





# ssVERDICT: Self-supervised VERDICT-MRI for enhanced prostate tumor characterization

Snigdha Sen<sup>1</sup>   | Saurabh Singh<sup>2</sup> | Hayley Pye<sup>3</sup> | Caroline M. Moore<sup>3</sup> | Hayley C. Whitaker<sup>3</sup> | Shonit Punwani<sup>2</sup> | David Atkinson<sup>2</sup>  | Eleftheria Panagiotaki<sup>1</sup>  | Paddy J. Slator<sup>1,4,5</sup>

<sup>1</sup>Center for Medical Image Computing, Department of Computer Science, University College London, London, UK

<sup>2</sup>Center for Medical Imaging, Division of Medicine, University College London, London, UK

<sup>3</sup>Department of Targeted Intervention, Division of Surgery and Interventional Science, University College London, London, UK

<sup>4</sup>Cardiff University Brain Research Imaging Center, School of Psychology, Cardiff University, Cardiff, UK

<sup>5</sup>School of Computer Science and Informatics, Cardiff University, Cardiff, UK

## Correspondence

Snigdha Sen, Center for Medical Image Computing, Department of Computer Science, University College London, London, UK.

Email: [snigdha.sen.20@ucl.ac.uk](mailto:snigdha.sen.20@ucl.ac.uk)

## Funding information

Prostate Cancer UK, Grant/Award Number: PG14-018-TR2; Engineering and Physical Sciences Research Council, Grant/Award Numbers: EP/N021967/1, EP/R006032/1, EP/S021930/1, EP/V034537/1; University College London Hospitals

## Abstract

**Purpose:** Demonstrating and assessing self-supervised machine-learning fitting of the VERDICT (vascular, extracellular and restricted diffusion for cytometry in tumors) model for prostate cancer.

**Methods:** We derive a self-supervised neural network for fitting VERDICT (ssVERDICT) that estimates parameter maps without training data. We compare the performance of ssVERDICT to two established baseline methods for fitting diffusion MRI models: conventional nonlinear least squares and supervised deep learning. We do this quantitatively on simulated data by comparing the Pearson's correlation coefficient, mean-squared error, bias, and variance with respect to the simulated ground truth. We also calculate in vivo parameter maps on a cohort of 20 prostate cancer patients and compare the methods' performance in discriminating benign from cancerous tissue via Wilcoxon's signed-rank test.

**Results:** In simulations, ssVERDICT outperforms the baseline methods (non-linear least squares and supervised deep learning) in estimating all the parameters from the VERDICT prostate model in terms of Pearson's correlation coefficient, bias, and mean-squared error. In vivo, ssVERDICT shows stronger lesion conspicuity across all parameter maps, and improves discrimination between benign and cancerous tissue over the baseline methods.

**Conclusion:** ssVERDICT significantly outperforms state-of-the-art methods for VERDICT model fitting and shows, for the first time, fitting of a detailed multi-compartment biophysical diffusion MRI model with machine learning without the requirement of explicit training labels.

## KEYWORDS

diffusion MRI, microstructure estimation, prostate cancer, self-supervised learning

Eleftheria Panagiotaki and Paddy J. Slator contributed equally to this work.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Author(s). *Magnetic Resonance in Medicine* published by Wiley Periodicals LLC on behalf of International Society for Magnetic Resonance in Medicine.

## 1 | INTRODUCTION

Prostate cancer (PCa) characterization is reliant on invasive biopsy, but in recent years, multiparametric MRI (mp-MRI) has become established in the diagnostic pathway for localization and staging of clinically significant PCa (csPCa).<sup>1</sup> Diffusion MRI (dMRI) is a powerful component of mp-MRI, measuring the motion of water molecules in biological tissues to infer information about local microstructure. Advanced multi-compartment models of the dMRI signal can estimate parameters relating to specific microstructural properties such as cell size, density, and vasculature.<sup>2</sup> Such models enable noninvasive analysis of similar metrics to those typically only accessible by histology and have been shown to reduce the need for invasive biopsies in breast<sup>3</sup> and PCa.<sup>4</sup>

Microstructural information can be extracted by designing models with parameters corresponding to these biophysically relevant metrics, which are then estimated by fitting the models to dMRI data. These models tend to be nonlinear with many free parameters, and dense  $q$ -space sampling is required for accurate description of the microstructure, which involves time-consuming processes. This means that parameter estimation becomes a difficult inverse problem, scaling with both voxel number and model complexity. Additionally, parameter estimation requires an optimization-based procedure, typically relying on nonlinear least squares (NLLS) curve fitting, which is computationally expensive and prone to estimation errors.<sup>5</sup> These challenges when obtaining and examining dMRI data hinder the clinical translation of these methods.

Recent work has used deep learning (DL) techniques to solve this parameter-estimation inverse problem. These algorithms learn the mapping between the  $q$ -space data and the microstructural parameters of the dMRI model. Pioneering work on  $q$ -space learning by Golkov et al.<sup>6</sup> estimated model parameters using a multilayer perceptron (MLP), an approach that has since been widely used for ultrafast model fitting.<sup>7-9</sup> Convolutional neural networks (CNNs) have also been used with supervised learning for dMRI model fitting,<sup>10</sup> as have transformers,<sup>11</sup> but these methods require large amounts of training data. Supervised DL approaches have been used to fit both simple exponential models as well as complex biophysical models, such as NODDI (neurite orientation dispersion and density imaging)<sup>12</sup> and the spherical mean technique.<sup>13</sup> However, supervised methods are significantly affected by the underlying distribution of the training data, which can introduce biases in the parameter estimates.<sup>5,13</sup>

Another branch of deep learning is referred to as self-supervised learning. These methods require the neural network to predict one part of the input data given another part,<sup>14</sup> removing the requirement for explicitly labeled training data and thus the risk of introducing bias to the estimates. This approach has been successful for microstructural parameter estimation with the intravoxel incoherent motion (IVIM) model.<sup>5,15-19</sup> However, despite these numerous IVIM examples, and in contrast to supervised model fitting, self-supervised model fitting has not been demonstrated for detailed multicompartment biophysical models.

Here we introduce a self-supervised approach to fit the vascular, extracellular, and restricted diffusion for cytometry in tumors (VERDICT) model for prostate: a multi-compartment biophysical dMRI model.<sup>20</sup> We refer to our method as ssVERDICT. The VERDICT framework, currently in clinical trials applied to PCa, requires robust model fitting to estimate microstructural metrics such as cell size, intracellular volume fraction, and diffusivity. These have previously been estimated via NLLS and supervised DL approaches,<sup>21-23</sup> but the complexity of VERDICT increases its susceptibility to the aforementioned limitations of these techniques.

This is the first work demonstrating self-supervised fitting of diffusion MRI models beyond simple exponential models. We show that ssVERDICT achieves higher accuracy and reduced bias when estimating microstructural parameters using ground-truth simulations. On real data, ssVERDICT achieves discrimination of cancerous tissue from benign at a higher confidence level on a data set of 20 PCa patients, highlighting the potential of the method for clinical translation.

## 2 | METHODS

In this section, we first introduce the VERDICT model for prostate, a three-compartment biophysical dMRI model. We then outline how the simulated data was generated and how the patient data was acquired. We discuss the two baseline fitting methods (conventional NLLS fitting and supervised deep learning), followed by our novel self-supervised fitting method, ssVERDICT. Finally, we give details on the preprocessing steps, region-of-interest (ROI) selection, and the evaluation metrics used.

