applied mutual-information rigid and affine registration using custom code in MATLAB [35].

### Mathematical Models

The ADC model is a simple mono-exponential characterising the diffusion signal decay as a function of the b-value. It assumes an isotropic Gaussian distribution of water molecule displacements and has one parameter to be estimated; the ADC, $d$. The normalised signal is given by

$$S = e^{-bd}$$

The IVIM model is a bi-exponential with the assumption that the diffusion signal is made up of two non-exchanging compartments of water molecules (one fast and one slow), each following an ADC model. There are three parameters to be estimated; $f$, the volume fraction associated with the fast ('vascular') component, $D^*$, the diffusivity of the 'fast' component and $D$, the diffusivity of the slow ('cellular') component (36). The normalised signal is given by

$$S = f e^{-b(D+D*)} + (1-f)e^{-bD}$$

The mean signal DKI model relaxes the assumption in the ADC model of Gaussian water dispersion. There are two parameters to be estimated; $D_K$ and $K$. The diffusivity parameter $D_K$ is similar to the ADC parameter $d$, whilst the kurtosis parameter $K$ quantifies the degree of deviation of the dispersion pattern from a Gaussian distribution (37). The normalised signal is

$$S = e^{-bD_K + \frac{1}{6}b^2 D_K^2 K}$$

The VERDICT model for prostate is the sum of three parametric models, each describing the diffusion magnetic resonance signal in a separate population of water from one of the three components:

- signal S1 comes from intracellular water trapped within cells (including epithelium)
- signal S2 comes from extracellular-extravascular water adjacent to, but outside cells and blood vessels (including stroma and lumen)
- signal S3 comes from water in blood undergoing microcirculation in the capillary network.

It is assumed that there is no water exchange between the three tissue compartments. The total signal for the multi-compartment VERDICT model is:

$$S = \sum_{i=1}^{3} f_i S_i$$

where fi is the proportion of signal with no diffusion weighting (b = 0) from water molecules in population $i$, where $i = IC, VASC$ or $EES$, $0 \le f_i \le 1$, $\sum_{i=1}^{3} f_i = 1$ . This results in

$$S = f_{VASC} S_{VASC}(d_{VASC}, b) + f_{IC} S_{IC}(d_{IC}, R, b) + f_{EES} S_{EES}(d_{EES}, b)$$

where $f_{VASC} + f_{IC} + f_{EES} = 1$.

The VERDICT model used in this work [38] represents the intracellular component as spheres of radius $R$ and intra-sphere diffusivity fixed at $d_{IC}$ = μm²/ms, the extracellular-extravascular component as Gaussian isotropic diffusion with effective diffusivity $d_{EES}$ (Ball) and the vascular component as randomly oriented sticks with intra-stick diffusivity fixed at $d_{VASC}$ = μm²/ms (AstroSticks). In total, there are four model parameters that are estimated by fitting Eq. 5 to DW-MRI data: $f_{EES}$, $f_{IC}$, $R$ and $d_{EES}$. The vascular signal fraction, $f_{VASC}$, is computed as $1 - f_{IC} - f_{EES}$ and a cellularity index is computed as $f_{IC}/R^3$. Several previous studies [23,31] investigated the validity of the assumptions made in this model under the experimental conditions of the optimised DW-MRI acquisition for VERDICT in prostate.

### Model Fitting

The IVIM, DKI and VERDICT models were fitted to the DW-MRI data using the signal averaged across three gradient directions. To obtain ultra-fast and robust parameter estimation, we performed the fitting using a DNN known as a multi-layer perceptron, implemented using the 'MLPregressor' in Python scikit-learn 0.23 (https://scikitlearn.org/stable/). The input of the DNN is a vector of DW-MRI signals for each combination of b, TE and TR (a total of 10 in this specific case). The DNN consists of three fully connected hidden layers with 150 neurons, each characterised by a linear matrix operation followed by an element-wise rectified linear unit function (ReLU), and a final regression layer with the number of output neurons equal to the number of tissue parameters to be estimated (i.e. four for the VERDICT models used here). We generate 100,000 synthetic DW-MRI signals for training the DNN using the signal equations above with different values for the model parameters randomly chosen between biophysically plausible intervals: $f$ = [0.01, 0.99], $D$ = [0.5, 3] μm²/ms and $D^*$ = [0.5, 3] μm²/ms for IVIM, $D_K$ = [0.5, 3] μm²/ms and $K$ = [0.01, 2.99] for DKI and $f_{EES}$ = [0.01, 0.99], $f_{IC}$ = [0.01, 0.99], $R$ =

[0.01, 15] µm and $d_{EES}$ = [0.5, 3] µm$^2$/ms for VERDICT. We also add Rician noise corresponding to SNR=35 to consider the effect of experimental noise [26,35,39].

