npj | digital medicine

Article

Published in partnership with Seoul National University Bundang Hospital

https://doi.org/10.1038/s41746-025-01890-x

Physics-informed machine learning digital twin for reconstructing prostate cancer tumor growth via PSA tests

Daniel Camacho-Gomez[1,2], Carlos Borau[1,3], Jose Manuel Garcia-Aznar[1], Maria Jose Gomez-Benito[1], Mark Girolami[2,4] & Maria Angeles Perez[1]

Existing prostate cancer monitoring methods, reliant on prostate-specific antigen (PSA) measurements in blood tests often fail to detect tumor growth. We develop a computational framework to reconstruct tumor growth from the PSA integrating physics-based modeling and machine learning in digital twins. The physics-based model considers PSA secretion and flux from tissue to blood, depending on local vascularity. This model is enhanced by deep learning, which regulates tumor growth dynamics through the patient's PSA blood tests and 3D spatial interactions of physiological variables of the digital twin. We showcase our framework by reconstructing tumor growth in real patients over 2.5 years from diagnosis, with tumor volume relative errors ranging from 0.8% to 12.28%. Additionally, our results reveal scenarios of tumor growth despite no significant rise in PSA levels. Therefore, our framework serves as a promising tool for prostate cancer monitoring, supporting the advancement of personalized monitoring protocols.

Prostate cancer (PCa) is one of the most prevalent forms of cancer that affect men[1]. It is estimated that prostate cancer accounted for 1.4 million new cases globally and resulted in more than 370,000 deaths in 2020 alone[2]. Prostate cancer is characterized by the uncontrolled growth and division of luminal cells within the prostate gland. Over time, these cancerous cells can invade nearby tissues and potentially spread to other parts of the body, mainly to the bones, lymph nodes, liver, and lungs[3–6] in a process called metastasis[7]. Therefore, predicting the evolution of prostate cancer is essential for timely detection of growth and halting the expansion of the disease.

The diagnosis of prostate cancer typically relies on the Prostate Imaging Reporting and Data System (PI-RADS), which assigns scores on a five-point scale to lesions observed in magnetic resonance imaging (MRI) images[8], and the Gleason score, which assesses the differentiation of cells in biopsy samples[9], serving as an indicator of tumor prognosis[10]. After confirming tumor presence through MRI and biopsy, prostate cancer monitoring is commonly based on the prostate-specific antigen (PSA) biomarker[11]. The PSA is a protein produced by both normal and cancerous cells within the prostate gland. Its main function is to liquefy semen, aiding in the mobility and transportation of sperm during ejaculation[12]. PSA levels in the blood have been extensively employed as a biomarker for both the detection and ongoing monitoring of prostate cancer. Elevated levels of PSA may indicate various prostate conditions, including prostate cancer[13]. While the exact reason for increased PSA levels in prostate cancer is not fully understood, it is believed that cancerous cells can disrupt the normal architecture of the prostate gland, leading to increased production and leakage of PSA into the bloodstream. Consequently, measuring PSA levels in the blood can help in both the detection and the monitoring of prostate cancer. Yet, tumor progression and growth often occur without a significant rise in PSA levels, thereby obscuring the prognosis of the tumor. The limited specificity and sensitivity of the PSA as biomarker[14–16] are linked with poor diagnosis, as well as treatment and screening-related adverse effects[17]. Consequently, it is essential to deepen our understanding of the relationship between PSA levels and tumor development.

Significant improvements have been made in the comprehension of prostate cancer through computational models[18,19]. Jain et al.[20] studied through a system of differential equations the progression of prostate cancer under continuous and intermittent antiandrogen treatment regimes, emphasizing the heterogeneous nature of the disease and adjusting personalized parameters to the average patients. Colli et al.[21] analyzed prostate cancer growth with chemotherapy and antiangiogenic therapy effects, suggesting that, while cytotoxic drugs may suffice to treat mild tumors, the combination of cytotoxic action with the reduction of intratumoral nutrient availability is essential to kill the aggressive tumor. Lorenzo et al.[22] formulated a model for the PSA dynamics after radiotherapy treatment,

[1]Department of Mechanical Engineering, Multiscale in Mechanical and Biological Engineering (M2BE), Aragon Institute of Engineering. Research (I3A), University of Zaragoza, Zaragoza, Spain. [2]Department of Engineering, University of Cambridge, Cambridge, UK. [3]Centro Universitario de la Defensa de Zaragoza, Zaragoza, Spain. [4]The Alan Turing Institute, London, UK. e-mail: angeles@unizar.es

showing good agreement with the patients' PSA. However, these works do not incorporate the anatomical and physiological characteristics of the prostate gland itself, nor do they fully capture personalized aspects of prostate cancer. By considering the specific features of the prostate through image-based biomarkers, such as its shape, vascularization, or the spatial distribution of cancer cells, computational models could provide more detailed and accurate insights into tumor growth dynamics and treatment response. This enhanced precision would allow for more personalized and effective treatment strategies that would ultimately benefit the patients.

Advancing along these lines, digital twins are emerging as a promising approach. Digital twins[23–25] are virtual representations of physical objects, processes, or systems, that mimic the behavior, characteristics, and functionalities of their physical counterparts[26]. In the context of biology, digital twins offer a means to integrate the anatomical and physiological characteristics of patients' tissues, together with medical records, and other relevant information into computational models[27–29]. Thus, digital twins are a powerful tool for providing a more accurate and personalized understanding of tumor growth and treatment response in prostate cancer, allowing for a more comprehensive representation of the disease. Various authors have made efforts in this direction. For example, Boubaker et al.[30] perform a finite element simulation of the interaction of the prostate and the surrounding organs to predict their deformation to improve radiotherapy delivery. Lorenzo et al.[31] simulate a patient's specific prostate generated from computed tomography (CT) images. Lorenzo et al.[32] extend the model to include the mechanical interaction of prostate cancer and benign prostatic hyperplasia. Hadjicharalambous et al.[33] data-driven framework for personalized brachytherapy simulation, integrating prostate biomechanical modeling, catheter insertion, and tissue deformation to optimize prostate cancer treatment planning. Despite significant advances, the aforementioned works still present several limitations. First, using only CT imaging to determine prostate geometry falls short in precisely identifying prostate cancer and pinpointing its location. Second, these models lack the capability to incorporate detailed information at the cellular level within the prostate tissue, such as cell density or cellularity, crucial for accurately simulating tumor growth. Moreover, the lack of perfusion data eliminates the possibility of accounting for the spatial arrangement of the vascular network. Recent work by Lorenzo et al.[34] address some of these limitations by leveraging multiparametric MRI to account for tumor cellularity in prostate cancer. Since tumor growth is driven directly by the proliferation of cancer cells, incorporating cell concentration is important for accurately capturing its progression and development. Advancing in this direction holds the potential to enhance patient-specific representation and the biological relevance of simulations, thereby better supporting clinicians in personalized disease management.

In this work, we present a physics-informed machine-learning framework to predict the prognosis of prostate cancer using digital twins. Initially, we generate patient-specific digital twins of the prostate from T2-weighted MRI images, incorporating cellular-level data such as cellularity, vascularization, and tumor location. These digital twins are then introduced into the computational framework to obtain the tumor's growth over time. In this physics-based model of prostate cancer, we consider the production of tissue PSA by cancer cells and propose a mathematical model to calculate serum PSA levels based on the prostate's vascularization. Consequently, serum PSA dynamics are primarily driven by tumor cell proliferation, and the actual variations in serum PSA levels arise from the complex interplay among tumor cell concentration, local vascularity, and the production and decay of PSA in both tissue and blood. To determine the tumor's growth, we incorporate a patient-agnostic deep-learning method that regulates its dynamics in the physics-based model. To achieve this, it considers the spatial interaction of both MRI-derived variables, such as cellularity, vascularization, and tumor location, as well as simulation-derived variables, including tissue PSA and the fraction of proliferating tumor cells, along with clinical patient data such as serum PSA levels and time points. We calibrate the physics-based model to reflect patient-specific tumor growth dynamics by integrating a serum PSA measurement alongside one additional follow-up MRI data point. This process enables the determination of personalized physical parameters, enhancing the accuracy of tumor prognoses. After this calibration, no further adjustments are needed, and tumor growth reconstructions can be made from serum PSA follow-up assessments. We illustrate our approach with two real patients sourced from the La Fe Hospital (HULAFE) database. Our findings demonstrate the efficacy of our methodology in accurately predicting the growth of prostate cancer, based on patient follow-ups using serum PSA data. This approach helps reduce the uncertainty of tumor growth during periods when monitoring relies solely on PSA measurements, enabling more informed clinical decisions. Ultimately, this method enhances the interpretability of PSA monitoring and contributes to more personalized patient care protocols.