### 2.1 | VERDICT model

The VERDICT prostate model is the sum of three parametric models describing the dMRI signal as arising from

intracellular (IC), extracellular-extravascular (EES), and vascular (VASC) water populations, as originally introduced in Panagiotaki et al.<sup>20</sup> The total normalized signal is calculated as

$$\frac{S}{S_0} = f_{\text{VASC}} S_{\text{VASC}}(d_{\text{VASC}}, b) + f_{\text{IC}} S_{\text{IC}}(d_{\text{IC}}, R, b, \Delta, \delta) + f_{\text{EES}} S_{\text{EES}}(d_{\text{EES}}, b), \quad (1)$$

where  $f_i$  is the volume fraction;  $S_i$  is the normalized signal from water molecules in population  $i$  (where  $i = \text{IC}, \text{VASC}, \text{or EES}$ ); and  $R$  represents the cell radius. The vascular signal fraction,  $f_{\text{VASC}}$ , is computed as  $1 - f_{\text{IC}} - f_{\text{EES}}$ , since  $\sum_{i=1}^3 f_i = 1$  and  $0 \leq f_i \leq 1$ , and  $S_0$  is the signal with no diffusion weighting.<sup>20</sup> The contribution of each compartment to the overall signal can be assessed via the volume fractions. Here,  $b$  is the  $b$ -value;  $\Delta$  is the gradient pulse separation; and  $\delta$  is the gradient pulse duration.

In this paper, we use a modified version of the original VERDICT model first introduced in Panagiotaki et al.,<sup>22</sup> where the mathematical signal forms are as follows:

$$S_{\text{VASC}} = \frac{\sqrt{\pi}}{2} \times \frac{\phi\left(\sqrt{b(-(\Delta - \delta/3)(\gamma\delta)^2 d_{\text{VASC}})}\right)}{\sqrt{b(-(\Delta - \delta/3)(\gamma\delta)^2 d_{\text{VASC}})}}, \quad (2)$$

$$S_{\text{IC}} = \exp\left(-\frac{2\gamma^2 G^2}{d_{\text{IC}}} \sum_{m=1}^{\infty} \frac{\alpha_m^4}{\alpha_m^2 R^2 - 2} \left[2\delta - \frac{2 + e^{-\alpha_m^2 d_{\text{IC}}(\Delta - \delta)} - 2e^{-\alpha_m^2 d_{\text{IC}}\delta} - 2e^{-\alpha_m^2 d_{\text{IC}}\Delta} + 2e^{-\alpha_m^2 d_{\text{IC}}(\Delta + \delta)}}{\alpha_m^2 d_{\text{IC}}}\right]\right), \quad (3)$$

$$S_{\text{EES}} = e^{-bd_{\text{EES}}}, \quad (4)$$

where  $\phi$  is the error function  $\phi(z) = \int_0^z \exp(-t^2) dt$ , and  $\alpha_m^2$  is the  $m^{\text{th}}$  root of  $(\alpha R)^{-1} J_{\frac{3}{2}}(\alpha R) = J_{\frac{5}{2}}$ , where  $J_n(x)$  is a Bessel function of the first kind.<sup>2,22,24</sup>

The spherical mean version of the VERDICT model represents the IC component as restricted diffusion within spheres of radius  $R$  (using the Gaussian phase distribution approximation<sup>25</sup>) with intrasphere diffusivity fixed at  $d_{\text{IC}} = 2 \mu\text{m}^2/\text{ms}$ . The EES component is represented as Gaussian isotropic diffusion with effective diffusivity  $d_{\text{EES}}$ , and the vascular component as isotropically restricted diffusion within spherically-averaged randomly oriented sticks with intrastick diffusivity fixed at  $d_{\text{VASC}} = 8 \mu\text{m}^2/\text{ms}$  as in Refs. 22 and 24. The fixed values are chosen in Panagiotaki et al.<sup>22</sup> via preliminary work on model selection to minimize fitting error, and here we use the VERDICT model with free  $d_{\text{EES}}$ , as in Bonet-Carne et al.,<sup>24</sup> which can be informative for describing cancer tissue.<sup>26</sup> By fitting the

model to dMRI data, we estimate four model parameters:  $f_{\text{EES}}, f_{\text{IC}}, R$ , and  $d_{\text{EES}}$ .

## 2.2 | Patient data

The study was performed with the approval of the local ethics committee embedded within the INNOVATE clinical trial (NCT02689271),<sup>27</sup> which included men suspected of having csPCa. For this study, we randomly selected 20 patients from the INNOVATE cohort with biopsy-confirmed csPCa. This sample size was chosen based on previous work,<sup>15</sup> allowing us to provide an initial demonstration of the fitting technique on a range of Gleason grades while optimizing radiologists' time. VERDICT-MRI was performed on a 3T MRI system (Achieva; Philips, Best, the Netherlands) using a pulsed-gradient spin-echo sequence. The imaging parameters, as published in Refs. 4 and 26–28 were as follows: TR = 2482–3945 ms, FOV = 200 × 220 mm, voxel size = 1.3 × 1.3 × 5 mm, no interslice gap, acquisition matrix = 176 × 176. The optimized VERDICT acquisition protocol for prostate is:  $b = 90, 500, 1500, 2000$ , and  $3000 \text{ s/mm}^2$ ;  $\delta = 3.9, 11.4, 23.9, 14.4$ , and  $18.9 \text{ ms}$ ; and  $\Delta = 23.8, 31.3, 43.8, 34.3$ , and  $38.8 \text{ ms}$ <sup>28</sup> (with some slight variation in  $\delta/\Delta$  due to machine settings). Images were acquired with diffusion weighting in three orthogonal directions. Six repeats were acquired per  $b$ -value in each direction (except for  $b = 90 \text{ s/mm}^2$ , for which four repeats were acquired). For each of the five combinations of  $b/\delta/\Delta$ , we used the minimum possible TE, giving TEs of 50–90 ms, and a  $b = 0$  image was acquired with each of the five datasets. The three directions were averaged for each  $b$ -value, resulting in 10 image volumes.

## 2.3 | Simulated data

We generated synthetic datasets for quantitative analysis using the VERDICT model with added Rician noise. We first simulated datasets with SNR levels ranging from 10 to 100, so we could test the robustness of the methods to noise. We then set SNR = 50 for the final simulated data set. We created 100 000 signals from uniform VERDICT parameter distributions within biophysically realistic parameter ranges:  $f_{\text{EES}} = [0.01, 0.99]$ ,  $f_{\text{IC}} = [0.01, 0.99]$ ,  $R = [0.01, 15] \mu\text{m}$ , and  $d_{\text{EES}} = [0.5, 3] \mu\text{m}^2/\text{ms}$ . We calculated  $f_{\text{VASC}} = 1 - f_{\text{IC}} - f_{\text{EES}}$  as before, with  $f_{\text{VASC}} < 0$  set to be equal to 0. We normalized each volume fraction by dividing by  $f_{\text{VASC}} + f_{\text{IC}} + f_{\text{EES}}$ , to scale them proportionally while ensuring they sum to one. We simulated dMRI data using the same acquisition protocol as the patient data.<sup>28</sup> The parameters were drawn from uniform (rather than

in vivo) distributions to minimize bias in the resulting parameter estimates.<sup>5,13</sup>

## 2.4 | Conventional iterative fitting

We fit the VERDICT model via NLLS using custom code in *MATLAB* (The MathWorks Inc., Natick, Massachusetts, USA). This used the “lsqcurvefit” function to implement the Levenberg–Marquardt algorithm as in Refs. 20 and 22 with parameter constraints as given in Section 2.3. Prediction for the whole unmasked dMRI data set (approximately  $5 \times 10^5$  voxels) took approximately 140 s per subject (Apple M1 Pro).

## 2.5 | Supervised deep learning

Supervised techniques approximate the function  $f$  that maps the measurement  $\mathbf{S}$  to its corresponding ground-truth parameters,  $\mathbf{x}$ , by minimizing the difference between the ground-truth parameter values (training labels) and the parameter estimates (network output). The training loss is calculated as the mean squared error (MSE) between the estimated and ground-truth values. We use an MLP architecture, implemented in *Python* 3.7.13 using the “MLPRegressor” in scikit-learn 0.23, to compare with Refs 21, 26, 29, and 30.