### Statistical Analysis

The performance of each parameter for differentiating between the three tissue types was evaluated via a Wilcoxon signed-rank test (preceded by a Shapiro-Wilk test for normality). This was performed using the scipy.stats package [40], and $p < 0.05$ was taken to indicate significance. The data is presented using boxplots, allowing for visualisation of the median, interquartile range and any outliers. ROC curves were plotted for all models to compare sensitivity and specificity, and the area under the curve (AUC) was calculated from these to compare the parameters' utility for tissue type discrimination. We also investigated the correlation between the VERDICT parameters and those from the DKI and IVIM models via the r$^2$-value.

## 3. Results

The aim of the first experiment was to analyse differences in parameter estimates in false positive and true positive lesions from the different models. In Fig. 1 we see comparisons between parameter estimates from the ADC, IVIM, DKI and VERDICT models in these two tissue types. All models provided discrimination between false positives and true positives: ADC $d$ with $p = 0.04$ (Wilcoxon signed-rank test), IVIM $f$ with $p = 0.02$ and $D$ with $p = 0.01$, DKI $D_K$ with $p = 0.0009$ and $K$ with $p = 0.001$ and VERDICT $f_{IC}$ with $p = 0.01$, $f_{EES}$ with $p = 0.02$ and $d_{EES}$ with $p = 0.0009$. $d$, $f$, $D$, $D_K$, $f_{EES}$ and $d_{EES}$ are all lower in cancerous lesions than in false positives, while $K$ and $f_{IC}$ values are higher in false positive lesions. The next experiment analysed differences in parameter estimates between false positive lesions and normal tissue – these are also presented in the boxplots in Fig. 1. The only parameter that shows statistically significant differences between false positives and normal tissue is the VERDICT $f_{IC}$ ($p = 0.02$), with normal tissue having lower values than false positives.

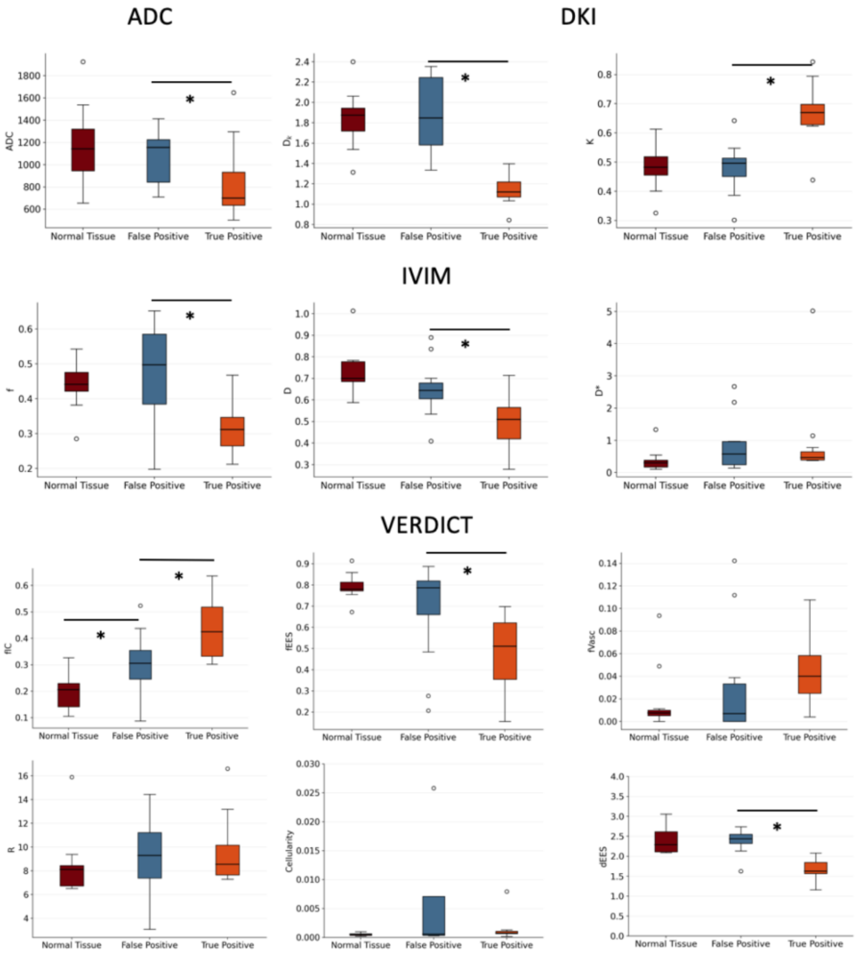

**Figure 1.** Boxplots showing the parameter estimates obtained using the ADC, DKI, IVIM and VERDICT models. We observe significant differences between true and false positives in the ADC $d$ ($p$ = 0.04), IVIM $D$ ($p$ = 0.01) and $f$ ($p$ = 0.02), DKI $D_K$ ($p$ = 0.0009) and $K$ ($p$ = 0.001) and VERDICT $f_{IC}$ ($p$ = 0.01), $f_{EES}$ ($p$ = 0.02) and $d_{EES}$ ($p$ = 0.0009). $d$, $D_K$, $D$, $f$, $f_{EES}$ and $d_{EES}$ are all lower in true positives than false positives while $K$ and $f_{IC}$ are higher. We also find statistically significant differences between false positives and normal tissue for the VERDICT $f_{IC}$ ($p$ = 0.02), with higher values in false positive lesions than in normal tissue. Outliers are indicated by a circle.