Results

Physics-informed machine learning digital twin framework for prostate cancer

We present a computational framework for prostate cancer in digital twins, combining a physics-based model for tumor growth with a machine-learning model that regulates its growth dynamics (Fig. 1). This patient-specific digital twin of the prostate is generated from the T2-weighted image (T2WI) sequences on magnetic resonance imaging (MRI), which include Diffusion Weighted (DW) and Dynamic Contrast Enhanced (DCE) image sequences, as well as the tumor segmented by expert radiologists from

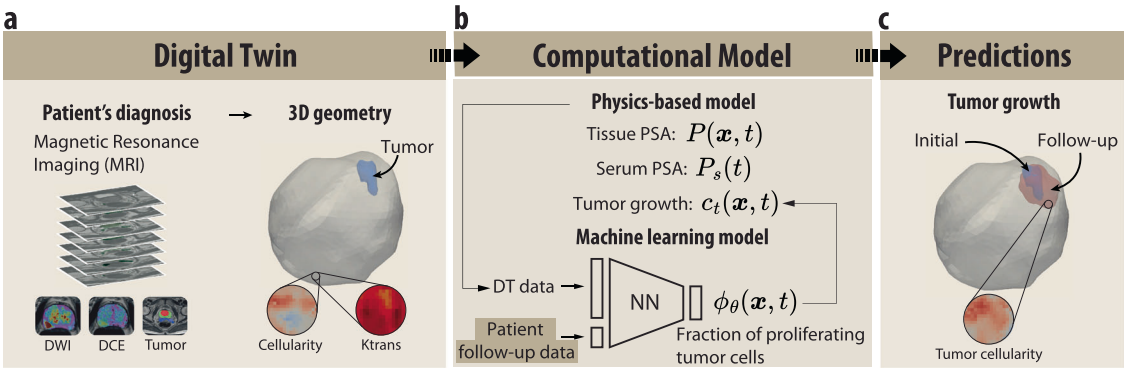

Fig. 1 | Prostate cancer digital twin framework. a The digital twin geometry is reconstructed from the T2-weighted image sequences on magnetic resonance imaging, including the spatial distributions of cellularity from the apparent diffusion coefficient maps in the diffusion-weighted images, k_{trans} from dynamic contrast-enhanced sequences, and the tumor mask. **b** The computational model consists of two main parts. First, a physics-based model simulates the evolution of tissue PSA $P(\boldsymbol{x}, t)$, serum PSA $P_s(t)$, and tumor growth $c_t(\boldsymbol{x}, t)$. Second, a machine learning model based on a neural network (NN) determines the fraction of proliferating tumor cells $\phi_\theta(\boldsymbol{x}, t)$ in the tumor growth equation based on the data from the digital twin and the patient follow-up serum PSA blood test. **c** The outcome of the model is the patient's tumor growth from diagnosis to the follow-up date.

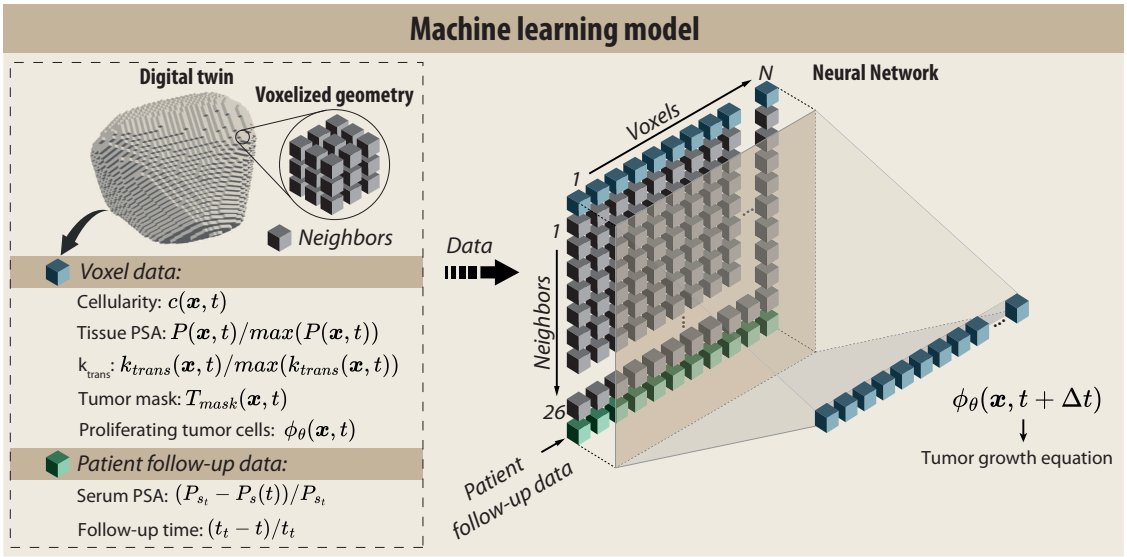

Fig. 2 | Machine learning model. This model approximates the evolution of the fraction of proliferating tumor cells in each location ($\phi_\theta(\boldsymbol{x},t)$). To do that, it employs data from the digital twin and the patient's follow-up serum PSA test to construct an input matrix. The matrix is then processed through a NN to determine the fraction of proliferating tumor cells. Spatial interactions are considered by incorporating data from neighboring voxels. This process is repeated at every time step in the physics-based simulation.

HULAFE, all common procedures in the evaluation and diagnosis of prostate cancer (Fig. 1a). Using these MRI sequences, a 3D voxelized geometry is generated, containing the patient's data such as the cellularity $c(\boldsymbol{x},t)$, the spatial distribution of vascularization considered through $k_{trans}(\boldsymbol{x})$, which provides crucial information about tissue perfusion and permeability, and the tumor binary mask $T_{mask}(\boldsymbol{x},t)$ (see *Imaging and postprocessingMRI sequences* in Methods for further details).

We propose a physics-informed machine learning model to simulate the growth of prostate cancer (Fig. 1b). The purpose of the physics-based model is to represent the biological processes involved in prostate cancer within the digital twin framework to predict the growth of the diagnosed tumor. In this model, we simulate the tissue PSA per unit volume of prostatic tissue $P(\boldsymbol{x},t)$ as a consequence of the PSA leakage from cancer cells. Additionally, we incorporate the exchange of PSA between the tissue and the bloodstream. To achieve this, we consider the capillaries located in the tissue, represented through $k_{\text{trans}}(\boldsymbol{x})$. This exchange depends on the concentration difference between the tissue PSA distribution ($P(\boldsymbol{x},t)$) and the serum PSA ($P_s(t)$), as well as the permeability of the capillaries. We also consider the natural decay of the tissue PSA. The serum PSA ($P_s(t)$) is obtained by integrating the flux between the bloodstream and the tissue, also considering its natural decay. Finally, we model the evolution of the concentration of tumor cells $c_t(\boldsymbol{x},t)$, which represents tumor growth. This tumor cell concentration is responsible for the production of tissue PSA, and therefore, for the variations in the simulated serum PSA concentration. Hence, the dynamics of tumor growth cause the variations in serum PSA (see *Physics-based model for prostate cancer* in Methods for details of the physics-based model).