The input of the deep neural network (DNN) is a vector of dMRI signals for each combination of  $b$ ,  $\Delta$ ,  $\delta$  (a total of 10 in this case), followed by three fully-connected hidden layers, each with 150 neurons,<sup>21,26,29,30</sup> and a final regression layer with four output neurons (equal to the number of parameters to be estimated). The DNN is trained on 100 000 synthetic signals (split into 80% for training and 20% for validation), with values for the model parameters randomly chosen from the ranges given in Section 2.3. We performed the optimization with the ADAM method for 1000 epochs (adaptive learning rate with initial value of 0.001; one update per minibatch of 100 voxels; early stopping to mitigate overfitting (as implemented in scikit-learn); and momentum = 0.9). For the final parameter computation, we used the DNN at the epoch with minimum validation loss. The creation of the training set and training of the DNN (which was performed only once) took roughly 200 s. Prediction of the trained DNN for the whole unmasked dMRI data set took approximately 30 s per subject.

## 2.6 | Self-supervised deep learning

Self-supervised methods compute  $f$  by minimizing the difference between the noisy MR signals (network inputs)

and noise-free signal estimates reconstructed from the estimated parameters (network outputs). The training loss is equivalent to the MSE between the predicted signal,  $\hat{S}$ , and the input signal  $S$ .<sup>15</sup> Here, network training and inference is performed on the same data set, mimicking the NLLS approach.

We implemented a fully connected neural network with three hidden layers, each with 10 neurons (equal to the number of image volumes), using *PyTorch* 1.12.1. The output layer is fed into the VERDICT model equation to generate the predicted signal  $\hat{S}$ . Crucially, this requires coding the VERDICT model in *PyTorch* in a differentiable form. For this, we formulate the intricate signal equations for VERDICT’s “sphere” and “astrosticks” compartments (Eqs. [2] and [3]) as *PyTorch* tensor functions, so that multidimensional tensors of batched parameter values can be inputted to yield output tensors of batched predicted signals. A schematic of ssVERDICT is given in Figure 1.

For the final parameter estimation, we used the normalized input data, the ADAM optimizer, and the DNN at the epoch with minimum validation loss. We optimized the DNN by backpropagating the MSE between  $S$  and  $\hat{S}$ , where  $\hat{S}$  is reconstructed via the VERDICT model from the parameter estimates. We chose a learning rate of 0.0001 and the network was trained until 10 consecutive epochs occurred without any improvement in loss, before terminating to prevent overtraining. We used dropout ( $p = 0.5$ ) to prevent overfitting and constrained the parameter values to the ranges in Section 2.3 using the *PyTorch* clamp function. Training and prediction for the whole unmasked dMRI data set took about 50 s per subject.

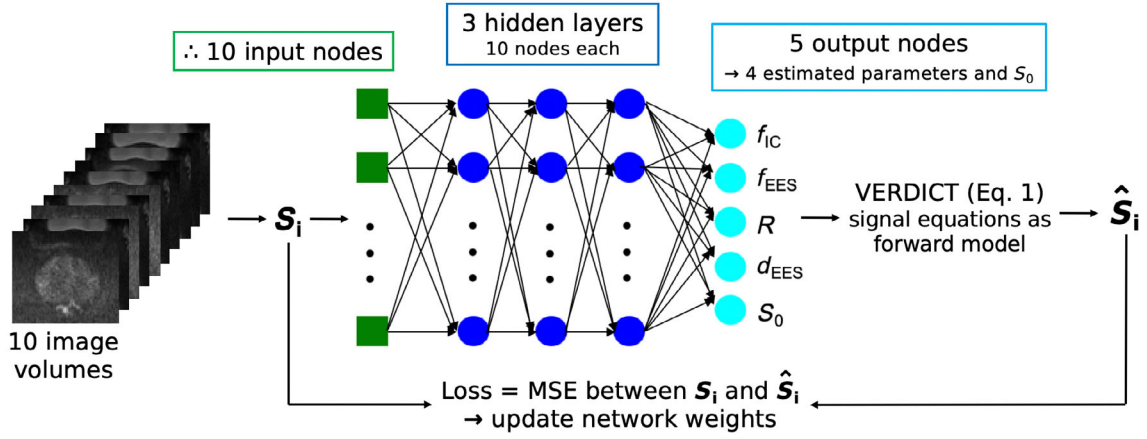
## 2.7 | Data preprocessing

The preprocessing pipeline included denoising of the raw DW-MRI using MP-PCA<sup>31</sup> as implemented within MrTrix3<sup>32</sup> “dwidenoise,” and then correction for Gibbs ringing<sup>33</sup> with custom code in *MATLAB*, as done in Refs. 4, 21, 26, and 27. To reduce potential artifacts caused by patient movement during scanning and eddy current distortions, we applied mutual-information rigid and affine registration using custom code in *MATLAB*.<sup>34</sup> We normalized the data by dividing the dMRI volumes by their matched  $b = 0$ . As we use the spherical mean version of the VERDICT model, we spherically averaged the data across the three orthogonal directions to produce 10 image volumes, where each volume was a 3D image consisting of 14 slices.

## 2.8 | ROIs

Patients were biopsied based on their mp-MRI score as reported by two board-certified experienced





**FIGURE 1** Schematic of our self-supervised network. The input to the neural network is the signal extracted from 10 signal volumes; therefore, there are 10 input nodes. The network has three hidden layers, each with 10 nodes. The final layer has five nodes, corresponding to the four estimated vascular, extracellular, and restricted diffusion for cytometry in tumors (VERDICT) parameters and  $S_0$ , the signal with no diffusion weighting. To reconstruct the signal ( $\hat{S}$ ), the complex VERDICT signal equations (Eqs. [1]–[4]) are written in differentiable form so that it can be incorporated as a layer in the network, such that batches of signals can be inputted and batches of parameters outputted.

uroradiologists (over 5 years of reporting experience at a specialist center and reporting more than 2000 prostate MR scans per year). The ROIs were drawn by a board-certified study radiologist (S. Singh) using a pictorial report made by the clinical uroradiologist and confirmed as cancerous retrospectively via targeted biopsy. Twenty patients with positive biopsies with a range of Gleason grades were selected for inclusion in the study. For each of the 20 patients, an additional ROI was located in an area of benign tissue to be used for comparison, after a review of the sampling biopsy result confirmed the absence of tumor on the contralateral side.

## 2.9 | Evaluation metrics

We quantitatively compared the performance of the three parameter estimation methods via a variety of evaluation metrics: (1) Pearson's correlation coefficient, (2) MSE, (3) bias, and (4) variance, all with respect to ground-truth parameter values used for the simulated data. The formulas for the metrics used are as follows:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (O_i - E_i)^2, \quad (5)$$

$$\text{Bias} = \frac{1}{N} \sum_{i=1}^N (O_i - E_i), \quad (6)$$

$$\text{Variance} = \frac{1}{N} \sum_{i=1}^N (E_i - \bar{E})^2, \quad (7)$$

where  $O$  is the ground-truth parameter value;  $E$  is the estimated value ( $\bar{E}$  is the mean); and  $N$  is the number of samples.

We discriminated between cancerous and benign tissue *in vivo* by comparing parameter values in respective ROIs using the Wilcoxon's signed-rank test (preceded by the Shapiro–Wilk test for normality).

