# Improved prediction of recurrence after prostate cancer radiotherapy using multimodal data and *in silico* simulations

Valentin Septiers[1,3], Carlos Sosa-Marrero[1], Renaud de Crevoisier[1], Aurélien Briens[2], Hilda Chourak[1], Maria A. Zuluaga[3], and Oscar Acosta[1]

[1] Univ Rennes, CLCC Eugène Marquis, Inserm, LTSI - UMR 1099, F-35000 Rennes, France
[2] CLCC Eugène Marquis, F-35000, Rennes, France
[3] EURECOM, Data Science Department, F-06410 Biot, France

**Abstract.** Prediction of biochemical recurrence (BCR) after prostate cancer radiotherapy is crucial for devising personalised treatments. BCR has been traditionally predicted using clinical data or *in vivo* imaging within AI frameworks such as radiomics approaches, but with limited results and reduced interpretability. These analysis are additionally hindered by the imbalanced and heterogeneous nature of data. In this paper, we present a novel approach to predict BCR at 5 years, based not only on clinical and image features, but also on a patient specific radiobiological mechanistic *in silico* model simulating tumour growth and radiation response. By combining all these data, we aim at i) improving the prediction of BCR after prostate cancer radiotherapy (RT), and ii) bringing interpretability to this prediction. A cohort of 254 patients was used. Pre-treatment T2-w MRIs, ADC maps and 7 clinicopathological characteristics were available. Patient specific digital twins of tumours were created from MRIs. The prescribed treatment was simulated with the mechanistic model yielding 414 features characterising the response of the tumour to RT. A first univariate feature selection analysis was conducted to select the most predictive features. Then, a machine learning algorithm was trained using selected features and compared with a deep learning (DL) approach based on clinicopathological characteristics and MRIs. Our approach achieved an AUC of 0.74 by training a random forest classifier combining most predictive features. The DL model achieved an AUC of 0.69. This methodology opens the road to interpretability of the response to radiotherapy and tailored treatments for prostate cancer patients.

**Keywords:** Recurrence prediction · prostate cancer · *in silico* modeling · artificial intelligence · interpretability/explainability.

## 1 Introduction

Prostate cancer is the second most diagnosed cancer in men in the world and the fifth leading cause of death [22]. External Beam Radiotherapy (RT) is the

clinical standard treatment for localized prostate cancer [12], which allows to control the tumour in the majority of cases. However, biochemical recurrence (BCR), defined as 2 consecutive elevations of Prostate Specific Antigen (PSA) $\geq$ 0.2ng/mL (Phoenix criteria [1]), may occur in 0-10%, 10-20% and 30-40% of patients with respectively low, intermediate and high-risk tumours (according to the D'Amico Classification) within 5 years after the treatment [5]. Predicting BCR prior to RT appears as crucial for assessing patient risk and personalising treatment. Several models have been previously proposed to predict BCR. The first tumour control probability models, [14, 4] were based on probability curves, describing the dose-effect relationship on a given population. Nevertheless, these models are limited as they only consider dose discarding the rich nature of the tumour. Recently, radiomics approaches introduced image biomarkers by extracting multiple tumour and organ features from available medical images [3, 16], improving performances but facing several issues [9, 6]. These approaches require a large amount of training data and they are known to be *black-box* methods lacking interpretability and explainability. Furthermore, they are highly dependent on the data (imbalance classes, image harmonisation, external validation, etc). Deep learning (DL) models have emerged as an appealing tool to predict BCR, yielding better results than radiomics [17],but sharing the same issues with data. In contrast to these data-driven techniques, mechanistic *in silico* models open new possibilities to predict tumour response by simulating the prescribed irradiation treatment. These models are based on the integration of several biological mechanisms to better understand the response of patients to RT. Their predictive capabilities have been explored in [18]. These models may offer better interpretability and could allow to simulate different patient-specific treatments.

The objective of the present work was thus to propose a novel approach to predict BCR after prostate cancer radiotherapy in a patient specific framework, by combining Magnetic Resonance Imaging (MRI), clinicopathological data, and a radiobiological mechanistic *in silico* model simulating tumour growth and radiation response. The model, based on [20], integrates the most relevant radiobiological mechanisms identified by a sensitivity analysis : oxygenation, tumour cells division and irradiation response. It allows us to create digital twins of patient tumours from MR images and simulate treatments and tissue response. This novel approach was compared to radiomics and deep-learning-based predictions from image and clinicopathological features.