In this physics-based model, we integrate a machine-learning model to determine the dynamics of tumor growth ($c_t(\boldsymbol{x},t)$) based on the patient's serum PSA tests from their follow-ups. This machine learning model approximates, through a fully connected neural network (NN), the fraction of proliferating tumor cells $\phi_\theta(\boldsymbol{x},t)$, which is incorporated into the equation describing tumor growth (see next section for further details). To accomplish this, the deep-learning model receives data from the physics-based model and the patient follow-up data based on serum PSA blood tests. Consequently, it regulates the tumor growth dynamics in the physics-based model so the simulated PSA accurately reproduces the patient's serum PSA values. Finally, we predict the tumor's growth from the time of diagnosis to the follow-up date which is responsible for the observed serum PSA variations.

Machine learning model for tumor growth dynamics

We propose a deep-learning algorithm to determine tumor growth by regulating the fraction of proliferating tumor cells ($\phi_\theta(\boldsymbol{x},t)$) in Eq. (3) (Fig. 2). To approximate this function that controls tumor growth, we employ data from the digital twin and the patient. First, for each voxel of the digital twin, we extract the following data: cellularity ($c(\boldsymbol{x},t)$), normalized tissue PSA ($P(\boldsymbol{x},t)/\max(P(\boldsymbol{x},t))$), normalized k_{trans} ($k_{\text{trans}}(\boldsymbol{x})/\max(k_{\text{trans}}(\boldsymbol{x}))$), the tumor mask ($T_{\text{mask}}(\boldsymbol{x},t)$), and the fraction of proliferating tumor cells ($\phi_\theta(\boldsymbol{x},t)$). Second, we use as patient data the serum PSA measured in the follow-up blood test ((P_{s_t})), which we aim to replicate. This is compared to the current simulated PSA at the time the NN is called (P_s), and normalized as $(P_{s_t}-P_s(t))/P_{s_t}$. Similarly, we consider the time when the follow-up blood test is conducted (t_t), compared to the current simulation time t, and normalize it as $(t_t-t)/t_t$.

Then, we construct the input matrix for the neural network using this data. For each voxel of the digital twin, we gather its extracted data and concatenate it with the data from adjacent neighboring voxels and the data from the patient, following a standardized sequence. Due to the voxel-based geometry, each voxel is typically surrounded by 26 neighbors, except for edge voxels. For these edge voxels, we insert -1 values in the positions of the non-existent neighboring voxels to substitute the missing data (see *Machine learning input matrix construction* in Supplementary Material). This method enables us to estimate the proportion of proliferating cells not solely from individual voxel data but by also considering the spatial context of surrounding voxels, thus incorporating spatial dynamics into our analysis. Additionally, it is patient-agnostic, meaning it is independent of the patient's specific digital twin geometry, not requiring changes in architecture or retraining due to geometrical changes in the patient's prostate. Finally, this matrix is fed forward to the NN to obtain the fraction of proliferating tumor cells for the next time step $\phi_\theta(\boldsymbol{x},t+\Delta t)$. This data extraction, input matrix construction and feed-forwarding is repeated at every time step of the physics-based simulation.

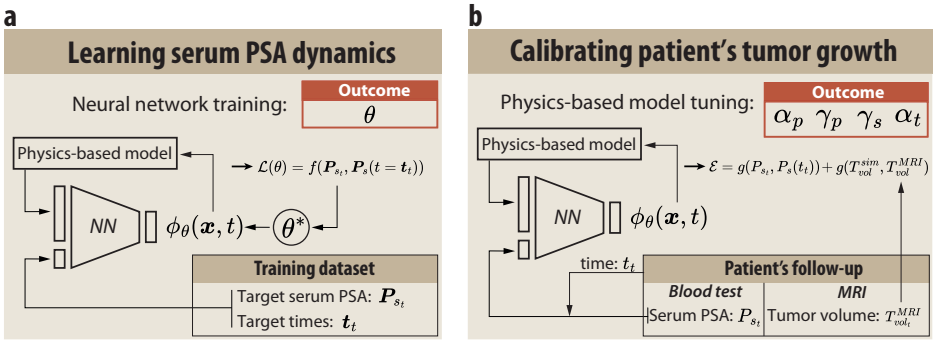

Fig. 3 | Calibration of the framework to reproduce patient-specific tumor growth. a Training of the NN to learn to replicate serum PSA dynamics by controlling tumor growth. This involves conducting simulations with each serum PSA dataset, followed by minimizing the loss between the simulated and dataset serum PSA. **b** Calibration of the patient-specific parameters of the physics-based model for personalized predictions. The serum PSA is integrated into the computational model to predict tumor growth at the follow-up. Subsequently, the error between the predicted tumor volume and serum PSA and the clinical outcomes is minimized.

Calibration for patient-specific tumor growth dynamics

The calibration of the framework involves two phases. The initial phase entails training the patient-agnostic NN of the machine learning model to learn serum PSA dynamics. The second phase involves obtaining patient-specific parameters for the physics-based model for personalized tumor growth (Fig. 3).

In the initial phase, the weights of the NN (θ) are obtained to learn serum PSA dynamics through adequate spatial and temporal regulation of the fraction of proliferating tumor cells (Fig. 3a). It is important to note that this training step uses a single case of initial conditions (either synthetic or from an actual patient) which is subjected to a multi-objective simulation: reach a certain blood level of PSA (target serum PSA dataset ($\boldsymbol{P}_{s_t}$)) within a defined period (target times dataset ($\boldsymbol{t}_t$)). Then, we calculate the loss between the simulated serum PSA and the target serum PSA at the target times. This loss is then backpropagated to optimize the NN weights (see *Model calibration* in the Methods section). In this way, the NN learns the underlying relationships between PSA production and the spatial distribution of cellularity, k_{trans}, and tumor growth that result in the different evolutions of the serum PSA. This deep-learning model addresses the inverse problem of predicting tumor growth from serum PSA at a certain date. At this point, the physics-based parameters are not required to be patient-specific, so this step is performed only once during the model setup.

Once the machine learning model is set up with the trained NN, the next phase involves obtaining the patient's specific parameters of the physics-based model to make personalized predictions from their blood test (Fig. 3b). This ensures that the predicted tumor aligns with the patient's actual tumor when replicating serum PSA levels. For this purpose, we conduct a simulation incorporating the patient's follow-up data as targets for the NN, resulting in the personalized growth of the tumor and evolution of the serum PSA. Subsequently, we calculate the error between the predicted volume and serum PSA and the actual values obtained from the follow-up and minimize it by optimizing the parameters of the physics-based model (refer to *Model calibration* in the Methods section). After determining the patient-specific parameters of the physics-based model with a single MRI follow-up, we can make further predictions of tumor growth using only subsequent serum PSA follow-ups.

Unveiling patient-specific tumor growth

To demonstrate the potential of the framework for predicting patient-specific prostate cancer growth, we utilize data from two anonymized patients from HULAFE, whom we will henceforth name Patient A and Patient B. Patient A was diagnosed at the age of 68, with a PI-RADS category 5 and a Gleason score of $3 + 3$. Patient B, diagnosed at the age of 60, had a PI-RADS category 4. Patient B's biopsy reveals a Gleason score of $3 + 3$ and a 2% concentration of tumor cells. Since the concentration of tumor cells was not measured in the biopsy of Patient A, we assume it to be the same as that of Patient B.

We apply the calibration method presented in Fig. 3b to determine the patient-specific parameters of the physics-based model which replicate the serum PSA and tumor growth from the diagnosis to the first follow-up (Fig. 4a). Notably, the obtained parameters outline substantial differences between the two patients. The tissue PSA production rate (α_p) is 49.05% lower in the case of Patient A compared to Patient B. Conversely, the tumor growth rate (α_t) is 124.07% higher in Patient A than in Patient B. This indicates that while Patient A's tumor secretes less PSA, it exhibits a higher growth rate. Regarding decay rates, Patient B has higher decay rates both in tissue and serum, although our parametric analysis suggests that the serum PSA decay rate (γ_s) is the least influential parameter in the computational model.