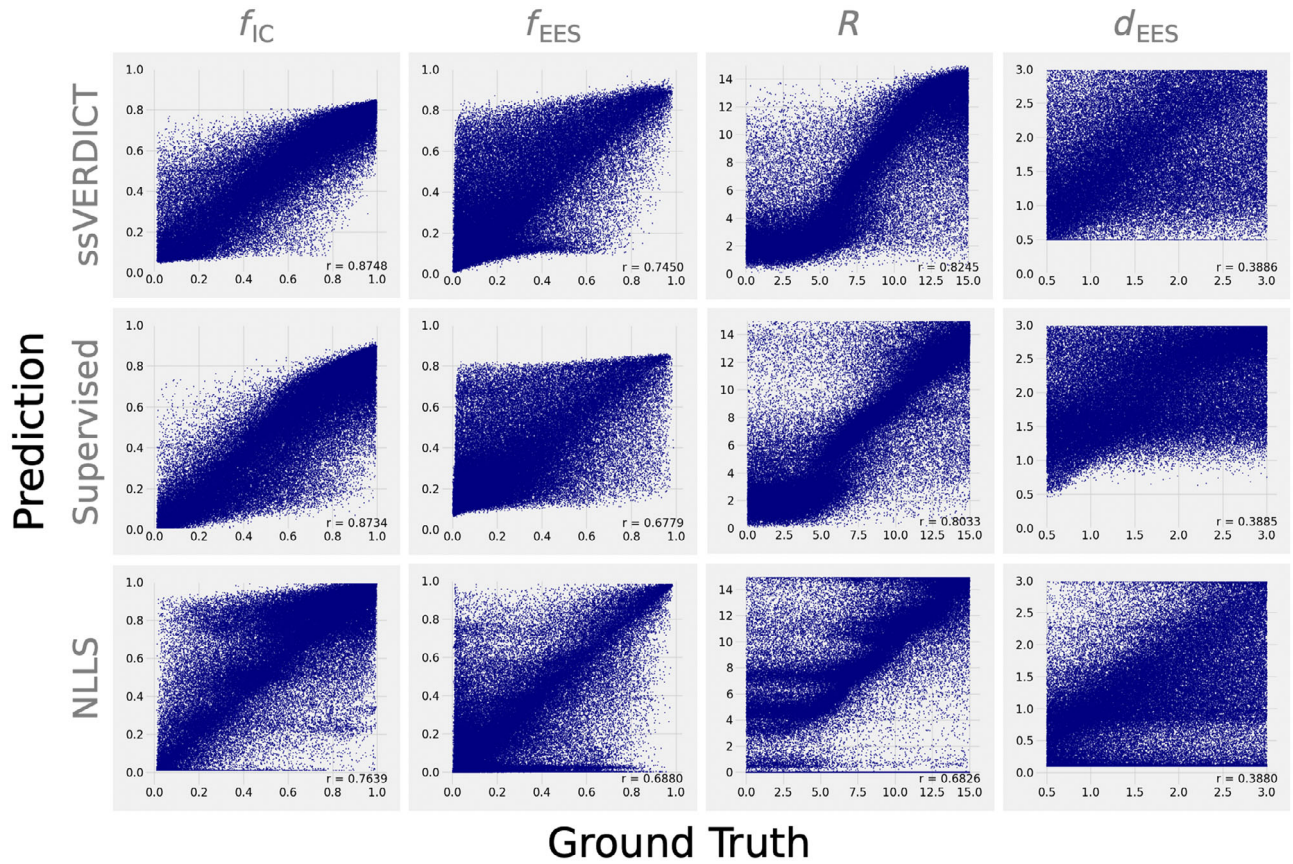
## 3 | RESULTS

Figure 2 shows estimated VERDICT parameters via each fitting method plotted against randomly generated ground-truth parameter values (Section 2.3). The Pearson's correlation coefficients  $r$  are highest for all four VERDICT parameters when fitted via ssVERDICT. We also observe higher  $r$  values for supervised DL fitting over NLLS.

Table 1 gives the MSE, bias, and variance values for all four fitted parameters with each of the fitting methods. We observe lower bias and MSE across all parameters via ssVERDICT, and lower variance in estimating  $f_{EES}$  and  $d_{EES}$ . However, supervised DL fitting achieves lowest variance in estimating  $f_{IC}$  and  $R$ .

Figure 3 shows *in vivo* maps of the four fitted VERDICT parameters and calculated  $f_{VASC}$ . ssVERDICT shows strong lesion conspicuity for the  $f_{IC}$  and  $f_{EES}$  maps, and reasonable conspicuity on the  $f_{VASC}$ ,  $R$ , and  $d_{EES}$  maps. The supervised DL method achieves strong lesion conspicuity for  $f_{EES}$  and  $f_{VASC}$ , and the NLLS method for  $f_{IC}$  and  $f_{EES}$ . In Figure S1, we show these trends continue for patients with different Gleason grades, and that the maps can aid in discriminating cancer grades.

Figure 4 shows boxplots of the fitted VERDICT parameters in benign and cancerous prostate tissue for a data set of 20 patients. All three methods can discriminate among tissue types to a high level of significance with  $f_{IC}$  and  $f_{EES}$ . ssVERDICT increases discrimination between benign and

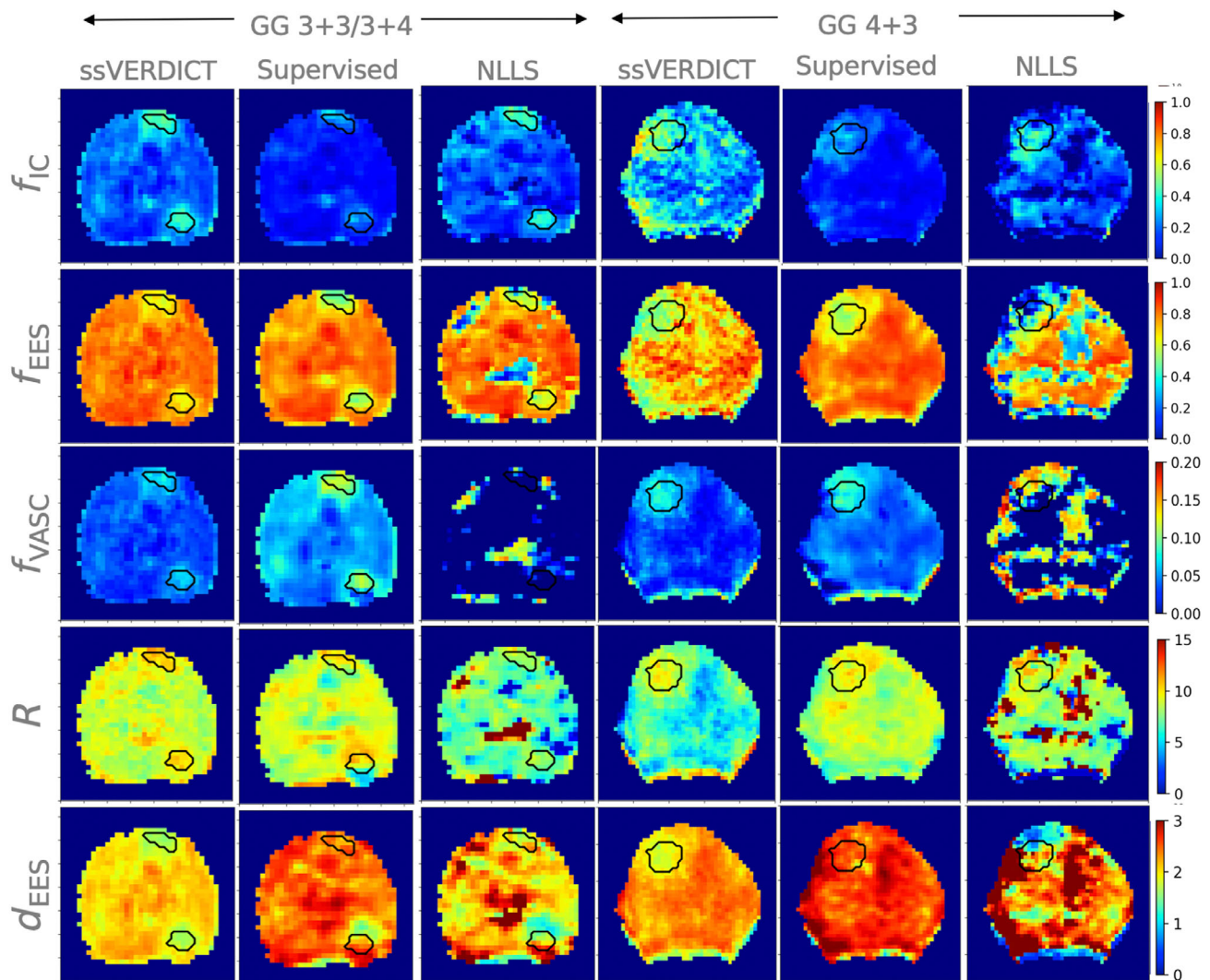


**FIGURE 2** Scatterplot of simulated ground-truth parameter values and predicted values via the three fitting methods, with a signal-to-noise ratio (SNR) of 50. We observe higher Pearson's correlation coefficient  $r$  when using self-supervised neural network for fitting VERDICT (vascular, extracellular and restricted diffusion for cytometry in tumors) (ssVERDICT) for all four estimated parameters. EES, extracellular-extravascular; IC, intracellular; NLLS, nonlinear least squares.