Next, we analysed the parameter maps obtained using the different models, focusing on those parameters with statistically significant differences between true positive and false positive lesions. Figure 2 illustrates parametric maps for two example patients: a 70-year-old with a false positive lesion in the left anterior prostate (atrophy and mild focal chronic inflammation) and a 72-year-old with a Gleason score 4+3 tumour in the right posterior peripheral zone (PZ). The parameter maps firstly show clear differences in the data between the true cancer ROIs and the surrounding tissue, improving lesion conspicuity over the mp-MRI images. They also demonstrate that the true positive lesions are more noticeably different from the surrounding healthy tissue than the false positive lesions. We observe that the $f_{IC}$ and $K$ are higher in lesions than the surrounding tissue, whilst for all the other parameters the values are lower in the lesions than in the surrounding tissue. The VERDICT $f_{IC}$, $f_{EES}$ and $d_{EES}$ strongly highlight the true positive tumour in comparison to the surrounding benign tissue, showing a clearer difference compared to the other models. The VERDICT $f_{IC}$ map also significantly highlights the false positive lesion, which is not evident in the other parametric maps.

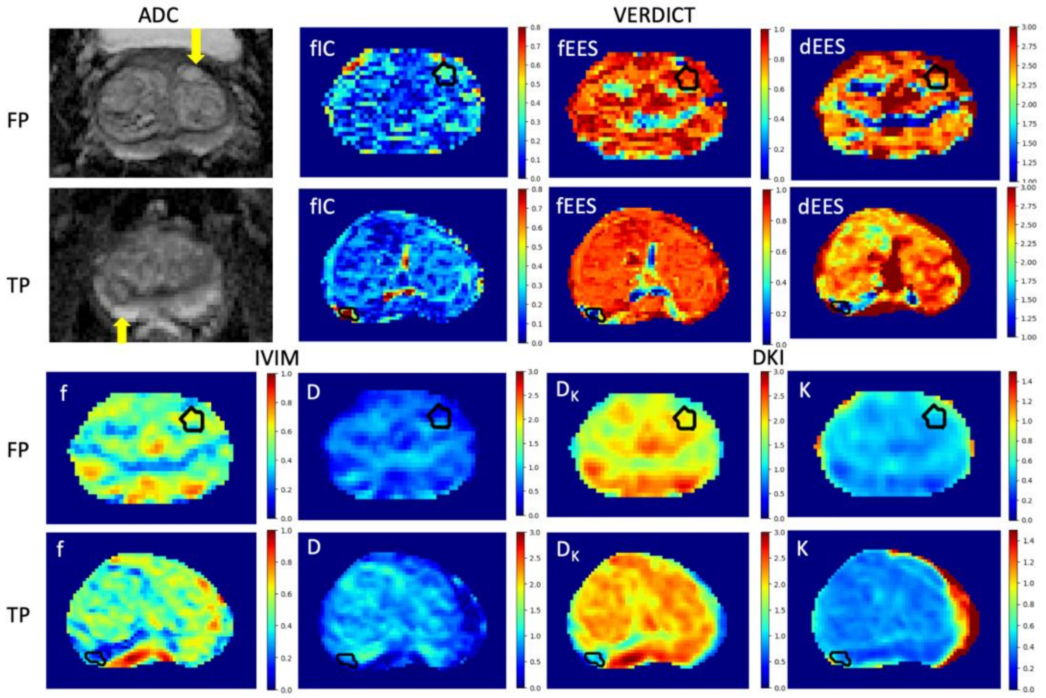

**Figure 2.** Parametric maps obtained using the IVIM, DKI and VERDICT models in a 70-year-old patient with a false positive lesion and a 72-year-old patient with a true positive lesion. Only the parameters which successfully differentiated between the two lesion types have been included, as well as the clinical ADC maps. We observe that the VERDICT maps highlight the true positive lesion most conspicuously, and the VERDICT $f_{IC}$ also distinguishes the false positive lesion from the surrounding tissue.

Following this, we investigate the diagnostic accuracy of the different parameters using ROC curves – we compare the performance of all parameters that can successfully discriminate between true and false positive lesions. Figure 3 presents discrimination between normal tissue and false positives on the left and between false positives and true positives on the right. We observe that the largest AUC value for discrimination between true positives and false positives is found for the DKI $D_K$ (0.9028), closely followed by the VERDICT $d_{EES}$ (0.8750). For the discrimination between false positives and normal tissue, the largest AUC is found for the VERDICT $f_{IC}$ (0.7778) – all other parameters show low diagnostic accuracy for this task.