## 2 Materials and methods

### 2.1 Population dataset

A cohort of 254 patients with localised prostate cancer having undergone RT was used for this study (performed in line with the principles of the Declaration of Helsinki). A detailed description of patient and tumour characteristics can be found in **Table 1**. Before starting the treatment, 3T MRIs were acquired as described in [9]. It included axial turbo spin echo T2-w and axial diffusion using multiple b-values. N4-bias field correction has been done on T2w images,

and apparent diffusion coefficient (ADC) maps were calculated. Prostate and tumour were manually segmented by experts on T2-w sequences and contours were propagated onto the co-registered ADC images (Step (0) in **Fig.1**). Patients were followed up through clinical examination and PSA analysis every 6 months for 5 years after the end of irradiation. A total of 39 patients suffered BCR, defined according to the Phoenix criteria [1]. More details about the clinicopathological features available are shown in the Supplementary Materials.

**Table 1.** Patients description and tumour characteristics.

| **Patients description** | **Tumour characteristics** (number of instances and their percentage in parenthesis) | | | |
|---|---|---|---|---|
| ***Number of patients*** | **Pre-treatment PSA (ng/mL)** | | | |
| 254 | $\boldsymbol{PSA} \leq 7$ | $7 < \boldsymbol{PSA} \leq 11$ | $11 < \boldsymbol{PSA} \leq 20$ | $\boldsymbol{PSA} > 20$ |
| ***Median age (years)*** | 66 (26%) | 63 (25%) | 66 (26%) | 59 (23%) |
| 71 | **Clinical stage (T stage)** | | | |
| **Recurrence** | ***T1*** | ***T2*** | ***T3*** | ***T4*** |
| ***Non BCR*** ***BCR*** | 23 (9%) | 108 (43%) | 120 (47%) | 3 (1%) |
| 215 39 | **Gleason score** | | | |
| ***Total Dose (Gray)*** | ***6*** | ***7*** | ***8*** | ***9*** |
| 74 - 80 | 38 (15%) | 157 (62%) | 30 (12%) | 29 (11%) |

### 2.2 Mechanisitic *in silico* model

A previously developed mechanistic *in silico* model simulating tumour growth and response to RT was used [20]. It was implemented in C++, based on the Multiformalism Modeling and Simulation Library (M2SL) [10], allowing the integration of different mechanisms arising at different temporal and spatial scales. It considered 2D digital twins of patient tumours in which each pixel ($20\mu m$ x $20\mu m$) represented a cell corresponding to 6 types : healthy (fibroblasts, macrophages, epithelial, muscle, etc), undamaged tumour, lethally damaged tumour, pre-existing endothelial (vessel cells), neo-created endothelial and dead cells. This mechanistic model integrated major radiobiological mechanisms, occurring at various temporal and spatial dimensions. (a) **Angiogenesis**, the recruitment, creation of new vessels from pre-existing endothelial cells, was based on the diffusion of vascular endothelial growth factor (VEGF). The VEGF distribution was modelled using a reaction-diffusion equation. (b) **Division of healthy and** (c) **tumour cells** based on the cell cycle, which consisted of phases G1 (Gap 1), S (Synthesis), G2 (Gap 2) and M (Mitosis). It also included a fifth phase G0 in which cells were quiescent. (d) **Oxygenation** modeled using a reaction-diffusion equation. (e) **Response to irradiation** given by the linear-quadratic model [8]. To identify the most important parameters and mechanisms involved in tumour response, a sensitivity analysis based Morris screening method was previously performed. Following this sensitivity analysis, a reduced

model of 18 parameters was obtained, which included only (c) **Division of tumour cells**, (d) **Oxygenation** and (e) **Response to irradiation**. This model is interpretable in terms of tumour response as it involves understandable equations (differential and linear quadratic equations) and explainable outputs (tumour volume, percentage of cells in the cycle, at each time step). In this paper, we used this reduced mechanistic *in silico* model to simulate the patient specific tumour growth and response to RT. All the equations and parameters details used are fully described in [20].

### 2.3 Radiomics feature extraction

IBSI-compliant [25] features were extracted (Step (1) in **Fig.1**) from both modalities within the tumour using the Pyradiomics library [24]. It included 18 first-order statistics, 23 shape descriptors and 66 textural features. Prior to the feature extraction, images were resampled to 1mm×1mm×1mm using BSpline interpolation. This provided a total of 214 features (107 features from T2 modality and 107 from ADC modality). Details about radiomics features extracted are shown in the Supplementary Materials.