We illustrate the tumor growth of Patient A from diagnosis to follow-up (Fig. 4b). The serum PSA level at diagnosis is 6.60 ng/mL with a corresponding tumor volume of 579 mm^3. Subsequently, after 393 days, the clinical tumor volume increases to 862 mm^3, accompanied by a rise in serum PSA to 7.40 ng/mL. We reproduce this clinical growth with our computational model, with a simulated serum PSA of 7.70 ng/mL and a tumor volume of 887 mm^3, resulting in a relative error of 2.82% compared to the clinical tumor volume at follow-up 1. Furthermore, our model not only accurately replicates the serum PSA levels and tumor volume, but also captures the shape of the tumor as shown in the comparison of the tumors in Fig. 4b. To quantitatively assess the differences in tumor shape between the data and the simulation, we calculate the principal moments of inertia. This analysis demonstrates that our model accurately captures and reproduces the tumor shape over time.

We also depict the tumor growth of Patient B from diagnosis to the follow-up 1 after 710 days (Fig. 4c). In this case, the diagnosed tumor volume is 249 mm^3 with a serum PSA level of 7.64 ng/mL. Subsequently, at the follow-up 1, the clinical serum PSA increases to 10.85 ng/mL and the tumor volume to 407 mm^3. Here, we also replicate both the serum PSA (10.94 ng/mL) and tumor volume (404 mm^3), with relative errors of 0.7% and 0.8%, respectively. Moreover, the simulated tumor mimics the real shape of the tumor at the follow-up, as indicated by the principal moments of inertia in Fig. 4c.

Interestingly, we can observe that the patients have different tumor and PSA dynamics. Patient A's tumor volume increased from diagnosis to follow-up 1 at a rate of 0.72 mm^3/day with only a 2.0×10^{-3} ng/mL/day increase in serum PSA, while Patient B's tumor volume and serum PSA increased at rates of 0.22 mm^3/day and 4.5×10^{-3} ng/mL/day, respectively. Thus, the growth rate of Patient A's tumor is 3.27 times higher, while the serum PSA increase rate is 2.25 times lower. These clinical prognoses of the patients align with the calibration results of the physics-based parameters in Fig. 4a, which indicate that Patient A's tumor grows faster but secretes less than Patient B's tumor. Consequently, our model demonstrates the capability to precisely reproduce prostate tumor growth, obtaining patient-specific parameters that account for interpatient prostate cancer heterogeneity.

Reconstructing long-term patient-specific tumor growth from serum PSA tests

Our framework's greatest contribution is the capability to predict long-term tumor growth solely from serum PSA measurements in subsequent

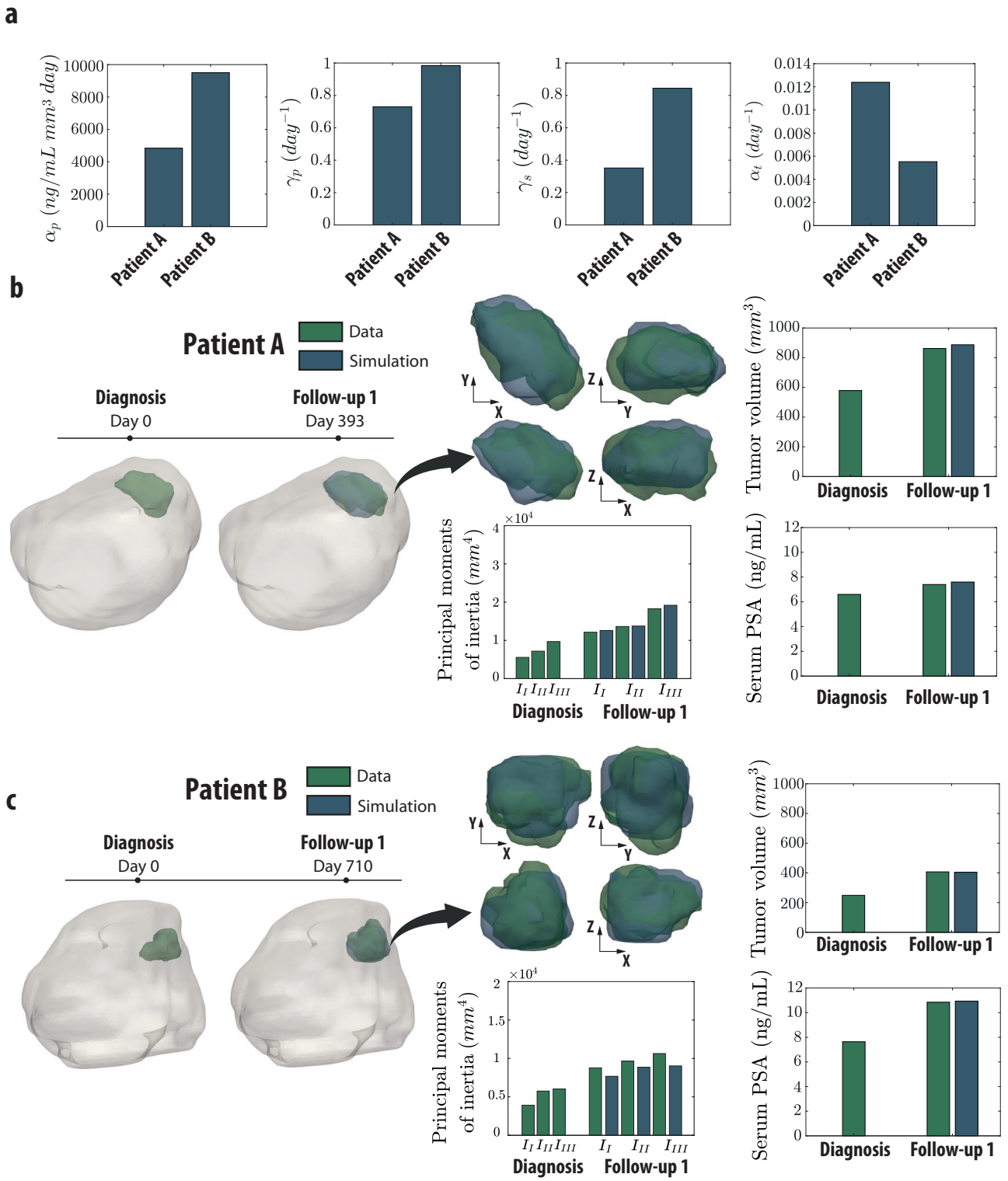

Fig. 4 | Patient-specific prostate tumor growth. a Calibration of the physics-based parameters for Patients A and B. **b** Tumor growth in Patient A. **c** Tumor growth in Patient B. For both patients, representations of the actual prostate geometry with the tumor at diagnosis and comparisons between the clinical and simulated tumors at follow-up are provided. Additionally, comparisons between tumor volume, serum PSA, and principal moments of inertia of the tumor data from clinical and simulation records at follow-up are presented.

follow-up blood tests, potentially reducing the need for frequent MRIs and thereby optimizing patient care. To do this, we fix the parameters of the physics-based model obtained from each patient and infer tumor growth from serum PSA levels at their subsequent follow-ups (Fig. 5). Using this approach, we predict the clinical evolution of Patient A (Fig. 5a). The clinical follow-ups for Patient A's serum PSA show a 5.94% increase from Follow-up 1 (7.40 ng/mL) to Follow-up 2 (7.84 ng/mL) at day 582. Subsequently, in Follow-up 3, the serum PSA increases by 7.22% (8.45 ng/mL) from Follow-up 2. We incorporate these measurements into our computational model and accurately replicate the serum PSA evolution, obtaining the growth of