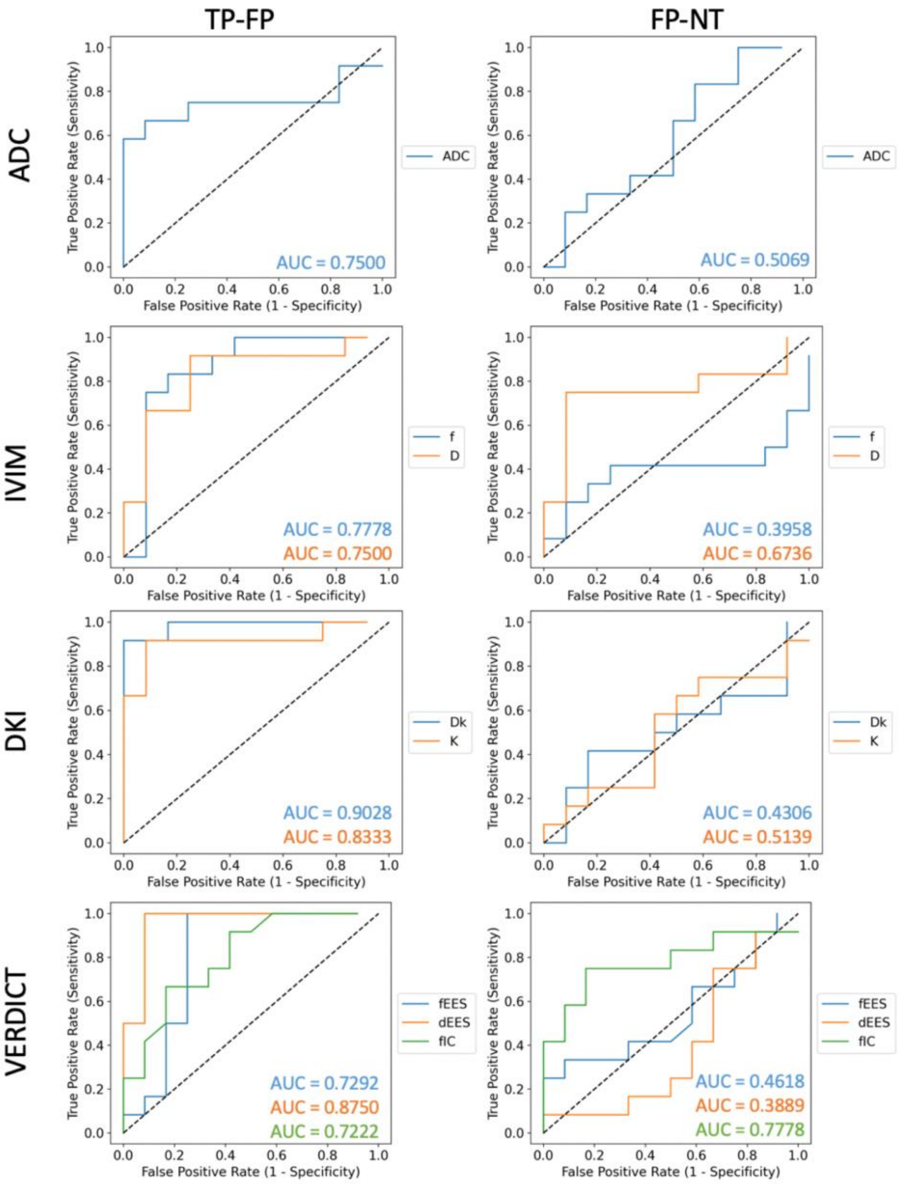

**Figure 3.** ROC curves for ADC, DKI, IVIM and VERDICT parameters – those on the left are for discriminating true (TP) and false positives (FP) and those on the right are for discriminating false positives and normal tissue (NT). We observe the largest AUC for discrimination between true and false positives is achieved by the DKI $D_K$ (AUC = 0.9028), closely followed by the VERDICT $d_{EES}$ (AUC = 0.8750). The largest AUC for discrimination between false positives and normal tissue is achieved by the VERDICT $f_{IC}$ (AUC = 0.7778).

The final experiment investigates the correlations between the VERDICT $f_{IC}$, $R$, $f_{EES}$, $f_{VASC}$, IVIM $D$, $f$, $D^*$ and DKI $D_K$, $K$ for each voxel within the ROIs of all patients – the strongest correlations are presented in Fig. 4. The colour coding distinguishes individual patients, with the cancer ROIs shown as circles whilst the benign ROIs are shown as crosses. We observe similar trends for $D_K$ and $D$ - strong negative correlation with $f_{IC}$ and positive correlation with $f_{EES}$ and $d_{EES}$. $K$ shows strong positive correlation with $f_{IC}$ and negative correlation with $f_{EES}$ and $d_{EES}$. Finally, $D^*$ shows negative correlation with $f_{EES}$ and positive correlation with $f_{VASC}$.

