



MOX-Report No. 65/2025

**MOSAİK: A computational framework for theranostic digital twin in  
renal cell carcinoma**

Pottier, A.; Gelardi, F.; Larcher, A.; Capitanio, U.; Rainone, P.; Moresco,  
R.M.; Tenace, N.; Colecchia, M.; Grassi, S.; Ponzoni, M.; Chiti, A.; Cavinato,  
L.

MOX, Dipartimento di Matematica  
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

[mox-dmat@polimi.it](mailto:mox-dmat@polimi.it)

<https://mox.polimi.it>

# MOSAİK: A computational framework for theranostic digital twin in renal cell carcinoma

Augustin Pottier <sup>1</sup>, Fabrizia Gelardi <sup>1</sup>, Alessandro Larcher <sup>2</sup>, Umberto Capitano <sup>2</sup>, Paolo Rainone <sup>3</sup>, Rosa Maria Moresco <sup>4</sup>, Nazario Tenace <sup>5</sup>, Maurizio Colecchia <sup>5</sup>, Stefano Grassi <sup>5</sup>, Maurilio Ponzoni <sup>5</sup>, Arturo Chiti <sup>1</sup>, Lara Cavinato <sup>6,\*</sup>

<sup>1</sup>Department of Nuclear Medicine, Ospedale San Raffaele, Milan, Italy

<sup>2</sup>Department of Urology, Ospedale San Raffaele, Milan, Italy

<sup>3</sup>Institute of Molecular Bioimaging and Physiology, National Research Council, Milan, Italy

<sup>4</sup>Department of Medicine and Surgery, University of Milan Bicocca, Milan, Italy

<sup>5</sup>Department of Pathological Anatomy, Ospedale San Raffaele, Milan, Italy

<sup>6</sup>MOX Laboratory, Department of Mathematics, Politecnico di Milano, Milan, Italy

\*Correspondence: [lara.cavinato@polimi.it](mailto:lara.cavinato@polimi.it)

## ABSTRACT

1 In nuclear oncology, radiopharmaceuticals (RP) emerged as theranostic tools able to bind specifically to cancer biomark-  
2 ers and inflict subsequently systematic and irreparable damage to the DNA of the targeted cells. That is why radioisotopes-  
3 based therapies entered the clinical practice to diagnose and treat tumours simultaneously and may potentially overcome  
4 therapeutic resistance encountered in various cancers. Despite these advancements, tumoural heterogeneity and poor anti-  
5 cancer drug penetration in solid tumours turns out to be overlooked pieces of the personalized oncology puzzle, leading to  
6 treatment failure. In this study, we propose MOSAİK, an oncological digital twin framework to simulate the intra-tumour up-  
7 take of radiopharmaceutical agent, specifically [<sup>89</sup>Zr]Zr-girentuximab in clear cell Renal Cell Carcinoma (ccRCC). Our com-  
8 prehensive approach integrates patient-based insights in space and time for reflecting the multi-faceted nature of RP up-  
9 take. We develop models to segment blood vessels and identify neoplastic regions, enabling the characterization of the  
10 biological domain. To discuss the intra-tumour heterogeneity contribution to the drug diffusion process, we spatially cor-  
11 relate immunochemistry images-derived parameters with the baseline drug accumulation captured through micro PET  
12 imaging. Additionally, the model is informed with temporal features leveraged from the compartmental model of the RP  
13 agent. The presented Deep Learning (DL) framework incorporates interpretable spatial and temporal inputs stemming from  
14 histopathology images. This work aims to provide a computational model with predictive capabilities in drug retention in  
15 tissues to move beyond the one-size-fits-all paradigm in nuclear medicine.

## 16 INTRODUCTION

17 Radiopharmaceuticals (RP) represent a cornerstone in the evolving landscape of precision medicine, where patient-centred ap-  
18 proaches align therapeutic decisions with individual genetic and biological profiles. By combining radioisotopes and pharmaceu-  
19 ticals, RP therapies bridge diagnostic imaging and targeted treatments, allowing the accurate delivery of radioactive payloads to  
20 cancer cells<sup>1</sup>. This dual diagnostictherapeutic function provides clinicians with valuable molecular insights into tumour behaviour  
21 and raises new expectations for managing advanced cancers often refractory to conventional therapies<sup>2</sup>. However, the clinical suc-  
22 cess of such approaches depends on a deep understanding of cancer biology and on modelling the complex cellular and micro-  
23 environmental processes, such as aberrant neovascularizationthat constrain drug diffusion and limit therapeutic efficacy in solid  
24 tumours. Within this framework, Digital Twins (DTs) emerge as a promising paradigm for representing tumour evolution and thera-  
25 peutic response. DTs, conceived as dynamic computational replicas of biological systems continuously updated with patient-specific  
26 data, enable predictive and knowledge-driven strategies to anticipate treatment resistance and failure<sup>3</sup>.