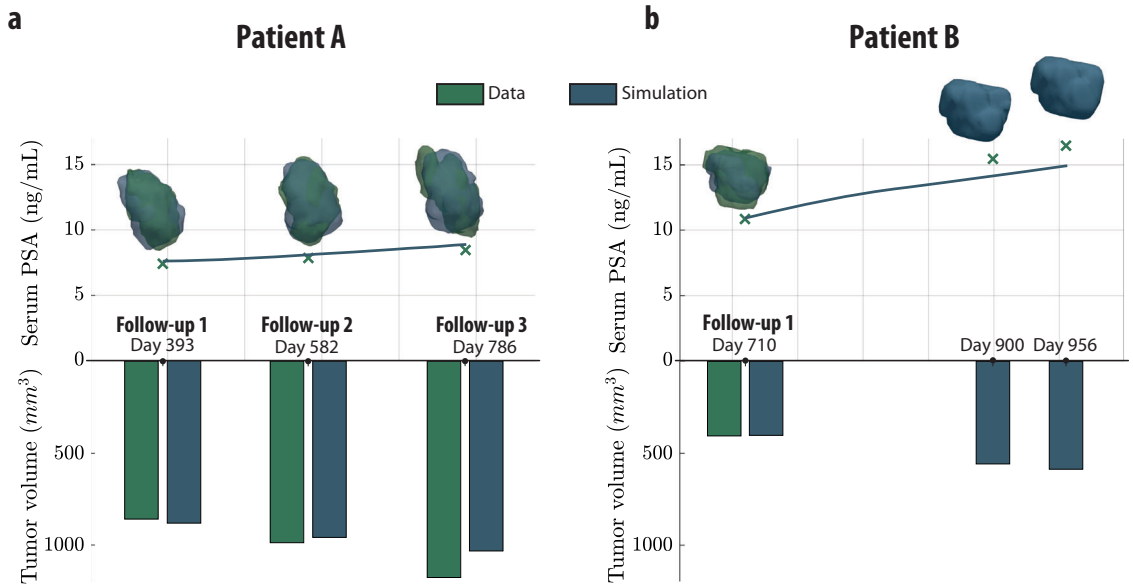

Fig. 5 | Reconstruction of patient-specific prostate tumor growth. a Predictions for Patient A. **b** Predictions for Patient B. For both patients, the clinical serum PSA measurements from follow-up blood tests and the simulated evolution of serum PSA are shown. Additionally, a comparison between the clinical and reconstructed tumor volumes is provided (no subsequent image-based follow-up was conducted for Patient B after day 710).

the tumor that produces these variations. For Patient A, follow-ups included both blood tests and MRI scans. Thus, we can compare not only the PSA levels but the tumors to further validate our predictions. At follow-up 1, the clinical tumor volume was 862 mm^3. By follow-up 2, it increases to 991 mm^3 (a 13% increase from follow-up 1), and at follow-up 3, it further grows to 1181 mm^3 (a 16% increase from follow-up 2), confirming that the nature of Patient A's tumor produces minimal PSA elevation despite continuous tumor growth. As shown in Fig. 5a, our predictions align with the actual evolution of Patient A, demonstrating relative errors of 3.10% and 2.90% for PSA and tumor volume, respectively, at follow-up 2, and 4.47% and 12.28% at follow-up 3. These results highlight the potential of our model to infer patient-specific parameters and future tumor growth and shape (see Supplementary Fig. 2a in Supplementary Material for the principal moments of inertia of the tumors) solely from serum PSA data (see Supplementary Movie 1 for the simulation of Patient A's tumor growth from diagnosis to follow-up 3).

We also predict the clinical evolution of Patient B based on blood tests conducted on days 900 and 956 (Fig. 5b). Patient B's follow-ups rely solely on serum PSA tests (no MRI data), so predicted tumor volume at follow-ups 2 and 3 cannot be compared with imaging data. In this case, the serum PSA levels increase by 29.82% on day 900 compared to Follow-up 1, followed by a further increase of 6.53% on day 956. Our approach predicts a tumor volume increase of 38.61% on day 900 compared to Follow-up 1, followed by an additional increase of 5.18% over the next 56 days, consistent with the percentage increases in serum PSA (see Supplementary Movie 2 for the simulation of Patient B's tumor growth from diagnosis to day 956 and Supplementary Fig. 2b in Supplementary Material for the principal moments of the inertia of the tumors). We confirm that the tumor reconstructions for both patients remain consistent, despite variations in serum PSA data arising from uncertainties in its measurement due to biological or instrumental factors (refer to *Tumor volume sensitivity to serum PSA variations* in Supplementary Material). We conclude that tumor growth in Patient B is directly correlated with increased serum PSA levels, allowing for better detection compared to Patient A.

Discussion

We have developed an innovative framework consisting of a digital twin for the prostate with a physics-informed machine-learning model to predict the prognosis of prostate cancer. Through this model, we can predict a patient's specific prostate tumor growth based on an initial MRI and subsequent PSA blood tests. This approach involves employing a deep learning model to regulate tumor dynamics within the physics-based model, determining the fraction of proliferating tumor cells in each location of the prostate. The machine-learning model we have developed takes into account the three-dimensional spatial interaction of physiological variables in the prostate, including cellularity levels, vascularization, tumor location, as well as simulation-derived variables, including tissue PSA and the fraction of proliferating tumor cells, along with clinical patient data such as serum PSA levels and time points, to capture the spatial-physical interactions driving disease progression. Importantly, our approach relies on a deep NN independent of the prostate's geometry (geometry agnostic). In our proposed method, we voxelize the prostate's geometry, standardizing voxel connections and, consequently, the number of neighboring voxels. To obtain the fraction of proliferating tumor cells in a voxel, we incorporate data from that voxel, its neighboring voxels, and the patient's follow-up data. Since the connections between voxels remain constant regardless of the specific prostate geometry, we can maintain a consistent input structure for the NN. Therefore, this allows the NN to be applied effectively across patients without changing architectures or requiring retraining due to geometrical changes, all while considering spatial interactions.

We have preliminarily validated our methodology by calibrating and then reconstructing tumor growth in two real patients, showcasing accurate simulated PSA evolutions and robust tumor growth reconstructions despite PSA uncertainty (refer to *Tumorvolumesensitivity to variations in PSA* in Supplementary Material), although uncertainty quantification will be explored in future work to further enhance its robustness. We have demonstrated its potential to accurately reproduce not only clinical tumor volume and shape over time (refer to Supplementary Fig. 2 in Supplementary Material for comparisons of the principal moments of inertia of the tumors), but also to obtain patient-specific parameters of the physics-based model that accounts for interpatient prostate cancer heterogeneity (Fig. 4a) This determination of the physics-based patient parameters enabled us to characterize two distinct patient profiles. In one scenario (Patient B), tumor growth is directly correlated with the increase in serum PSA levels, which could be readily detected through blood tests. Conversely, in Patient A tumor grows without a significant elevation in PSA levels. This mechanism elucidates the well-known lack of sensitivity of PSA as a standalone

biomarker for screening[14–16]. With further characterization of these parameters across a broader patient population, we could potentially establish correlations between the model's parameters and tumor aggressiveness metrics, such as the Gleason score. Therefore, our model allows for the identification of patients with hidden tumor growth, where serum PSA levels remain stable despite tumor development. This addresses the challenges of monitoring the disease through serum PSA in periods between MRIs, enabling more timely and personalized interventions.