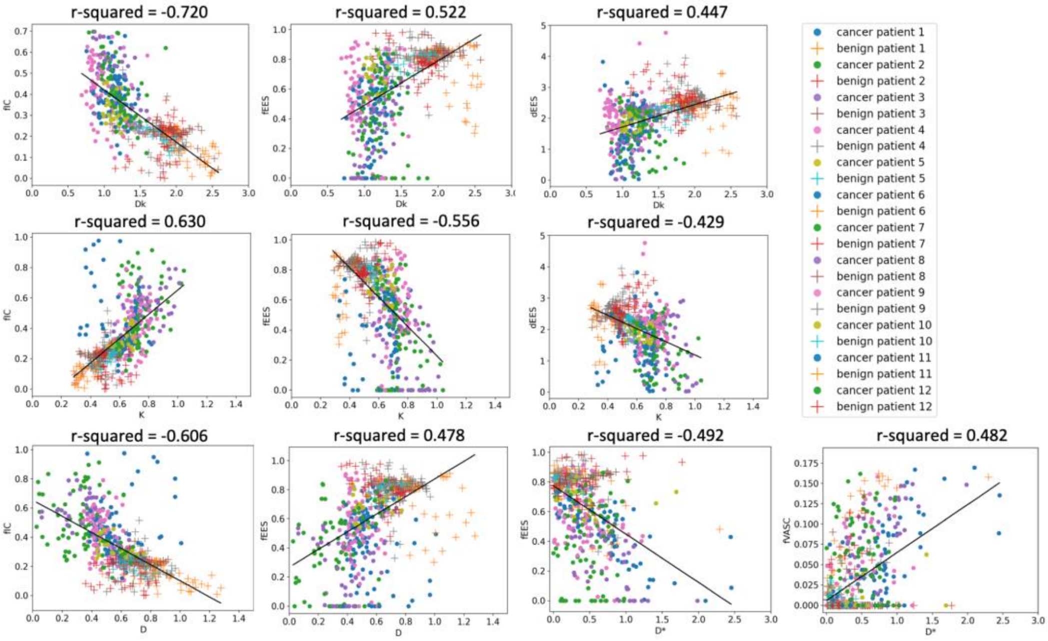

**Figure 4.** Scatter plots showing the correlation between VERDICT parameters and the DKI and IVIM parameters. For $D_K$ and $D$, we observe strong negative correlation with $f_{IC}$ and positive correlation with $f_{EES}$ and $d_{EES}$. $K$ shows strong positive correlation with $f_{IC}$ and negative correlation with $f_{EES}$ and $d_{EES}$. Finally, $D^*$ shows negative correlation with $f_{EES}$ and positive correlation with $f_{VASC}$.

## 4. Discussion

Benign prostate pathologies such as atrophy, inflammation and high-grade PIN can cause false positives on mp-MRI by having signal characteristics that mimic PCa. We aimed to firstly differentiate false positives from true cases of PCa, and then to discriminate false positives from normal tissue using various diffusion MRI models. We analysed the clinical ADC and then fitted the IVIM, DKI and VERDICT models to the acquired DW-MRI from 23 patients using a deep learning approach. We then compared the parameter estimates of each of the models between tissue types using statistical tests to draw conclusions about the diagnostic utility of the different models for characterising false positive cases of PCa.

All diffusion models can discriminate false positives from true positives. The strongest statistical significance is observed for the DKI $D_K$ and the VERDICT $d_{EES}$ ($p = 0.0009$), followed by the DKI $K$ ($p = 0.001$), VERDICT $f_{IC}$ and IVIM $D$ ($p = 0.01$), VERDICT $f_{EES}$ and IVIM f ($p = 0.02$) and ADC $d$ ($p = 0.04$). The $D_K$, $d_{EES}$, $D$, $f_{EES}$, $f$ and $d$ are lower in true cancer than in false positives, whilst $K$ and $f_{IC}$ are higher. This reflects the reduced diffusivity and larger deviations from Gaussian dispersion due to the increased cellularity in prostate carcinoma than in non-cancer diseases [41]. No significant differences were found in IVIM $D^*$, VERDICT $f_{VASC}$, $R$ or Cellularity. The best diagnostic performance was found for $D_K$ (AUC = 0.9028), closely followed by the VERDICT $d_{EES}$ (AUC = 0.8750), and the parametric maps obtained using VERDICT emphasise the true positive lesion in comparison to the surrounding tissue most clearly. Our finding of significantly decreased ADC values in true positives agree with work by Stavrinides et al., who concluded that ADC could predict clinically-significant PCa in biopsy-naive men with indeterminate lesions [42]. Falaschi et al. similarly found lower ADC ratio in tumours than false positives, however they could not draw any firm conclusions about the usefulness of ADC ratio in detecting cancer [43]. These results demonstrate that quantitative diffusion MRI has the potential to reduce the number of unnecessary biopsies occurring in men with clinically insignificant disease.

Only the VERDICT $f_{IC}$ could discriminate false positives from normal tissue. Significantly higher values of $f_{IC}$ were found in false positives (p = 0.02), which agrees with histological findings of increased cellularity in the false positive disease types in comparison to healthy tissue [44]. The diagnostic accuracy of $f_{IC}$ (AUC = 0.7778) is the highest of all parameters, and the $f_{IC}$ parameter map also highlights the false positive lesion significantly. We did not find any significant differences in ADC values, potentially due to its limited biological specificity. Neither Falaschi et al. or Stavrinides et al. considered comparisons between these tissue types, nor did they investigate any other mathematical or biophysical models in their study. None of the IVIM or DKI parameters could perform this discrimination successfully either. For IVIM, this could be due to the fact it does not account for cellular geometry and compartmentalisation, leading to unreliable estimates of both diffusion components [14,45-47]. In the case of DKI, the model lacks specificity in a similar way to ADC, as it does not acknowledge the underlying microstructural features of the tissue [41]. These findings demonstrate the increased diagnostic accuracy of VERDICT, indicating its potential to identify unique diffusion signatures of diseases that mimic PCa.