27 Despite remarkable progress, current literature still faces major challenges in integrating tumour heterogeneity into predictive mod-  
28 els of radiopharmaceutical diffusion and retention. Tumour heterogeneity, arising both from genetic variability among tumour cells  
29 (inter- and intra-tumoural) and from spatial differences in the tumour micro-environment (TME), including vasculature, permeabil-  
30 ity, hypoxia, and necrosis remains a principal determinant of therapy failure<sup>4–7</sup>. Numerous studies have underlined the prognostic  
31 and diagnostic relevance of TME heterogeneity metrics across cancer types. For instance, in renal cell carcinoma (RCC), Carbonic  
32 Anhydrase IX (CAIX) expression not only defines the clear cell RCC subtype but also correlates with improved survival outcomes<sup>8</sup>.  
33 Yet, standard compartmental models, widely used to analyse dynamic Positron Emission Tomography (PET) data following RP ad-  
34 ministration, assume homogeneous tumour regions and disregard spatial variations in tracer kinetics. Consequently, such models

overlook how intra-tumoural heterogeneity shapes the spatio-temporal evolution of radiopharmaceutical uptake, hindering their ability to yield clinically actionable predictions about treatment efficacy.

Several recent works have sought to overcome these limitations through novel computational and imaging methodologies. Approaches combining mathematical modelling with high-resolution 3D imaging, for example, have reconstructed tumour vascular networks and simulated intratumoural drug diffusion using optical projection tomography and MRI-based diffusion analyses<sup>9,10</sup>. These studies demonstrated that tumour tissues exhibit non-uniform and temporally varying diffusion behaviours, confirming the need for spatially resolved models that account for the heterogeneous tumour architecture. Nevertheless, existing models largely remain constrained by limited interpretability, reliance on coarse morphologic indicators (e.g., tumour size or bulk morphology), or dependence on partial histological sections, often from animal models, insufficient to represent global tumoural responses<sup>11,12</sup>. Moreover, most in-silico frameworks addressing radiopharmaceutical pharmacokinetics have been developed at the whole-body scale, neglecting the complex intra-tumoural interactions between vasculature, neoplastic cells, and supporting stroma<sup>13,14</sup>.

To address these gaps, we introduce MOSAIK, an in-silico framework designed to simulate intra-tumoural uptake of [<sup>89</sup>Zr]Zr-girentuximab in clear cell renal cell carcinoma (ccRCC). Our approach integrates deep learning-based spatial characterization of tumour histology with immunokinetic modelling of radiopharmaceutical dynamics, enabling spatially resolved predictions of tracer accumulation. Specifically, baseline drug uptake is inferred from ex-vivo micro-PET imaging of tumour slices, while segmentation models automatically delineate blood vessels and tumour regions on whole-slide histology images. By mapping histology-derived parameters that describe the TME composition and structure, MOSAIK quantitatively links microscopic tissue architecture to macroscopic radiopharmaceutical distribution. This multimodal framework thus contributes to the broader goal of implementing oncological Digital Twins capable of forecasting therapeutic outcomes and supporting personalized treatment planning.

The contributions of this work are threefold. First, we provide a hybrid deep learning and mechanistic modelling framework that captures intra-tumoural heterogeneity in RP diffusion and retention. Second, we establish a pipeline linking histology-based tissue characterization with immunokinetic modelling, allowing interpretable prediction of radiopharmaceutical uptake patterns. Third, we demonstrate the feasibility of the proposed approach in ccRCC, quantitatively relating spatial tracer distribution to CAIX expression as a biological marker of tumour aggressiveness.

The remainder of this paper is organized as follows. Section 2 details the methodological framework and dataset, including the modelling of tracer kinetics and histological analyses. Section 3 presents the experimental setup and evaluation metrics. Section 4 reports and discusses the results, highlighting the predictive capabilities and interpretability of MOSAIK. Finally, Section 5 concludes with perspectives on the integration of such frameworks into future oncological Digital Twin systems for precision radiopharmaceutical therapy.