**TABLE 1** Mean-squared error (MSE), bias, and variance values calculated between simulated ground truth and predictions obtained via each fitting method, with the best performing method highlighted in bold. We find self-supervised neural network for fitting VERDICT (ssVERDICT) achieves the lowest MSE and bias across all four parameters, and lowest variance for  $f_{EES}$  and  $d_{EES}$ .

<b>MSE</b>				
Method	$f_{IC}$	$f_{EES}$	$R$ ( $\mu m$ )	$d_{EES}$ ( $\mu m^2/ms$ )
NLLS	0.1232	0.1137	17.6976	0.9905
Supervised DL	0.0714	0.0994	6.8860	0.7489
ssVERDICT	<b>0.0289</b>	<b>0.0362</b>	<b>5.5278</b>	<b>0.7160</b>
<b>Bias</b>				
Method	$f_{IC}$	$f_{EES}$	$R$ ( $\mu m$ )	$d_{EES}$ ( $\mu m^2/ms$ )
NLLS	-0.0742	0.0680	-0.9442	-0.3624
Supervised DL	-0.1008	0.1571	-0.7152	0.3994
ssVERDICT	<b>-0.0070</b>	<b>-0.0162</b>	<b>0.4002</b>	<b>0.2522</b>
<b>Variance</b>				
Method	$f_{IC}$	$f_{EES}$	$R$ ( $\mu m$ )	$d_{EES}$ ( $\mu m^2/ms$ )
NLLS	0.0958	0.0942	17.5022	0.8649
Supervised DL	<b>0.0459</b>	0.0627	<b>13.3562</b>	0.6388
ssVERDICT	0.0655	<b>0.0542</b>	16.7244	<b>0.4378</b>

Abbreviations: DL, deep learning; NLLS, nonlinear least squares; VERDICT, vascular, extracellular and restricted diffusion for cytometry in tumors.



**FIGURE 3** Parameter maps of the four fitted vascular, extracellular and restricted diffusion for cytometry in tumors (VERDICT) parameters and calculated  $f_{VASC}$  for 2 patients. Data set 1 shows a Gleason 3 + 3 grade tumor in the left anterior and 3 + 4 grade tumor in the right posterior peripheral zone, and data set 2 shows a Gleason 4 + 3 grade tumor in the right peripheral zone. We observe improved lesion conspicuity overall when using self-supervised neural network for fitting VERDICT (ssVERDICT), whereas supervised deep learning only shows strong tumor conspicuity for  $f_{EES}$  and  $f_{VASC}$ , and nonlinear least squares (NLLS) only for  $f_{IC}$  and  $f_{EES}$ . We observe greater variations in the NLLS maps due to the fitting technique's susceptibility to noise.<sup>6</sup> EES, extracellular-extravascular; IC, intracellular; VASC, vascular.

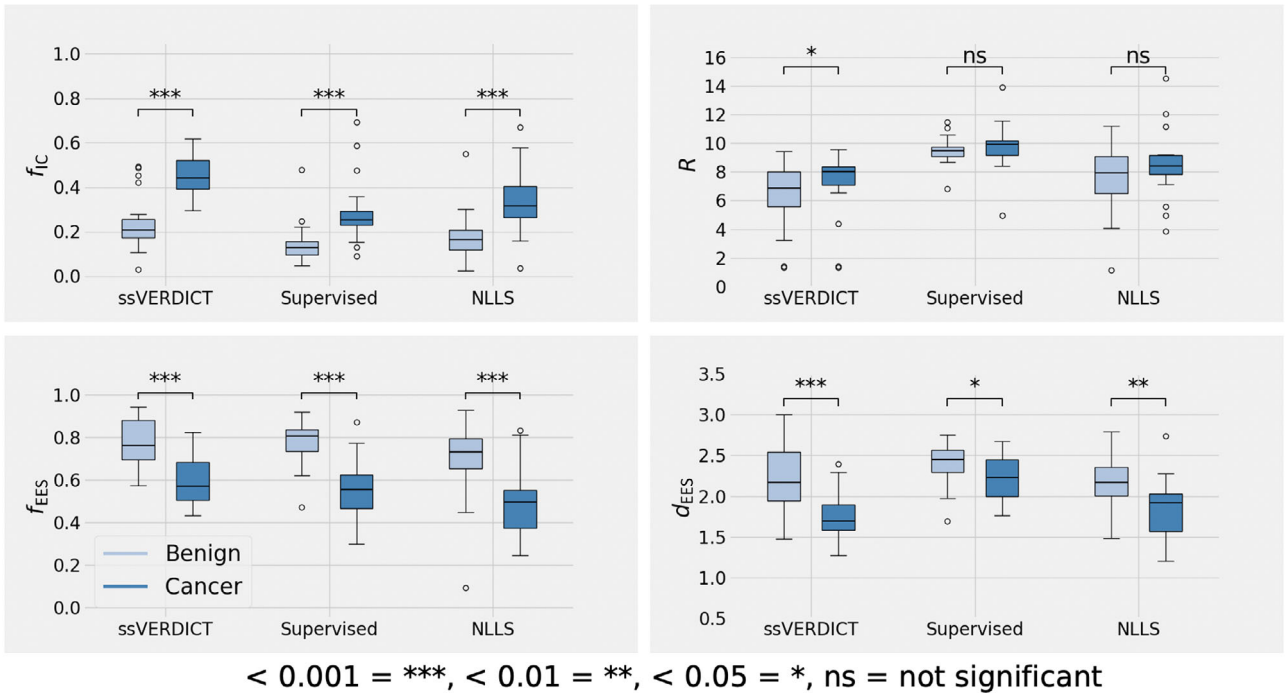
cancerous prostate tissue when compared with NLLS and supervised fitting in two ways:

1. Shows statistically significant differences at  $p < 0.001$  with extracellular-extravascular diffusivity ( $d_{EES}$ ), which is not achieved by the other techniques; and
2. Shows statistically significant differences at  $p < 0.05$  for cell radius ( $R$ ), which is not seen with supervised DL or NLLS fitting.

Figure 5 shows boxplots of the difference between the fitted VERDICT parameters and the ground-truth values for 10 000 simulated signals via the three fitting methods at varying SNR. In general, we observe that ssVERDICT

results in estimates with a median difference closest to zero and smaller interquartile ranges, including at low SNR. More specifically, the supervised and NLLS methods result in differences with interquartile ranges over half the span of the parameter range from a relatively high SNR, whereas ssVERDICT maintains robust fitting until  $SNR = 10$  for  $f_{IC}$ ,  $f_{EES}$ , and  $R$  (and  $SNR = 20$  for  $d_{EES}$ ). This suggests more accurate estimation via ssVERDICT across a range of SNR values, with reduced robustness only at very low SNR. This also supports our decision to simulate data with an SNR of 50, as we show that parameter estimation remains robust across a wide range of SNRs, and this value is close to that observed in practice.