This study demonstrates the utility of various diffusion MRI models for tackling the specific diagnostic obstacle of false positive cases of PCa. All diffusion models (ADC, IVIM, DKI and VERDICT) were able to discriminate false positives from true positives, however only VERDICT revealed significant differences between false positives and normal tissue. The VERDICT parameters also allow for inferences to be made about microstructural differences between tissue types, as shown by the correlation analysis. We observe strong negative correlation of $D_K$ and $D$ with $f_{IC}$, but positive correlation with $f_{EES}$. This is expected as diffusivities tend to decrease as the proportion of water trapped in cells increases. The DKI $K$ showed strong positive correlation with $f_{IC}$ and negative correlation with $f_{EES}$, which is also expected due to larger deviations from Gaussian dispersion as the proportion of water trapped in cells decreases. These observations emphasise VERDICT's improved biological specificity over simpler diffusion models, highlighting its utility in clinical practice to discriminate false positives.

The main limitation of this work is the number of participants – a larger patient cohort would allow us to improve the significance level of the results obtained, as well as potentially allowing for the identification of unique diffusion signatures for the different benign pathologies. However, we still achieve statistical significance for all the diffusion models. Also, this analysis was performed on retrospective data with an acquisition protocol optimised for VERDICT, which may mean that the choice of b-values was not optimal for IVIM and DKI parameter estimation. Furthermore, the range of benign pathologies in our study was limited to atrophy, inflammation and high-grade PIN, but the inclusion of others such as benign prostatic hyperplasia may allow for more comprehensive benign disease characterisation. Future work will increase the size of the patient cohort, encompassing a wider range of prostatic diseases.

## 5. Conclusions

In this work, we demonstrate that quantitative diffusion MRI (ADC, IVIM, DKI and VERDICT) can successfully discriminate false positive prostate lesions from cancerous tumours, showing potential to avoid unnecessary biopsies. The best diagnostic accuracy for discriminating false positives and true positives was observed for the DKI $D_K$, closely followed by the VERDICT $d_{EES}$. Among the different diffusion models only VERDICT was able to also differentiate false positive lesions from normal prostate tissue, correctly identifying benign diseases that mimic cancer.

**Author Contributions:** Conceptualization, Saurabh Singh and Eleftheria Panagiotaki; Data curation, Saurabh Singh; Formal analysis, Snigdha Sen; Funding acquisition, Eleftheria Panagiotaki; In-

vestigation, Snigdha Sen; Methodology, Snigdha Sen, Paddy Slator and Eleftheria Panagiotaki; Project administration, Snigdha Sen; Resources, Hayley Pye, Alistair Grey, Alex Freeman, Caroline Moore, Hayley Whitaker, Shonit Punwani and Saurabh Singh; Software, Snigdha Sen; Supervision, Saurabh Singh and Eleftheria Panagiotaki; Validation, Snigdha Sen; Visualization, Snigdha Sen; Writing – original draft, Snigdha Sen; Writing – review & editing, Snigdha Sen, Vanya Valindria, Paddy Slator, Hayley Pye, Alistair Grey, Alex Freeman, Caroline Moore, Hayley Whitaker, Shonit Punwani, Saurabh Singh and Eleftheria Panagiotaki.

**Funding:** This research was funded by EPSRC, grant numbers EP/N021967/1, EP/R006032/1, EP/V034537/1 and by Prostate Cancer UK: Targeted Call 2014: Translational Research St.2, grant number PG14-018-TR2.