Furthermore, our computational framework allows for the reconstruction of tumor growth even in cases with multiple primary tumors (see *Multiple primary foci tumor growth reconstruction* in Supplementary Material) from periodic blood tests, offering the potential to predict long-term outcomes over periods spanning months or even years with remarkable computational speed (a simulation spanning 956 real-world days takes ~23 s to complete (see *Model implementation* in Methods for further details)). This efficiency enables rapid prognosis evaluations, an important feature to assist clinicians in making real-time decisions for personalized cancer management. In our proposed physics-based model, we integrated biologically relevant aspects of prostate cancer, including the increase in tissue PSA caused by the PSA leakage from cancer cells. Under normal conditions, PSA produced by healthy cells is delivered to the urethra, however, in cancer pathological conditions, tumor cells proliferate without generating the necessary ducts for PSA delivery to the urethra, resulting in leakage. Depending on the prostate's vascularization, as described by the k_{trans} distribution, this abnormal concentration of PSA passes into the bloodstream. Thus, k_{trans} reflects the spatial distribution of vascularization, tissue perfusion, and permeability in dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) within our physics-based model[35]. The flux of fluid and substances between tissue and blood through the k_{trans} distribution has already been modeled[36–38]. However, to the best of our knowledge, this is also the first computational work in prostate cancer considering the intravasation of PSA from the tissue to the blood while taking into account the spatial distribution of vascularization through the k_{trans} distribution.

We proposed a mathematical methodology to simulate these biological processes that occur on different timescales and perform long-term simulations with reduced computational burden. In this model, certain assumptions were made. We inferred the concentration of tumor cells from cellularity and the percentage of tumor cells from biopsy results. If the biopsy is randomly sampled from the prostate, there may be samples without tumor, leading to an underestimation of the percentage. Ideally, our estimation would be more precise if the biopsy were guided to the tumor location. Additionally, the updating algorithm of the tumor mask considers an upper threshold for expansion, representing the carrying capacity that the voxel cannot exceed in that area. Although this upper threshold aligns with evidence that the cell population reaches a maximum tissue carrying capacity[39,40], other algorithms could be introduced, such as considering this threshold as a stochastic parameter or increasing a certain percentage of the initial concentration. Nonetheless, we evaluated the effect of the carrying capacity on tumor reconstructions and demonstrated the model's ability to replicate tumor growth with different values by recalibrating the physics-based parameters. Despite recalibrating the physics-based parameters for the new carrying capacities, the trends in the parameters across patients remain consistent (see *Carrying capacity sensitivity analysis* in Supplementary Material.) A limitation of this study is that we did not perform longitudinal image registration for each patient. Incorporating this in future work would enhance model validation by enabling voxel-wise comparisons of tumor growth and spatial maps, such as cellularity concentration, across multiple time points. Furthermore, clinical segmented tumors may have inaccuracies, resulting in small deviations in clinical tumor volumes, which can be minimized with the utilization of available automated segmentation tools for prostate cancer[41,42] and for geometry reconstruction[43]. Regarding the machine learning model, larger datasets of serum PSA at different time points could be employed, enhancing the capacity to reproduce various dynamic behaviors or training on different geometries. However, despite these simplifications, we have demonstrated the predictive capabilities of our model in determining the patient's prognosis. Thus, the proposed model framework can be applied to patients with prostate cancer. To simulate patients with concomitant benign prostatic hyperplasia or episodes of prostatitis, we should include other factors not included in the current model.

Therefore, this framework, which integrates a physics-informed machine learning model into the prostate digital twin, offers a novel approach to improving the monitoring of prostate cancer. By providing a clearer understanding of tumor growth trends from serum PSA blood tests, it can guide clinicians in determining the optimal timing for further diagnostic actions, such as MRI or biopsy.

Methods

Imaging and postprocessing MRI sequences

The 3D geometry of the prostate is reconstructed from the T2-weighted MRI sequences using the *im2mesh* library[43] and then meshed with voxels. Then, the tumor mask $T_{mask}(\boldsymbol{x}, t)$ is interpolated onto this voxelized geometry from the binary mask of the tumor segmented by clinicians in the T2W MRI sequences. The cellularity $c(\boldsymbol{x}, t)$ is calculated from the Apparent Diffusion Coefficient (ADC) maps extracted from DW sequences[44], which measure the diffusion of water molecules, following[45] and then interpolated in this voxelized mesh:

$$c(\boldsymbol{x}, t) = \frac{ADC_w - ADC(\boldsymbol{x})}{ADC_w - ADC_{min}}, \tag{1}$$

with $ADC_w = 3 \times 10^{-3}\,\mathrm{mm^2/s}$ the magnitude of diffusion of free water molecules from ref. 45 and ADC_{min} the minimum value of the ADC maps of the MRI at time t.

Finally, we employ the Standard Tofts Model as the pharmacokinetic model to describe the behavior of gadolinium in the DCE-MRI sequences[35] and characterize tissue vascularization. This model helps assess key parameters such as tissue perfusion, vascular permeability, and the rate of contrast washout. By adjusting the parameters of the model to fit the concentration of the contrast agent in the tissue over time, derived from the time-intensity curves, we can obtain $k_{trans}(\mathbf{x})$ and interpolate it onto the geometry, which represents the rate at which the contrast agent moves from the blood vessels into the tissue, influenced by both blood flow and vessel permeability[46].

Imaging acquisition details for each MRI performed on the patients can be found in Supplementary Tables 1 and 2.

Physics-based model for prostate cancer

The physics-based model has three main variables, the tissue PSA per unit volume of prostatic tissue $P(\boldsymbol{x}, t)$, the serum PSA $P_s(t)$, and the concentration of tumor cells $c_t(\boldsymbol{x}, t)$. First, the initial concentration of tumor cells $c_t(\boldsymbol{x}, t_0)$ is calculated from the diagnostic image-data derived cellularity $c(\boldsymbol{x}, t_0)$, the mean percentage of tumor cells estimated in the patient's biopsy p_b, and the tumor mask $T_{mask}(\boldsymbol{x}, t_0)$:

$$c_t(\boldsymbol{x}, t_0) = p_b T_{mask}(\boldsymbol{x}, t_0) c(\boldsymbol{x}, t_0). \tag{2}$$

Then, we introduce a mathematical model to simulate the evolution of the concentration of tumor cells:

$$\frac{\partial c_t(\boldsymbol{x}, t)}{\partial t} = \phi_\theta(\boldsymbol{x}, t) \alpha_t c_t(\boldsymbol{x}, t), \tag{3}$$

where $\phi_\theta(\boldsymbol{x}, t)$ is a function approximated by a NN that determines temporarily and spatially the fraction of tumor cells that are proliferating and α_t the tumor growth rate. We consider that when the concentration of cells reaches a maximum tissue-carrying capacity equal to 0.8[39,40], the tumor mask is expanded to the neighboring voxels (see *Carrying capacity sensitivity analysis* in Supplementary Material). Then, we compute the concentration

of tumor cells in the expanded regions. The temporal growth of the tumor volume $T_{vol}(t)$ can be obtained by integrating tumor mask across the prostate domain (Ω_x):

$$T_{vol}(t) = \int_{\Omega_x} T_{mask}(\boldsymbol{x}, t)d\boldsymbol{x}. \quad (4)$$

The tissue PSA increase is a consequence of the PSA leakage of cancer cells. Thus, the tissue PSA per unit volume of prostatic tissue ($P(\boldsymbol{x}, t)$) is calculated from:

$$\frac{\partial P(\boldsymbol{x}, t)}{\partial t} = \alpha_p c_t(\boldsymbol{x}, t) - \gamma_p P(\boldsymbol{x}, t) - \frac{m_{w_{gad}}}{m_{w_{PSA}}} k_{trans}(\boldsymbol{x})\left(P(\boldsymbol{x}, t) - \frac{P_s(t)}{\Omega_{vox}}\right). \quad (5)$$

Here, α_p is the tissue PSA production rate per unit volume of prostatic tissue by tumor cells, γ_p denotes the natural tissue PSA decay rate, $m_{w_{PSA}}$ represents the molecular weight of PSA[47] ($m_{w_{PSA}} = 26$ Da), $m_{w_{gad}}$ denotes the molecular weight of gadolinium-based contrast agent[48], and Ω_{vox} the volume of the interchanging region. The gadolinium-based contrast agent used for the MRI scans was Gadoterate Meglumine[49], which has a molecular weight of 753.9 Da. The last term represents the PSA exchange between the tissue and blood in vascularized regions of the prostate, as depicted by the spatial distribution of $k_{trans}(\boldsymbol{x})$. We assume constant and sufficient blood flow, allowing for a steady concentration of serum PSA ($P_s(t)$) through the capillaries. The serum PSA's evolution ($P_s(t)$) is determined by integrating the PSA exchange between the tissue and blood:

$$\frac{dP_s(t)}{dt} = \int_{\Omega_x} \frac{m_{w_{gad}}}{m_{w_{PSA}}} k_{trans}(\boldsymbol{x})\left(P(\boldsymbol{x}, t) - \frac{P_s(t)}{\Omega_{vox}}\right)d\boldsymbol{x} - \gamma_s P_s(t) \quad (6)$$

with γ_s the natural serum PSA decay rate.