**FIGURE 4** Boxplots of four fitted vascular, extracellular and restricted diffusion for cytometry in tumors (VERDICT) parameter values in benign and cancerous tissue regions in a data set of 20 prostate cancer patients, calculated via the three fitting methods. We find that self-supervised neural network for fitting VERDICT (ssVERDICT) maintains the high level of statistical significance achieved by the baseline methods when using  $f_{IC}$  and  $f_{EES}$  for tissue discrimination. ssVERDICT also improves the level of statistical significance with  $d_{EES}$  and achieves statistical significance with  $R$ . EES, extracellular-extravascular; IC, intracellular; NLLS, nonlinear least squares.

We also investigate correlations between the VERDICT parameters using ssVERDICT (with results included as Figures S2 and S3). We propose that these arise due to the inevitable way in which that biophysical compartments interact; for example, a higher volume fraction for one compartment will always lead to lower volume fractions for other compartments. The same trends are observed in both simulations (see Figure S2) and real data (see Figure S3); therefore, we conclude that they are not an artifact of the fitting process.

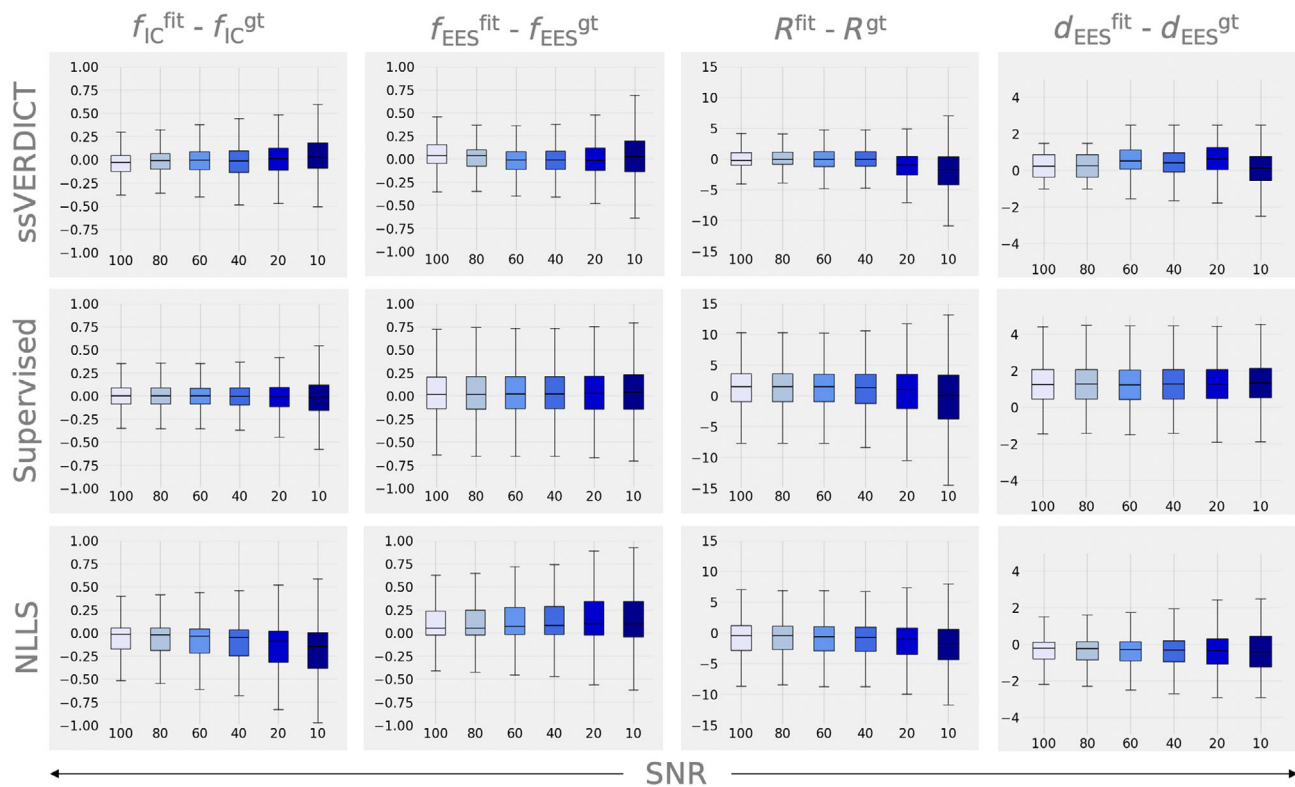
## 4 | DISCUSSION

PCa diagnosis can be significantly improved by the introduction of noninvasive biomarkers derived from quantitative diffusion MRI.<sup>20,35</sup> However, clinical adoption of such techniques requires robust model fitting to avoid misdiagnosis.<sup>13,36</sup> This study presents a self-supervised fitting strategy (ssVERDICT) that can support biophysical multi-compartment dMRI models, demonstrated with the three-compartment VERDICT prostate.<sup>20,22</sup> Previously, self-supervised model fitting was limited only to simple exponential dMRI models.<sup>5,15–19</sup> This is likely due to the difficulty involved in formulating complex signal

equations (typical of biophysical models) as a differentiable forward model. Our work is a key step-change for self-supervised machine learning for dMRI model fitting, moving from simple models to more detailed multi-compartment biophysical models. We use a version of the VERDICT model with four free parameters previously used in Bonet-Carne et al.,<sup>24</sup> to achieve a trade-off between biophysical realism and robust model fitting with traditional methods. This version has  $d_{EES}$  unfixed, which has shown biomarker potential for aiding discrimination of false-positive cases of PCa.<sup>26</sup>

We demonstrate that ssVERDICT outperforms the two baseline approaches for VERDICT model fitting (conventional iterative fitting [NLLS] and supervised DL fitting)<sup>20–22,26,29,30,37</sup> across a range of quantitative metrics. We also use ssVERDICT on clinical in vivo prostate data, showing excellent tissue discrimination between benign and cancerous tissue. Our work investigates DL model fitting-estimation bias in prostate imaging, achieving reduced bias in comparison to supervised DL in simulations. Both DL methods outperform the NLLS fit both in terms of efficiency and robustness, in line with results from Refs. 5 and 15, strongly motivating the use of machine learning for diffusion MRI model-fitting tasks and challenging the widespread use of NLLS as the standard





**FIGURE 5** Boxplots of difference between fitted vascular, extracellular and restricted diffusion for cytometry in tumors (VERDICT) parameter values and simulated ground truth for 10 000 simulated signals using the three fitting strategies. We find median differences closest to zero and smaller interquartile ranges in general across the four parameters when using self-supervised neural network for fitting VERDICT (ssVERDICT), suggesting more accurate fitting by our method across a range of signal-to-noise ratio (SNR) values. EES, extracellular-extravascular; IC, intracellular; NLLS, nonlinear least squares.

fitting procedure. *PyTorch* code for the VERDICT prostate model, as well as instructions on how to implement self-supervised fitting of other VERDICT-based biophysical models,<sup>21,37</sup> is available at <https://github.com/snigdhya-sen/ssVERDICT>. The differentiable form of the compartments can also be used to enable self-supervised fitting of other complex diffusion models, such as the “sphere” for NODDI<sup>38</sup> and “astrosticks” for the SANDI (Soma And Neurite Density Imaging) model.<sup>39</sup>

Our simulation results suggest ssVERDICT can estimate the underlying microstructure more accurately than supervised DL and NLLS, as we observe stronger correlations between parameter estimates and ground truth and improved robustness to additional noise in the acquisition (i.e., at low SNR). We also found reduced bias and MSE across all four fitted VERDICT parameters when using ssVERDICT in comparison to the other methods, as well as lower variance when estimating  $f_{EES}$  and  $d_{EES}$  (Table 1). We note these trends are generally also reflected at varying SNR. These results agree with Refs. 5 and 15, which found that self-supervised fitting of the simple IVIM model resulted in more accurate estimation than NLLS and lower bias than supervised DL. Our results demonstrate that this

improvement in estimation translates to a significantly more complex multi-compartment model.