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki and the UK Research Governance Framework version 2. INNOVATE received UK Research Ethics Committee approval on 23rd December 2015 by the NRES Committee London—Surrey Borders with REC reference 15/LO/0692.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Data is available on request.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Saltman A, Zegar J, Haj-Hamed M, Verma S, Sidana A. Prostate cancer biomarkers and multiparametric MRI: is there a role for both in prostate cancer management? Therapeutic Advances in Urology. 2021; 13, 1-11.
2. Lee SI, Hectors SJ. Prostate MRI: Toward Imaging Tumor Histology. Radiology. 2020; 296:2, 356-357.
3. Ahmed HU, El-Shater Bosaily A, Brown LC, Gabe R, Kaplan R, Parmar MK, et al. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. Lancet. 2017; 389, 815-822.
4. Graham J, Kirkbride P, Cann K, Hasler E, Prettyjohns M. Prostate cancer: summary of updated NICE guidance. BMJ. 2014; 348.
5. Kitzing YX, Prando A, Varol C, Karczmar GS, Maclean F, Oto A. Benign Conditions That Mimic Prostate Carcinoma: MR Imaging Features with Histopathologic Correlation. Radiographics. 2016; 36, 162-175.
6. Billis A. Prostatic atrophy. Clinicopathological significance. International Brazilian Journal of Urology. 2010; 36, 401-409.
7. Bostwick DG, Liu L, Brawer MK, Qian J. High-Grade Prostatic Intraepithelial Neoplasia. Reviews in Urology. 2004; 6, 171-179.
8. Chatterjee A, Thomas S, Oto A. Prostate MR: pitfalls and benign lesions. Abdominal Radiology. 2020; 45, 2154-2164.
9. Nickel JC. Prostatitis. Can Urol Assoc J. 2011; 5, 306-315.
10. Sato C, Naganawa S, Nakamura T, Kumada H, Miura S, Takizawa O, et al. Differentiation of noncancerous tissue and cancer lesions by apparent diffusion coefficient values in transition and peripheral zones of the prostate. Journal of Magnetic Resonance Imaging. 2005; 21, 258-262.
11. Lim HK, Kim JK, Kim KA, Cho KS. Prostate cancer: apparent diffusion coefficient map with T2-weighted images for detection—a multireader study. Radiology. 2009; 250, 145-151.
12. Bourne R, Panagiotaki E. Limitations and Prospects for Diffusion-Weighted MRI of the Prostate. Diagnostics. 2016; 6, 21.
13. Gibbs P, Pickles MD, Turnbull LW. Diffusion imaging of the prostate at 3.0 tesla. Investigative Radiology. 2006; 41, 185-188.
14. Kim CK, Park BK, Lee HM, Kwon GY. Value of diffusion-weighted Imaging for the prediction of prostate cancer location at 3 T using a phased-array coil—preliminary results. Investigative Radiology. 2007; 42, 842-847.
15. Koh DM, Collins DJ, Orton MR. Intravoxel incoherent motion in body diffusion-weighted MRI: reality and challenges. Americal Journal of Roentgenology. 2011; 196, 1351-1361.
16. Sigmund EE, Cho GY, Finn M, Moccaldi M, Jensen JH, Sodickson DK, et al. Intravoxel incoherent motion imaging of tumor microenvironment in locally advanced breast cancer. Magnetic Resononance in Medicine. 2011; 65, 1437-1447.
17. Riches S, Hawtin K, Charles-Edwards E, et al. Diffusion-weighted imaging of the prostate and rectal wall: comparison of biexponential and monoexponential modelled diffusion and associated perfusion coefficients. NMR in Biomedicine. 2009; 22, 318-325.
18. Klauss M, Lemke A, Grunberg K, Simon D, Re TJ, Wente MN, et al. Intravoxel incoherent motion MRI for the differentiation between mass forming chronic pancreatitis and pancreatic carcinoma. Investigative Radiology. 2011; 46, 57-63.
19. Rosenkrantz AB, Sigmund EE, Johnson G, Babb JS, Mussi TC, Melamed J, et al. Prostate cancer: feasibility and preliminary experience of a diffusional kurtosis model for detection and assessment of aggressiveness of peripheral zone cancer. Radiology. 2012; 264, 126-135.
20. Liang S, Panagiotaki E, Bongers A, Shi P, Sved P, Watson G, et al. Information-based ranking of 10 compartment models of diffusion-weighted signal attenuation in fixed prostate tissue. NMR in Biomedicine. 2016; 29, 660-671.
21. Jensen JH, Helpern JA, Ramani A, Lu H, Kaczynski K. Diffusional kurtosis imaging: the quantification of non-Gaussian water diffusion by means of magnetic resonance imaging. Magnetic Resonance in Medicine. 2005; 53, 1432-1440.