Finally, the temporal scale of the PSA exchange between blood and tissue through $k_{trans}(\boldsymbol{x})$ occurs within seconds, whereas the dynamics of tumor growth typically occur over months. To simulate time frames spanning years and enable follow-ups at any time point, we elaborate a methodology to address the different timescales in this physics-based model (refer to *PSA dynamics simulation* in Supplementary Material).

Model calibration

The NN is trained with 20% dropout employing a training and validation dataset. We employed the Python package Optuna[50] for hyperparameter tuning. The loss function used was mean squared error (squared L2 norm). At each epoch, we used a stochastic gradient optimizer (Adam)[51] for the training dataset and then tested over the validation dataset. The training dataset included three different serum PSA values, $\boldsymbol{P}_{s_{train}} = [1.5, 3, 5]$ ng/mL at 60 days, representing a slow, medium and fast serum PSA increase from an initial serum PSA of 1 ng/mL. The validation dataset comprised three different serum PSA values, $\boldsymbol{P}_{s_{val}} = [1, 2, 4]$ ng/mL at day 60, starting from an initial serum PSA of 1 ng/mL. We tuned the hyperparameters until minimizing the loss over the validation dataset. The optimized NN hyperparameters resulted in 2 hidden layers of 353 nodes each one, trained with a learning rate of 0.001 and 4585 epochs. The patient-specific parameters of the physics-based model are also optimized using the Optuna framework. In this case, we minimize the error based on the mean squared error (squared L2 norm) between the predicted and the clinical tumor volume and the serum PSA.

Model implementation

The digital twin's geometry reconstruction from T2-weighted MRI sequences was done using *im2mesh* library[43] in Python and the voxelization was performed in Matlab R2023b. The entire computational model was implemented in Python, utilizing the library PyTorch for the machine learning model. The physics-based model was integrated in time using the forward Euler method, with a time step of $\Delta t = 1$ day. The simulations were performed using an Intel(R) Core(TM) i9-7900X CPU @ 3.30GHz, 32.0 GB RAM, and NVIDIA GeForce GTX 1050 Ti GPU. For Patient A, a full simulation spanning 786 real-world days, with a geometry discretization of 1 mm^3 and 66,691 voxels, takes ~25 s on the GPU and 53 s on the CPU. For Patient B, a simulation covering 956 real-world days, with a geometry discretization of 1 mm^3 and 52,709 voxels, takes ~25 s on the GPU and 52 s on the CPU.

Data availability

The data used for the reproducibility of the model are openly available on GitHub https://github.com/daniel-camacho-gomez/ProstateNet. Additional data can be provided upon request.

Code availability

The code is openly available on GitHub at https://github.com/daniel-camacho-gomez/ProstateNet.

Received: 6 May 2024; Accepted: 8 July 2025;
Published online: 29 July 2025

References

1. Rawla, P. Epidemiology of prostate cancer. *World J. Oncol.* **10**, 63 (2019).
2. Wang, L. et al. Prostate cancer incidence and mortality: global status and temporal trends in 89 countries from 2000 to 2019. *Front. Public Health* **10**, 176 (2022).
3. Bubendorf, L. et al. Metastatic patterns of prostate cancer: an autopsy study of 1589 patients. *Hum. Pathol.* **31**, 578–583 (2000).
4. Logothetis, C. J. & Lin, S.-H. Osteoblasts in prostate cancer metastasis to bone. *Nat. Rev. Cancer* **5**, 21–28 (2005).
5. Saitoh, H. et al. Metastatic patterns of prostatic cancer: correlation between sites and number of organs involved. *Cancer* **54**, 3078–3084 (1984).
6. Datta, K., Muders, M., Zhang, H. & Tindall, D. J. Mechanism of lymph node metastasis in prostate cancer. *Future Oncol.* **6**, 823–836 (2010).
7. Woodhouse, E. C., Chuaqui, R. F. & Liotta, L. A. General mechanisms of metastasis. *Cancer Interdiscip. Int. J. Am. Cancer. Soc.* **80**, 1529–1537 (1997).
8. Gupta, R. T., Mehta, K. A., Turkbey, B. & Verma, S. Pi-rads: past, present, and future. *J. Magn. Reson. Imaging* **52**, 33–53 (2020).
9. Egevad, L., Granfors, T., Karlberg, L., Bergh, A. & Stattin, P. Prognostic value of the gleason score in prostate cancer. *BJU Int.* **89**, 538–542 (2002).
10. Draisma, G., Postma, R., Schröder, F. H., van der Kwast, T. H. & de Koning, H. J. Gleason score, age and screening: modeling dedifferentiation in prostate cancer. *Int. J. cancer* **119**, 2366–2371 (2006).
11. Lilja, H., Ulmert, D. & Vickers, A. J. Prostate-specific antigen and prostate cancer: prediction, detection and monitoring. *Nat. Rev. Cancer* **8**, 268–278 (2008).
12. Stenman, U.-H., Leinonen, J., Zhang, W.-M. & Finne, P. Prostate-specific antigen. *Semin. Cancer Biol.* **9**, 83–93 (1999).
13. Gretzer, M. B. & Partin, A. W. Psa levels and the probability of prostate cancer on biopsy. *Eur. Urol. Suppl.* **1**, 21–27 (2002).
14. Thompson, I. M. & Ankerst, D. P. Prostate-specific antigen in the early detection of prostate cancer. *Cmaj* **176**, 1853–1858 (2007).
15. Holmström, B. et al. Prostate specific antigen for early detection of prostate cancer: longitudinal study. *Bmj* **339**, b3537 (2009).
16. Adhyam, M. & Gupta, A. K. A review on the clinical utility of psa in cancer prostate. *Indian J. Surg. Oncol.* **3**, 120–129 (2012).
17. Ilic, D., Neuberger, M. M., Djulbegovic, M. & Dahm, P. Screening for prostate cancer. *Cochrane Database of Syst. Rev* **1**, 1–78 (2013).
18. Brady-Nicholls, R. et al. Prostate-specific antigen dynamics predict individual responses to intermittent androgen deprivation. *Nat. Commun.* **11**, 1750 (2020).