Analysis of real patient data with ssVERDICT shows promising results in vivo, achieving the best tumor conspicuity over all VERDICT maps (e.g., Figure 3) with enhanced tissue type discrimination. Importantly, we found higher statistical significance for  $d_{EES}$  with ssVERDICT in comparison to other methods—a parameter that has been shown to aid the characterization of false-positive lesions.<sup>26</sup> The improved lesion conspicuity and characterization with ssVERDICT shows great promise for the translation of the technique into the standard PI-RADS (prostate imaging reporting & data system)<sup>40</sup> for the scoring of indeterminate lesions (PI-RADS 3 and 4).<sup>23</sup> This strongly suggests that the benefits of our technique will translate to clinical practice, improving noninvasive tumor characterization and hence further reducing invasive biopsies.

This work is limited primarily by the small size of the patient data set and the range of prostatic disease included. We originally chose a cohort of 20 patients to enable an initial analysis of the efficacy of a new technique and include a range of cancer grades. Our initial cohort

demonstrates statistical significance in several parameters. However, a larger cohort study is necessary to determine the effect sizes with more confidence. We also only focus on voxel-wise methods rather than extending to architectures that learn spatial correspondences in images such as CNNs or spatial transformers. Although a self-supervised CNN has been demonstrated for the IVIM model,<sup>16,17</sup> and supervised CNN methods have been used widely for dMRI model fitting,<sup>10,12</sup> we instead focus on voxel-wise fitting methods to enable a clear comparison between ssVERDICT and the currently used VERDICT fitting techniques in a controlled environment.

Future work will aim to increase statistical significance with a larger patient cohort and incorporate a wider range of prostatic diseases<sup>26,29</sup> to test ssVERDICT's ability to accurately characterize tissue microstructure and maximize its potential clinical impact. Our results in Figure 4 show improved separation between benign and cancerous tissue with ssVERDICT, which may translate to improved discrimination between Gleason grades (see Figure S1) and characterization of indeterminate cases. We will also investigate fitting more complex biophysical dMRI models such as VERDICT with tissue relaxation and the brain VERDICT model<sup>21,37</sup> via a self-supervised CNN approach similar to Refs. 16 and 17, to investigate potential further gains in fitting speed and accuracy. Finally, we will conduct a comparison between the supervised and self-supervised methods across the entire range of machine learning hyperparameters, to optimize the network for VERDICT model fitting.

In conclusion, our work shows that self-supervised fitting of the VERDICT prostate model performs better in simulations and in vivo data than baseline methods. This study is the first to extend self-supervised model fitting beyond multi-exponential models. Our results demonstrate that ssVERDICT characterizes prostate tumor microstructure to enable improved discrimination between benign and cancerous tissue, contributing toward the ultimate goal of reducing the number of biopsies and improving patient care.

## ACKNOWLEDGMENTS

This work was supported by the EPSRC-funded UCL Center for Doctoral Training in Intelligent, Integrated Imaging in Healthcare (i4health) (EP/S021930/1) and the Department of Health's NIHR-funded Biomedical Research Centre at University College London Hospitals. This work is also funded by EPSRC (EP/N021967/1, EP/R006032/1, EP/V034537/1) and by Prostate Cancer UK, Targeted Call 2014, Translational Research St.2 (PG14-018-TR2).

## DATA AVAILABILITY STATEMENT

*PyTorch* code for the VERDICT prostate model, as well as instructions on how to implement self-supervised fitting of other VERDICT-based biophysical models,<sup>21,37</sup> is available at <https://github.com/snigdha-sen/ssVERDICT>.

## ORCID

Snigdha Sen  <https://orcid.org/0009-0009-1621-3979>

David Atkinson  <https://orcid.org/0000-0003-1124-6666>

Eleftheria Panagiotaki  <https://orcid.org/0000-0002-2697-3447>

## TWITTER

Snigdha Sen  [snigdhasen98](https://twitter.com/snigdhasen98)

## REFERENCES

- Ghai S, Haider MA. Multiparametric-MRI in diagnosis of prostate cancer. *Indian J Urol.* 2015;31:194-201. doi:10.4103/0970-1591.159606
- Panagiotaki E, Schneider T, Siow B, Hall MG, Lythgoe MF, Alexander DC. Compartment models of the diffusion mr signal in brain white matter: a taxonomy and comparison. *Neuroimage.* 2012;59:2241-2254. doi:10.1016/j.neuroimage.2011.09.081
- Rahbar H, Zhang Z, Chenevert TL, et al. Utility of diffusion-weighted imaging to decrease unnecessary biopsies prompted by breast MRI: a trial of the ECOG-ACRIN cancer research group. *Clin Cancer Res.* 2019;25:1756-1765. doi:10.1158/1078-0432.CCR-18-2967
- Singh S, Rogers H, Kanber B, et al. Avoiding unnecessary biopsy after multiparametric prostate MRI with VERDICT analysis: the INNOVATE study. *Radiology.* 2022;305:623-630. doi:10.1148/radiol.212536
- Epstein SC, Bray TJP, Hall-Craggs M, Zhang H. Choice of training label matters: how to best use deep learning for quantitative MRI parameter estimation. *Machine Learning for Biomedical Imaging.* 2024;2:586-610. doi:10.59275/j.melba.2024-geb5
- Golkov V, Dosovitskiy A, Sperl JI, et al. Q-space deep learning: twelve-fold shorter and model-free diffusion MRI scans. *IEEE Trans Med Imaging.* 2016;35:1344-1351. doi:10.1109/TMI.2016.2551324
- Bertleff M, Domsch S, Weingärtner S, et al. Diffusion parameter mapping with the combined intravoxel incoherent motion and kurtosis model using artificial neural networks at 3 T. *NMR Biomed.* 2017;30:3833. doi:10.1002/nbm.3833
- Grussu F, Battiston M, Palombo M, et al. Deep learning model fitting for diffusion-relaxometry: a comparative study. *Computational Diffusion MRI.* Springer International Publishing; 2020:159-172.
- Ye C, Cui Y, Li X. Q-space learning with synthesized training data. *Computational Diffusion MRI.* Springer International Publishing; 2019:123-132.
- Aliotta E, Nourzadeh H, Patel SH. Extracting diffusion tensor fractional anisotropy and mean diffusivity from 3-direction dwi scans using deep learning. *Magn Reson Med.* 2021;85:845-854. doi:10.1002/mrm.28470