## **SIMULATION PIPELINE**

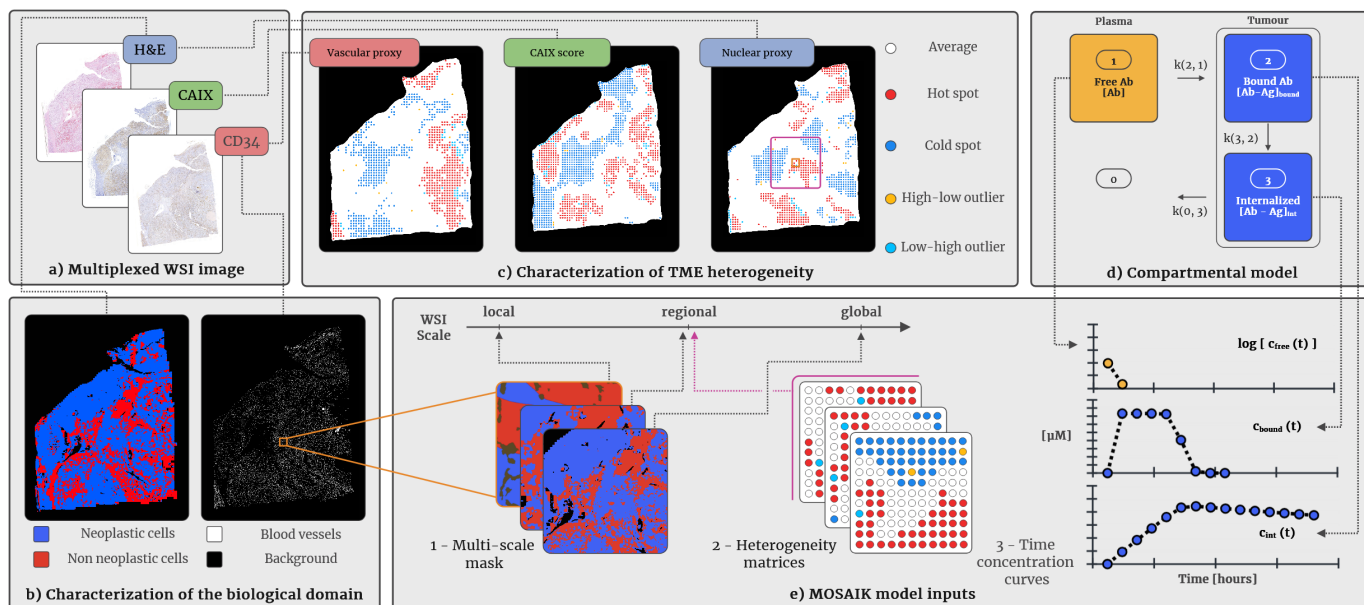
### **Overview of the workflow**

Building upon the challenges identified in digital pathology analysis, we propose a precision medicine framework to simulate tumoural response to RP therapy. The framework constructs multimodal inputs for a digital twin network by combining: (i) characterisation of the biological domain from histology, (ii) mapping of intra-tumour heterogeneity, and (iii) resolution of an immunokinetic (compartmental) model for [<sup>89</sup>Zr]Zr-girentuximab. The workflow leverages IHC-derived proxies of intra-tumour heterogeneity to translate microscopic-scale information into macroscopic-scale predictors of radiopharmaceutical uptake and retention. Implementation was performed in Python (v3.12.7) using OpenSlide for WSI handling and pydicom for PET data.

A high-level overview of the processing pipeline is provided in Figure 1: two spatial inputs (multi-scale masks and heterogeneity matrices) are obtained from histological analysis, while the temporal input (time concentration curves) is derived from the immunokinetic compartmental model described in Section 2.6.

### **Clinical protocol and data acquisition**

The clinical protocol is summarised in Figure 2. Two patients with different grades of clear-cell renal cell carcinoma (ccRCC) received a single intravenous injection of [<sup>89</sup>Zr]Zr-DFO-girentuximab (37 MBq ± 10%). Whole-body PET/CT acquisitions were obtained at



**Figure 1:** Overview of the framework for constructing multimodal inputs to the digital twin model. Two spatially resolved inputs related to radiopharmaceutical diffusion are obtained from histological analysis, while the temporal input is derived from the  $^{89}\text{Zr}$ -girentuximab immunokinetic model.

days 3, 5 and 7 post-injection. Surgical resection was scheduled at day 14 post-administration to reflect radionuclide behaviour in vivo. Primary tumours were excised and bisected along the imaging plane; ex vivo biosamples were imaged using a high-resolution preclinical microPET scanner to obtain PET reconstructions of the tumour halves (see Step 3 in Figure 2). From the bisected specimen, 4  $\mu\text{m}$  sections were prepared from formalin-fixed paraffin-embedded tissue, stained with H&E, and processed for CD34 and CAIX immunostaining to visualise vasculature and biomarker expression, respectively. All sections were digitised with WSI scanners, producing the histology images used throughout the pipeline.