22. Panagiotaki E, Walker-Samuel S, Siow B, Johnson SP, Rajkumar V, Pedley RB, et al. Non-invasive quantification of solid tumour microstructure using VERDICT MRI. Cancer Research. 2014; 74, 1902-1912.
23. Panagiotaki E, Chan RW, Dikaios N, Ahmed HU, O'Callaghan J, Freeman A, et al. Microstructural Characterization of Normal and Malignant Human Prostate Tissue With Vascular, Extracellular, and Restricted Diffusion for Cytometry in Tumours Magnetic Resonance Imaging. Investigative Radiology. 2015; 50, 218-227.
24. Johnston E, Pye H, Bonet-Carne E, Panagiotaki E, Patel D, Galazi M, et al. INNOVATE: A prospective cohort study combining serum and urinary biomarkers with novel diffusion-weighted magnetic resonance imaging for the prediction and characterization of prostate cancer. BMC Cancer. 2016; 16, 816.
25. Johnston EW, Bonet-Carne E, Ferizi U, Yvernault B, Pye H, Patel D, et al. VERDICT MRI for Prostate Cancer: Intracellular Volume Fraction versus Apparent Diffusion Coefficient. Radiology. 2019; 291, 391-397.
26. Valindria V, Palombo M, Chiou E, Singh S, Punwani S, Panagiotaki E. Synthetic Q-Space Learning with Deep Regression Networks for Prostate Cancer Characterisation with VERDICT. In International Symposium on Biomedical Imaging (ISBI); 2021.
27. Turkbey B, Rosenkrantz AB, Haider MA, Padhani AR, Villeirs G, Macura KJ, et al. Prostate Imaging Reporting and Data System Version 2.1: 2019 Update of Prostate Imaging Reporting and Data System Version 2. European Urology. 2019; 76, 340-351
28. Harada T, Abe T, Kato F, Matsumoto R, Fujita H, Murai S, et al. Five-point Likert scaling on MRI predicts clinically significant prostate carcinoma. BMC Urology. 2015; 15, 91.
29. Scheenen TWJ, Rosenkrantz AB, Haider MA, Futterer JJ. Multiparametric Magnetic Resonance Imaging in Prostate Cancer Management: Current Status and Future Perspectives. Investigative Radiology. 2015; 50, 594-600.
30. Cook PA, Bai Y, Nedjati-Gilani S, Seunarine KK, Hall MG, Parker GJ, et al. Camino: Open-Source Diffusion-MRI Reconstruction and Processing. In 14th Scientific Meeting of the International Society for Magnetic Resonance in Medicine (ISMRM); 2006; Seattle, WA, USA. p. 2759.
31. Panagiotaki E, Ianus A, Johnston E, Chan RW, Stevens N, Atkinson D, et al. Optimised VERDICT MRI protocol for prostate cancer characterisation. In International Society for Magentic Resonance in Medicine (ISMRM); 2015.
32. Veraart J, Fieremans E, Novikov DS. Diffusion MRI noise mapping using random matrix theory. Magnetic Resonance in Medicine. 2015; 76, 1582-1593.
33. Tournier JD, Smith R, Raffelt D, Tabbara R, Dhollander T, Peitsch M, et al. MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. NeuroImage. 2019; 202, 116137.
34. Kellner E, Dhital B, Kiselev VG, Reisert M. Gibbs-ringing artifact removal based on local subvoxel-shifts. Magnetic Resonance in Medicine. 2015; 76, 1574-1581.
35. Palombo M, Singh S, Whitaker H, Punwani S, Alexander DC, Panagiotaki E. Relaxed-VERDICT: decoupling relaxation and diffusion for comprehensive microstructure characterization of prostate cancer. In International Society for Magnetic Resonance in Medicine (ISMRM); 2020.
36. Le Bihan D, Breton E, Lallemand D, Aubin ML, Vignaud J, Laval-Jeantet M. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. Radiology. 1988; 168, 497-505.
37. Henriques RN, Jespersen SN, Shemesh N. Microscopic anisotropy misestimation in spherical-mean single diffusion encoding MRI. Magnetic Resonance in Medicine. 2019; 81, 3245-3261.
38. Bonet-Carne E, Johnston E, Daducci A, Jacobs JG, Freeman A, Atkinson D, et al. VERDICT-AMICO: Ultrafast fitting algorithm of non-invasive prostate microstructure characterization. NMR in Biomedicine. 2019; 32, 4019.
39. Valindria V, Singh S, Palombo M, Chiou E, Mertzanidou T, Kanber B, et al. Non-invasive Gleason Score Classification with VERDICT-MRI. In International Society for Magnetic Resonance in Medicine (ISMRM); 2021.
40. Virtanen P, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods. 2020; 17, 261-272.
41. de la Rosette JJMCH, Manyak MJ, Harisinghani MG, Wijkstra H. Imaging in Oncological Urology. London, UK: Springer; 2009.
42. Stavrinides V, Syer T, Hu Y, Giganti F, Freeman A, Karapanagiotis S, et al. False Positive Multiparametric Magnetic Resonance Imaging Phenotypes in the Biopsy-naïve Prostate: Are They Distinct from Significant Cancer-associated Lesions? Lessons from PROMIS. European Urology. 2021; 79, 20-29.
43. Falaschi Z, Venti M, Lanzo G, Attansio S, Valentini E, Garcia Navarro LI, et al. Accuracy of ADC ratio in discriminating true and false positives in multiparametric prostatic MRI. European Journal of Radiology. 2020; 128, 109024.
44. Qupn JS, Moosavi B, Khanna M, Flood TA, Lim CS, Schieda N. False positive and false negative diagnoses of prostate cancer at multi-parametric prostate MRI in active surveillance. Insights Imaging. 2015; 6, 449-463.
45. Mazaheri Y, Vargas HA, Akin O, Goldman DA, Hricak H. Reducing the influence of b-value selection on diffusion-weighted imaging of the prostate: evaluation of a revised monoexponential model within a clinical setting. Journal of Magnetic Resonance Imaging. 2012; 35, 660-668.
46. Dopfert J, Lemke A, Weidner A, Schad LR. Investigation of prostate cancer using diffusion-weighted intravoxel incoherent motion imaging. Magnetic Resonance Imaging. 2011; 29., 1053-1058.
47. Pang Y, Turkbey B, Bernardo M, Kruecker J, Kadoury S, Merino MJ, et al. Intravoxel incoherent motion MR imaging for prostate cancer: An evaluation of perfusion fraction and diffusion coefficient derived from different b-value combinations. Magnetic Resonance in Medicine. 2013; 69, 553-562.