### 2.4 Digital twin and patient-specific *in silico* simulation

A total of 254 2D digital tissues representing the 254 patients specific tumours of the cohort were built (Step (2) in **Fig.1**). These digital twins were initialized using different parameters (i.e. cell size, the tumour area, the tumour density and the vascular density). Every 3D tumour volume was mapped to a disk and considered as a perfect circle (ratio between each axis of the circle equals to 1). The cell size is set to 20.0 reflecting the average size of a cell (i.e. 20.0 $\mu m$). The initial area value was computed from the spherical tumour volume. The mean density value was obtained from the average intensity value inside the Volume of Interest (VOI) of the T2-w MR image through linear transformations [15]. These transformations provided 10 supplementary features (5 from T2 and 5 from ADC). An initial prostate-specific vascular density of 3.8% [19] was considered for every virtual tissue (simulating poorly-vascularized tumour core). Then, the prescribed standard irradiation treatment (74-80 Gy administered in 2 Gray fractions from Monday to Friday) was simulated through the 8 weeks of treatment (Step (3) in **Fig.1**). The *in silico* simulations produced several outputs at each time point from the beginning of the treatment simulation until 4 weeks after the end of the treatment. Outputs of the mechanistic model were the tumour volume and density, the volume of not damaged tumour cells, the hypoxic cell density, the killed cells percentage, or the percentage of cells in each phase of the cycle.

Exploratory computation of other features, based on *in silico* simulation outputs, was performed in order to produce features that might be good predictors of BCR. Relevant time points were chosen : t = 0; 2; 4; 6; 8 (end of the treatment); 10 and 12 weeks. Values of *in silico* simulation outputs were extracted at each time point, as well as the difference values and the fraction values between

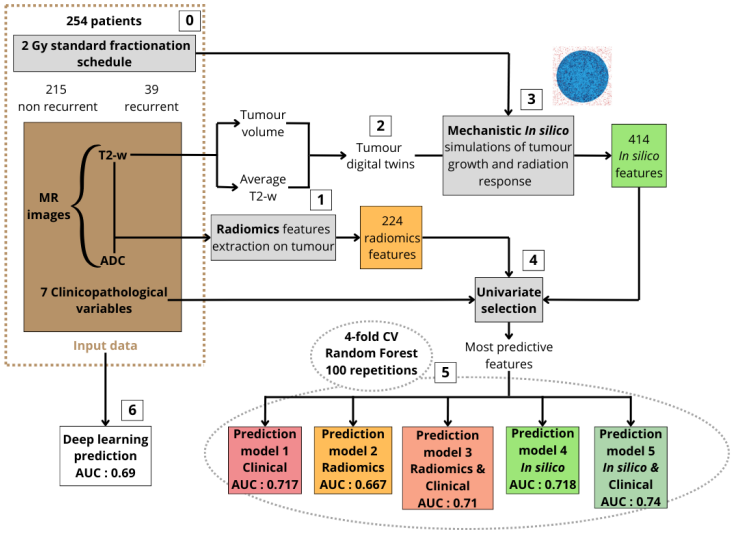


**Fig. 1.** Workflow. Data preparation and tumour segmentation (0). Image feature extraction (1). Tumour digital twins characterisation and construction (2). *In silico* simulations (3). Univariate feature selection (4). BCR prediction with different features combination with AUC prediction scores (5). BCR prediction with deep learning (6).

each time points. The curve tendency between each time point was also computed. A fitting of each output data was performed between t = 0 until t = 10 weeks (2 weeks after treatment end) using functions that best fit (for instance, periodic functions for the phases of the cell cycle). Parameters of these fitted curves were extracted and tested as features to predict BCR. Values and times when the fitted function is equal to 5% of the initial output value were also computed. This exploratory feature computation provided 414 features. Details about these mechanistic features are shown in the Supplementary Materials.