19. Brady-Nicholls, R. et al. Predicting patient-specific response to adaptive therapy in metastatic castration-resistant prostate cancer using prostate-specific antigen dynamics. *Neoplasia* **23**, 851–858 (2021).
20. Jain, H. V., Clinton, S. K., Bhinder, A. & Friedman, A. Mathematical modeling of prostate cancer progression in response to androgen ablation therapy. *Proc. Natl. Acad. Sci. USA* **108**, 19701–19706 (2011).
21. Colli, P. et al. Mathematical analysis and simulation study of a phase-field model of prostate cancer growth with chemotherapy and antiangiogenic therapy effects. *Math. Models Methods Appl. Sci.* **30**, 1253–1295 (2020).
22. Lorenzo, G. et al. Mechanistic modelling of prostate-specific antigen dynamics shows potential for personalized prediction of radiation therapy outcome. *J. R. Soc. Interface* **16**, 20190195 (2019).
23. Kamel Boulos, M. N. & Zhang, P. Digital twins: from personalised medicine to precision public health. *J. Personal. Med.* **11**, 745 (2021).
24. Laubenbacher, R., Mehrad, B., Shmulevich, I. & Trayanova, N. Digital twins in medicine. *Nat. Comput. Sci.* **4**, 184–191 (2024).
25. Katsoulakis, E. et al. Digital twins for health: a scoping review. *npj Digital Med.* **7**, 77 (2024).
26. Willcox, K. & Segundo, B. The role of computational science in digital twins. *Nat. Comput. Sci.* **4**, 147–149 (2024).
27. Gimpel, A. L., Stark, W. J., Heckel, R. & Grass, R. N. A digital twin for dna data storage based on comprehensive quantification of errors and biases. *Nat. Commun.* **14**, 6026 (2023).
28. Maksymenko, K., Clarke, A. K., Mendez Guerra, I., Deslauriers-Gauthier, S. & Farina, D. A myoelectric digital twin for fast and realistic modelling in deep learning. *Nat. Commun.* **14**, 1600 (2023).
29. Hervas-Raluy, S., Sainz-DeMena, D., Gómez-Benito, M. J. & García-Aznar, J. M. Image-based biomarkers for engineering neuroblastoma patient-specific computational models. *Eng. Comput.* **40**, 3215-3231 (2024).
30. Boubaker, M. B., Haboussi, M., Ganghoffer, J.-F. & Aletti, P. Finite element simulation of interactions between pelvic organs: predictive model of the prostate motion in the context of radiotherapy. *J. Biomech.* **42**, 1862–1868 (2009).
31. Lorenzo, G. et al. Tissue-scale, personalized modeling and simulation of prostate cancer growth. *Proc. Natl. Acad. Sci. USA* **113**, E7663–E7671 (2016).
32. Lorenzo, G., Hughes, T. J., Dominguez-Frojan, P., Reali, A. & Gomez, H. Computer simulations suggest that prostate enlargement due to benign prostatic hyperplasia mechanically impedes prostate cancer growth. *Proc. Natl. Acad. Sci. USA* **116**, 1152–1161 (2019).
33. Hadjicharalambous, M. et al. Personalised in silico biomechanical modelling towards the optimisation of high dose-rate brachytherapy planning and treatment against prostate cancer. *Front. Physiol.* **15**, 1491144 (2024).
34. Lorenzo, G. et al. A pilot study on patient-specific computational forecasting of prostate cancer growth during active surveillance using an imaging-informed biomechanistic model. *Cancer Res. Commun.* **4**, 617–633 (2024).
35. Tofts, P. S. Modeling tracer kinetics in dynamic gd-dtpa mr imaging. *J. Magn. Reson. imaging* **7**, 91–101 (1997).
36. Pishko, G. L., Astary, G. W., Mareci, T. H. & Sarntinoranont, M. Sensitivity analysis of an image-based solid tumor computational model with heterogeneous vasculature and porosity. *Ann. Biomed. Eng.* **39**, 2360–2373 (2011).
37. LoCastro, E. et al. Computational modeling of interstitial fluid pressure and velocity in head and neck cancer based on dynamic contrast-enhanced magnetic resonance imaging: feasibility analysis. *Tomography* **6**, 129–138 (2020).
38. Borau, C. et al. A multiscale orchestrated computational framework to reveal emergent phenomena in neuroblastoma. *Comput. Methods Prog. Biomed.* **241**, 107742 (2023).
39. Zelhof, B. et al. Correlation of diffusion-weighted magnetic resonance data with cellularity in prostate cancer. *BJU Int.* **103**, 883–888 (2009).
40. Chatterjee, A. et al. Changes in epithelium, stroma, and lumen space correlate more strongly with gleason pattern and are stronger predictors of prostate adc changes than cellularity metrics. *Radiology* **277**, 751–762 (2015).
41. Isensee, F., Jaeger, P. F., Kohl, S. A., Petersen, J. & Maier-Hein, K. H. nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat. methods* **18**, 203–211 (2021).
42. Jimenez-Pastor, A. et al. Automated prostate multi-regional segmentation in magnetic resonance using fully convolutional neural networks. *Eur. Radiol.* **33**, 5087–5096 (2023).
43. Sainz-DeMena, D., García-Aznar, J. M., Pérez, M. Á. & Borau, C. Im2mesh: a python library to reconstruct 3d meshes from scattered data and 2d segmentations, application to patient-specific neuroblastoma tumour image sequences. *Appl. Sci.* **12**, 11557 (2022).
44. Anderson, A. et al. Effects of cell volume fraction changes on apparent diffusion in human cells. *Magn. Reson. Imaging* **18**, 689–695 (2000).
45. Atuegwu, N. et al. Incorporation of diffusion-weighted magnetic resonance imaging data into a simple mathematical model of tumor growth. *Phys. Med. Biol.* **57**, 225 (2011).
46. Sainz-DeMena, D., Ye, W., Pérez, M. Á. & García-Aznar, J. M. A finite element based optimization algorithm to include diffusion into the analysis of dce-mri. *Eng. Comput.* **38**, 3849–3865 (2022).
47. Bélanger, A. et al. Molecular mass and carbohydrate structure of prostate specific antigen: studies for establishment of an international psa standard. *Prostate* **27**, 187–197 (1995).
48. Böck, J. C., Kaufmann, F. & Felix, R. Comparison of gadolinium-dtpa and macromolecular gadolinium-dtpa-polylysine for contrast-enhanced pulmonary time-of-flight magnetic resonance angiography. *Investig. Radiol.* **31**, 652–657 (1996).
49. National Center for Biotechnology Information. Pubchem compound summary for cid 6918037, gadoterate meglumine. https://pubchem.ncbi.nlm.nih.gov/compound/Gadoterate-Meglumine (2024). Retrieved March 8, 2024.
50. Akiba, T., Sano, S., Yanase, T., Ohta, T. & Koyama, M. Optuna: a next-generation hyperparameter optimization framework. In *The 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2623–2631 (ACM stands for Association for Computing Machinery (ACM, 2019).
51. Kingma, D. P. & Ba, J. Adam: a method for stochastic optimization. *arXiv preprint* https://doi.org/10.48550/arXiv.1412.6980 (2014).

Acknowledgements

The authors acknowledge the support of the Aragon Government through Group T50_23R. We would like to express our gratitude to our collaborators from the Biomedical Imaging Research Group (GIBI230) at HULAFE, and Quibim S.L. for their essential contribution in retrieving and processing the patient-specific clinical data required for the development of this work. We would also like to thank Justin Bunker (University of Cambridge) for useful discussions. This research was funded by Next Generation EU (ProCanAid Grant No. PLEC2021-007709). D.C.G. was also supported by the Erasmus+ program managed by Campus Iberus and Fundación Ibercaja-Cai (No IT 5/23) and J.M.G.A. was also supported by the European Research Council (ICoMICS Adv grant agreement: 101018587).

Author contributions

D.C.G. designed the study, implemented the framework, performed all coding, performed the simulations and analysis, prepared the figures, and wrote the original manuscript. C.B. supervised the computational work and reviewed the manuscript. J.M.G.A. supervised the computational work and reviewed the manuscript. M.J.G.B. supervised the computational work and reviewed the manuscript. M.G. supervised the computational work and

reviewed the manuscript. M.A.P. supervised the computational work, reviewed the manuscript, and obtained the resources. All authors read and approved the final manuscript.

Competing interests
The authors declare no competing interests

Ethical approval
The studies involving humans were approved by the Técnica del Comité de Ética de la Investigación con Medicamentos (CEIM) at Hospital Universitario y Politécnico La Fe. The studies were conducted in compliance with local legislation and institutional requirements. Written informed consent for participation was not required from the participants or their legal guardians/next of kin, in accordance with national legislation and institutional guidelines.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41746-025-01890-x.

Correspondence and requests for materials should be addressed to Maria Angeles Perez.