11. Zheng T, Yan G, Li H, et al. A microstructure estimation transformer inspired by sparse representation for diffusion MRI. *Med Image Anal.* 2023;86:102788. doi:10.1016/j.media.2023.102788
12. HashemizadehKolowri SK, Chen R, Adluru G, DiBella EVR. Jointly estimating parametric maps of multiple diffusion models from undersampled q-space data: a comparison of three deep learning approaches. *Magn Reson Med.* 2022;87:2957-2971. doi:10.1002/mrm.29162
13. Gyori NG, Palombo M, Clark CA, Zhang H, Alexander DC. Training data distribution significantly impacts the estimation of tissue microstructure with machine learning. *Magn Reson Med.* 2021;87:932-947. doi:10.1002/mrm.29014
14. Ericsson L, Gouk H, Loy CC, Hospedales TM. Self-supervised representation learning: introduction, advances, and challenges. *IEEE Signal Process Mag.* 2022;39:42-62. doi:10.1109/MSP.2021.3134634
15. Barbieri S, Guernsey-Champion OJ, Klassen R, Thoeny HC. Deep learning how to fit an intravoxel incoherent motion model to diffusion-weighted MRI. *Magn Reson Med.* 2020;83:312-321. doi:10.1002/mrm.27910
16. Huang H. An unsupervised convolutional neural network method for estimation of intravoxel incoherent motion parameters. *Phys Med Biol.* 2022;67:215018. doi:10.1088/1361-6560/ac9a1f
17. Vasylechko SD, Warfield SK, Afacan O, Kurugol S. Self-supervised IVIM DWI parameter estimation with a physics based forward model. *Magn Reson Med.* 2022;87:904-914. doi:10.1002/mrm.28989
18. Troelstra MA, Van Dijk A, Witjes JJ, et al. Self-supervised neural network improves tri-exponential intravoxel incoherent motion model fitting compared to least-squares fitting in non-alcoholic fatty liver disease. *Front Physiol.* 2022;13:942495. doi:10.3389/fphys.2022.942495
19. Lee W, Choi G, Lee J, Park HW. Registration and quantification network (RQnet) for IVIM-DKI analysis in MRI. *Magn Reson Med.* 2023;89:250-261. doi:10.1002/mrm.29454
20. Panagiotaki E, Walker-Samuel S, Siow B, et al. Non-invasive quantification of solid tumor microstructure using verdict MRI. *Cancer Res.* 2014;74:1902-1912. doi:10.1158/0008-5472.CAN-13-2511
21. Palombo M, Valindria V, Singh S, et al. Joint estimation of relaxation and diffusion tissue parameters for prostate cancer with relaxation-VERDICT MRI. *Sci Rep.* 2023;13:2999. doi:10.1038/s41598-023-30182-1
22. Panagiotaki E, Chan RW, Dikaos N, et al. Microstructural characterization of normal and malignant human prostate tissue with vascular, extracellular, and restricted diffusion for cytometry in tumours. *Invest Radiol.* 2015;50:218-227. doi:10.1097/RLI.0000000000000115
23. Johnston EW, Bonet-Carne E, Ferizi U, et al. VERDICT MRI for prostate cancer: intracellular volume fraction versus apparent diffusion coefficient. *Radiology.* 2019;291:391-397. doi:10.1148/radiol.2019181749
24. Bonet-Carne E, Johnston E, Daducci A, et al. VERDICT-AMICO: ultrafast fitting algorithm for non-invasive prostate microstructure characterization. *NMR Biomed.* 2019;32:4019. doi:10.1002/nbm.4019
25. Murday JS, Cotts RM. Self-diffusion coefficient of liquid lithium. *J Chem Phys.* 1984;48:4938-4945.
26. Sen S, Valindria V, Slator PJ, et al. Differentiating false positive lesions from clinically significant cancer and normal prostate tissue using VERDICT MRI and other diffusion models. *Diagnostics.* 2022;12:1631. doi:10.3390/diagnostics12071631
27. Johnston E, Pye H, Bonet-Carne E, 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. doi:10.1186/s12885-016-2856-2
28. Panagiotaki E, Ianus A, Johnston E, et al. Optimised verdict MRI protocol for prostate cancer characterisation. In: *Proceedings of the 23rd Annual Meeting of ISMRM.* Vol 2872.) International Society of Magnetic Resonance for Medicine; 2015.
29. Sen S, Valindria V, Pye H, et al. Verdict-MRI analysis of false positives in prostate mp-MRI. In: *Proceedings of the 31st Annual Meeting of ISMRM.* Vol 926. International Society of Magnetic Resonance for Medicine; 2022.
30. Valindria V, Chiou E, Palombo M, Singh S, Punwani S, Panagiotaki E. Synthetic q-space learning with deep regression networks for prostate cancer characterisation with verdict. In: *Proceedings of the 18th IEEE International Symposium on Biomedical Imaging (ISBI).* International Society of Magnetic Resonance for Medicine; 2021.
31. Veraart J, Fieremans E, Novikov DS. Diffusion MRI noise mapping using random matrix theory. *Magn Reson Med.* 2016;76:1582-1593. doi:10.1002/mrm.26059
32. Tournier JD, Smith R, Raffelt D, et al. MRtrix3: a fast, flexible and open software framework for medical image processing and visualisation. *Neuroimage.* 2019;202:116137. doi:10.1016/j.neuroimage.2019.116137
33. Kellner E, Dhital B, Kiselev VG, Reiser M. Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magn Reson Med.* 2015;76:1574-1581. doi:10.1002/mrm.26054
34. 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: *Proceedings of the 28th Annual Meeting of ISMRM [Virtual].* Vol 710. International Society of Magnetic Resonance for Medicine; 2020.
35. Riches SF, Hawtin K, Charles-Edwards EM, De Souza NM. Diffusion-weighted imaging of the prostate and rectal wall: comparison of biexponential and monoexponential modelled diffusion and associated perfusion coefficients. *NMR Biomed.* 2009;22:318-325. doi:10.1002/nbm.1328
36. Barbieri S, Donati OF, Froehlich JM, Thoeny HC. Impact of the calculation algorithm on biexponential fitting of diffusion-weighted MRI in upper abdominal organs. *Magn Reson Med.* 2016;75:2175-2184. doi:10.1002/mrm.25765
37. Figini M, Castellano A, Bailo M, et al. Comprehensive brain tumour characterisation with VERDICT-MRI: evaluation of cellular and vascular measures validated by histology. *Cancer.* 2023;15:2490. doi:10.3390/cancers15092490
38. Zhang H, Schneider T, Wheeler-Kingshott CAM, Alexander DC. NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage.* 2012;61:1000-1016. doi:10.1016/j.neuroimage.2012.03.072
39. Palombo M, Ianus A, Guerreri M, et al. SANDI: a compartment-based model for non-invasive apparent



soma and neurite imaging by diffusion MRI. *Neuroimage*. 2020;215:116835. doi:10.1016/j.neuroimage.2020.116835

40. Turkbey B, Rosenkrantz AB, Haider MA, et al. Prostate imaging reporting and data system version 2.1: 2019 update of prostate imaging reporting and data system version 2. *Eur Urol*. 2019;76:340-351. doi:10.1016/j.eururo.2019.02.033

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**Figure S1.** Parameter maps of a patient with a Gleason Grade 4 + 4 tumor in the patient's left (*image right*) posterior peripheral zone and a patient with a Gleason Grade 4 + 4 tumor in the patient's left transition zone (*upper*) and a Gleason Grade 4 + 5 tumor in the patient's left peripheral zone (*lower*). These maps show that the trends of clearer tumor conspicuity with ssVERDICT (self-supervised neural network for fitting VERDICT [vascular, extracellular and restricted diffusion for cytometry in tumors]) continue over a range of Gleason grades, and we may see trends such as higher intracellular volume fraction  $f_{IC}$  for higher Gleason grades.

**Figure S2.** Correlations between VERDICT (vascular, extracellular and restricted diffusion for cytometry in tumors) parameters on 10 000 simulated voxels. We observe strong correlation between  $f_{EES}$  and  $f_{IC}$  ( $r = -0.8298$ ), as the sum of the three volume fractions must equal to one; therefore, as one increases, the others will decrease. We also observe weak correlation between  $R$  and  $d_{EES}$ .

**Figure S3.** Correlations between VERDICT (vascular, extracellular and restricted diffusion for cytometry in tumors) parameters on the data set of 20 prostate cancer (PCa) patients, with all voxels plotted for each patient in benign and tumor regions of interest (ROIs). We observe strong correlation between  $f_{EES}$  and  $f_{IC}$  ( $r = -0.7416$ ), as the sum of the three volume fractions must equal to one; therefore, as one increases, the others will decrease. We see some weaker correlation among other parameters.

**How to cite this article:** Sen S, Singh S, Pye H, et al. ssVERDICT: Self-supervised VERDICT-MRI for enhanced prostate tumor characterization. *Magn Reson Med*. 2024;1-12. doi: 10.1002/mrm.30186