### 2.5 Recurrence prediction

In total, 645 features (7 clinical, 224 radiomics-based, 414 mechanistic-based) were extracted. A first univariate feature selection analsyis was performed on all these features using a Random Forest Classifier (RFC) stratified 4-fold cross-validation, 20 repetitions (Step (4) in **Fig.1**). Through this univariate analysis, the 20 most predictive features were identified (3 clinicopathological, 7 radiomics and 10 mechanistic features). A correlation analysis was performed in order to remove very strongly correlated features (i.e. Spearman's correlation coefficient $> 0.8$ [2]). Seven very strongly correlated features were identified (1 clinicopathological, 3 radiomics and 3 mechanistic). This resulted in 13 features split in the different prediction models according to the workflow in **Fig.1**. A RFC (stratified 4-fold cross-validation, 100 repetitions) was used to predict BCR with different combinations of features selected (Step (5) in **Fig.1**). **Model 1** Only

clinical variables were used. **Model 2** Only radiomics features. **Model 3** Model 2 + clinical variables. **Model 4** Only mechanistic features. **Model 5** Model 4 + clinical variables (proposed approach). Performance evaluation was quantified through different metrics : the area under the ROC curve (AUC), the classification accuracy (ACC), the average precision score (APS), the F1 score (F1), the precision (PREC) and the recall (REC).

### 2.6 Comparison with a Deep Learning approach

For comparison with other data-based approaches, we developed a DL neural network fed with raw MR images and clinicopathological features to predict binary BCR.

**Model construction** The DL model consisted of 3 parts : (i) A Convolutional Neural Network (CNN) taking as input the cropped T2-w MR images. (ii) A CNN taking as input the cropped ADC maps. The cropped images were centered on the tumour. (iii) A Fully Connected Neural Network (FCNN) taking as input the clinicopathological data. The CNNs (i and ii) used were built with 4 convolutional layers. MaxPooling layers were added to decrease the dimensionality. Flattening layers were inserted at the end of each CNN to produce two vectors of features reflecting the inner images data. The FCNN (iii) used for the clinicopathological data was built with 2 dense layers using L2 regularisation (to prevent overfitting). The outputs of these 3 parts were concatenated as a vector of features and put into another FCNN to predict the binary BCR status. Dense layers were used as in the previous FCNN part. The Sigmoid activation function was applied in the output layer, as the prediction was binary. The rectified linear unit activation function was used in the other layers to overcome the vanishing gradient problem. The dropout regularisation technique was used to prevent overfitting [21]. Batch normalisation layers were added for further regularisation and to reduce the internal covariate shift [11]. The Adam optimiser was used [13]. The number of parameters of this DL model was 25 566 721, reducing model interpretation. **Training and testing** The dataset, consisting in 215 non BCR patients and 39 BCR patients was split with a 75:25% train/test split. During the training, the binary cross-entropy loss was monitored and early stopping was used. As the dataset was imbalanced class weights were applied, defined as :

$$\frac{numberOfPatients}{numberOfClasses * numberOfInstancesOfEachClass}$$

Stratified 4-fold cross-validation, 5 repetitions, was performed. Through the evaluation process, performance was assessed with the same metrics as the other models (Step (6) in **Fig.1**).

## 3 Results and Discussion

The most predictive clinical feature found was the patient ISUP grade group [7]. For the mechanistic-based features, most predictive identified were variables

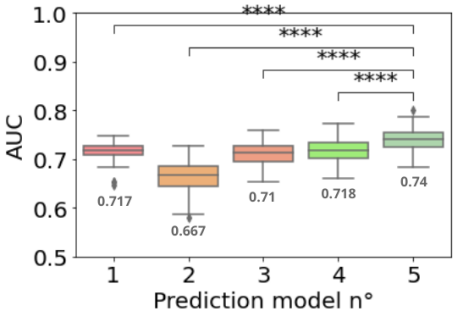


**Fig. 2.** AUC of prediction models trained with a random forest classifier. Same colour code and model numbers as in **Fig.1** are used. (1) Clinicopathological features; (2) Radiomics features; (3) Radiomics and Clinicopathological features. (4) *In silico* features. (5) *In Silico* and clinicopathological features. The stars correspond to the Wilcoxon signed-rank test significance. **** : p-value $\leq 0.0001$

underlying the percentage of tumour cells in phase M such as the curve tendency between 4 and 6 weeks, or the percentage of phase M tumour cells 2 weeks after the end of the treatment. These features are explainable as they have a biological and clinical interpretation. They bring insights into tumour response during RT, that are not visible from any clinical pre-treatment image and, to the best of our knowledge, have never been considered in any predictive model. The most predictive radiomics features identified were the skewness and the texture feature cluster prominence which are more complex to clinically interpret. Univariate feature selection and correlation matrix details are shown in the Supplementary Materials. Results of BCR prediction for each model trained with the RFC are shown in **Fig.2**. Results of the model classification performances are shown in **Table 2**. The proposed approach (Model 5), based on clinical and mechanistic data achieved a mean AUC of 0.74 over the 100 repetitions, outperforming the four other models trained with the other features combinations. It was significantly better than these models (p-value $\leq 0.0001$, given by a Wilcoxon signed-rank test with Bonferroni correction). It was not significantly better than models 2 and 3 in term of precision. The DL model achieved a mean AUC of 0.69, not outperforming the proposed approach in terms of AUC, accuracy and average precision. However, in term of F1 score and recall, the proposed novel approach was outperformed by the DL model and the radiomics ones.