## Pre-processing of Whole-Slide Images

Whole-Slide Images (WSIs) were processed in a tile-based fashion to generate spatially consistent inputs for downstream analysis and modelling. The following sub-steps describe the extraction and harmonisation of image tiles and the generation of labels from PET. The full procedure is illustrated in Figure 3 and yields a consistent extraction mask for tile sampling across multiple stains.

### Slide mask extraction

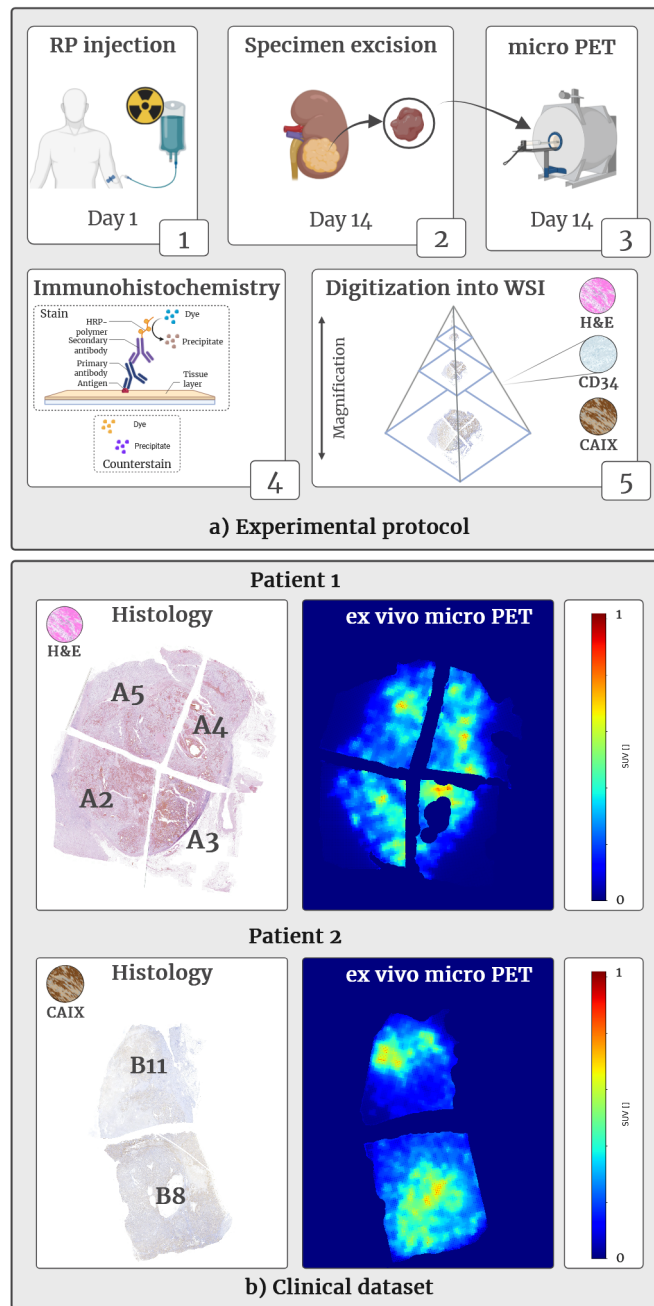
The tissue region on each WSI was delineated by generating a downsampled thumbnail (downsample factor = 50) and converting it to grayscale. Contrast inversion was applied when required, followed by Otsu thresholding to obtain a primary binary mask. Morphological operations (dilation, hole filling, removal of small objects) refined this mask, and a kernel-based filling operation completed the delineation of the region of interest.

### Tile dataset extraction

Patches were sampled from the highest-quality WSI level. Overlapping tiles were arranged on a regular grid covering the extraction mask, with 25% overlap horizontally and vertically. The initial grid of "inner tiles" excluded tiles covering less than 95% of their area by the tissue mask. "Border tiles" were appended along the mask boundary to ensure coverage, subject to a minimum 50% tissue overlap. Because multiple stains correspond to adjacent sections, the CD34-derived mask was chosen as the reference; masks from other stains were aligned and rotated by centroid matching to this reference to ensure consistent tile coordinates across stains.

### H&E stain normalization

To mitigate colour variability and staining artifacts that impair automated analysis, an adapted Macenko stain-normalisation method<sup>15</sup> was applied to H&E tiles. The algorithm operates in the optical density (OD) space, applies colour deconvolution to estimate reference staining matrices and stain concentrations, and returns harmonised tiles suitable for segmentation and feature extraction.