This study presents several limitations to be explored in future work. First of all, representing tumours via spherical mapping may seem simplistic, however this was also used by [19], and simulation results were in good agreement with their *in vivo* observations in mice. Initialisation of digital twins of patient tumours could have included ADC maps, which may provide additional information on cell density. The use of IntraVoxel Incoherent Motion DW-MRI could also be explored to better initialise patient-specific digital twins. Secondly,

**Table 2.** BCR prediction performance. Bold denotes top or significantly close to top results (Wilcoxon signed-rank test shows no significant difference)

| Model | ACC | APS | F1 | PREC | REC | AUC |
|---|---|---|---|---|---|---|
| Random Forest Classifier | | | | | | |
| Model 1 (Clinical Data) | 0.837 | 0.320 | 0.033 | 0.029 | 0.018 | 0.717 |
| Model 2 (Radiomics Data) | 0.840 | 0.376 | 0.237 | **0.491** | 0.174 | 0.667 |
| Model 3 (Radiomics + Clinical) | 0.839 | 0.389 | 0.233 | **0.451** | 0.167 | 0.710 |
| Model 4 (Mechanistic Data) | 0.830 | 0.325 | 0.064 | 0.194 | 0.004 | 0.718 |
| Model 5 (Mechanistic + Clinical) | **0.848** | **0.418** | 0.198 | **0.494** | 0.133 | **0.740** |
| Deep Learning model | 0.834 | 0.338 | **0.476** | **0.497** | **0.475** | 0.690 |

AUC: area under the ROC curve; ACC: classification accuracy;
APS: average precision score; F1: F1 score; PREC: precision; REC: recall

the mechanistic *in silico* model used has to be enriched with other biological mechanisms, such as the immune response. The model can be validated with animal experiments, as so far all the parameter values are set from the literature. Other parameters, such as the vascular density (which was set to a constant value), have to be calibrated properly from other image modalities. This could be done by assessing the hypoxia in the tumour which is well-known as a critical prognosis factor related to radio resistance and recurrence [23]. Furthermore, the exploratory feature selection can be improved using other strategies, such as dimensional reduction. The results presented in **Table 2** may seem underwhelming, in terms of F1 and recall, this is mainly due to the highly imbalanced nature of the dataset. Artificial oversampling techniques, such as SMOTE, can be included in this work. Finally, the approach has to be validated on an external population.

## 4 Conclusion and Future work

We proposed a novel approach to predict BCR after prostate cancer RT. It was based on the combination of clinical features and an *in silico* model of tumour growth and radiation response, initialised with MRI-based digital twins of tumours. This approach brings interpretability and explainability to the prediction, outperforming models based on clinical data, radiomics or DL techniques. Hence, this method enables to personalise treatments by stratifying patients with low or high risk of recurrence, but also opens the possibility to simulate tailored plannings.

**Acknowledgments** With financial support from ITMO Cancer of Aviesan within the framework of the 2021-2030 Cancer Control Strategy, on funds administered by Inserm.

**Disclosure of interests** The authors have no competing interests to declare that are relevant to the content of this article.

## References


[1] Matthew C. Abramowitz et al. "The Phoenix Definition of Biochemical Failure Predicts for Overall Survival in Patients with Prostate Cancer". In: *Cancer* 112.1 (2008), pp. 55–60.

[2] Haldun Akoglu. "User's Guide to Correlation Coefficients". In: *Turkish Journal of Emergency Medicine* 18.3 (Aug. 7, 2018).

[3] Michele Avanzo et al. "Machine and Deep Learning Methods for Radiomics". In: *Medical Physics* 47.5 (May 2020).

[4] M.-A. Chanrion et al. "The Influence of the Local Effect Model Parameters on the Prediction of the Tumor Control Probability for Prostate Cancer". In: *Physics in Medicine & Biology* 59.12 (May 2014), p. 3019.

[5] A. V. D'Amico et al. "Biochemical Outcome after Radical Prostatectomy, External Beam Radiation Therapy, or Interstitial Radiation Therapy for Clinically Localized Prostate Cancer". In: *JAMA* 280.11 (Sept. 16, 1998).

[6] Savannah R. Duenweg et al. "T2-Weighted MRI Radiomic Features Predict Prostate Cancer Presence and Eventual Biochemical Recurrence". In: *Cancers* 15.18 (2023), p. 4437.

[7] Jonathan I. Epstein et al. "The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: Definition of Grading Patterns and Proposal for a New Grading System". In: *The American Journal of Surgical Pathology* 40.2 (Feb. 2016), p. 244.

[8] John F. Fowler. "The Linear-Quadratic Formula and Progress in Fractionated Radiotherapy". In: *British Journal of Radiology* 62.740 (Aug. 1, 1989), pp. 679–694.

[9] Khémara Gnep et al. "Haralick Textural Features on $T_2$ -Weighted MRI Are Associated with Biochemical Recurrence Following Radiotherapy for Peripheral Zone Prostate Cancer: Impact of MRI in Prostate Cancer". In: *Journal of Magnetic Resonance Imaging* 45.1 (Jan. 2017), pp. 103–117.

[10] Alfredo I. Hernández et al. "A Multiformalism and Multiresolution Modelling Environment: Application to the Cardiovascular System and Its Regulation". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 367.1908 (Dec. 13, 2009), pp. 4923–4940.

[11] Sergey Ioffe and Christian Szegedy. "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift". In: *International Conference on Machine Learning*. pmlr, 2015, pp. 448–456.

[12] Michael C. Joiner and Albert J. van der Kogel. *Basic Clinical Radiobiology*. CRC Press, Aug. 28, 2018.

[13] Diederik P. Kingma and Jimmy Ba. "Adam: A Method for Stochastic Optimization". 2014.

[14] Patrick A. Kupelian et al. "Effect of Increasing Radiation Doses on Local and Distant Failures in Patients with Localized Prostate Cancer". In: *International Journal of Radiation Oncology* Biology* Physics* 71.1 (2008), pp. 16–22.

[15] Jin Tae Kwak et al. "Prostate Cancer: A Correlative Study of Multiparametric MR Imaging and Digital Histopathology". In: *Radiology* 285.1 (Oct. 2017).

[16] Philippe Lambin et al. "Radiomics: Extracting More Information from Medical Images Using Advanced Feature Analysis". In: *European Journal of Cancer* 48.4 (Mar. 1, 2012), pp. 441–446.

[17] Hye Won Lee et al. "Novel Multiparametric Magnetic Resonance Imaging-Based Deep Learning and Clinical Parameter Integration for the Prediction of Long-Term Biochemical Recurrence-Free Survival in Prostate Cancer after Radical Prostatectomy". In: *Cancers* 15.13 (2023), p. 3416.

[18] Chiara Nicolò et al. "Machine Learning and Mechanistic Modeling for Prediction of Metastatic Relapse in Early-Stage Breast Cancer". In: *JCO Clinical Cancer Informatics* 4 (Nov. 2020), pp. 259–274.

[19] Perrine Paul-Gilloteaux et al. "Optimizing Radiotherapy Protocols Using Computer Automata to Model Tumour Cell Death as a Function of Oxygen Diffusion Processes". In: *Scientific Reports* 7.1 (May 23, 2017), p. 2280.

[20] Carlos Sosa-Marrero et al. "Towards a Reduced in Silico Model Predicting Biochemical Recurrence after Radiotherapy in Prostate Cancer". In: *IEEE Transactions on Biomedical Engineering* 68.9 (2021), pp. 2718–2729.

[21] Nitish Srivastava et al. "Dropout: A Simple Way to Prevent Neural Networks from Overfitting". In: *The Journal of Machine Learning Research* 15.1 (Jan. 1, 2014), pp. 1929–1958.

[22] Hyuna Sung et al. "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries". In: *CA: A Cancer Journal for Clinicians* 71.3 (2021), pp. 209–249.

[23] James L Tatum. "Hypoxia: Importance in Tumor Biology, Noninvasive Measurement by Imaging, and Value of Its Measurement in the Management of Cancer Therapy". In: *International Journal of Radiation Biology* 82.10 (Jan. 2006), pp. 699–757.

[24] Joost J.M. Van Griethuysen et al. "Computational Radiomics System to Decode the Radiographic Phenotype". In: *Cancer Research* 77.21 (Nov. 1, 2017), e104–e107.

[25] Alex Zwanenburg et al. "The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping". In: *Radiology* 295.2 (May 2020), pp. 328–338.