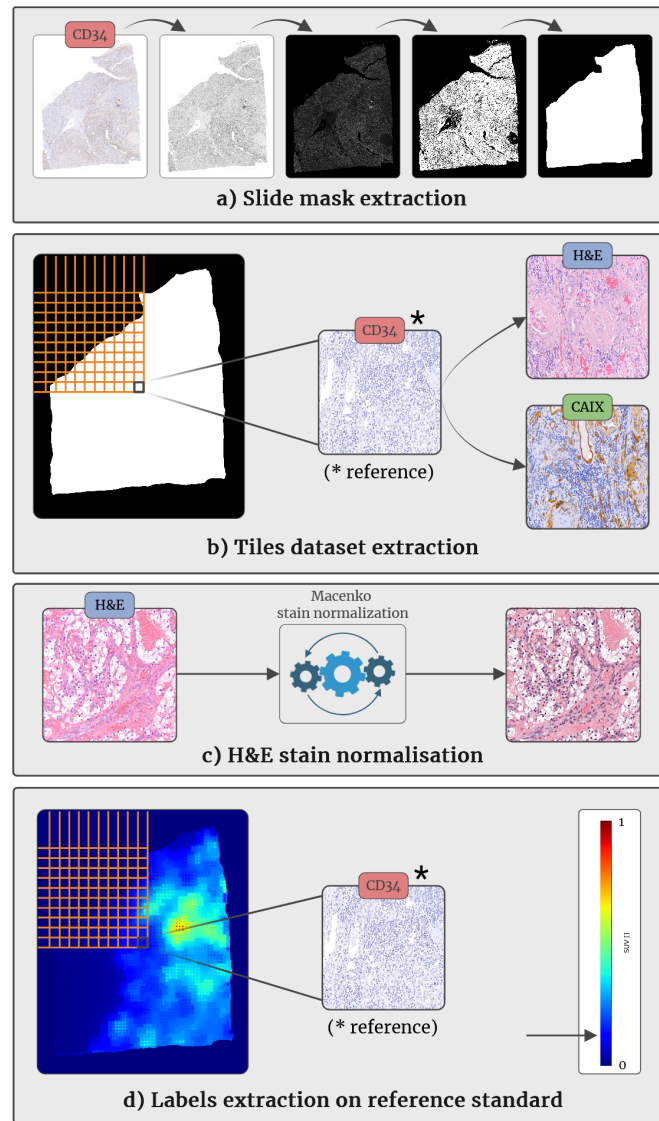
**Figure 2:** Overview of the clinical protocol applied to each patient. Step 3 yields an ex vivo microPET image of the excised biosample, while Step 5 results in H&E staining and two immunostaining procedures for specific tissue property annotation on Whole-Slide Images (WSIs). The clinical dataset is illustrated for patient 1.

### 103 Label extraction and PET normalization

104 PET images were converted from activity concentration (Bq/mL) to Standardised Uptake Value (SUV) to enable inter-subject com-  
 105 parison. SUV normalisation used body surface area (BSA) computed with the Du Bois formula<sup>16</sup>, incorporating tissue activity, ad-  
 106 ministered dose and patient anthropometrics. To reconcile the large resolution mismatch between histology tiles (0.25  $\mu\text{m}/\text{px}$ ) and  
 107 microPET (0.4 mm/px), tiles were assigned labels by mapping tile coordinates to PET voxels: each high-resolution tile approximately  
 108 corresponds to a single PET voxel, using reference coordinates to compute physical distances from a common origin.

### 109 Characterization of the biological domain

110 A biologically informed representation of the tissue domain was constructed by segmenting vasculature, delineating neoplastic vs  
 111 non-neoplastic compartments, and segmenting cell nuclei. These outputs form the multi-scale semantic masks and microscopic



**Figure 3:** Procedural stages for extracting and preprocessing the tiles dataset from a Whole-Slide Image. Slide mask extraction ensures consistent tile sampling across multiple stains representing distinct biological elements. Stain normalisation on standard H&E tiles yields a harmonised dataset suitable for subsequent automatic segmentation tasks.

parameter sets used downstream. [Figure 4](#) overviews the characterization pipeline.

### Blood vessel segmentation

CD34 immunostaining highlights endothelial cells and was used to detect vasculature supplying the tumour. On extracted tiles, CD34-positive endothelial regions appear brown against hematoxylin counterstain. Tiles were converted to HSV colour space and a brown-hue range was thresholded to identify candidate vascular objects. Morphological closing and size filtering removed spurious elements and retained relevant vessels.

### Tissue segmentation

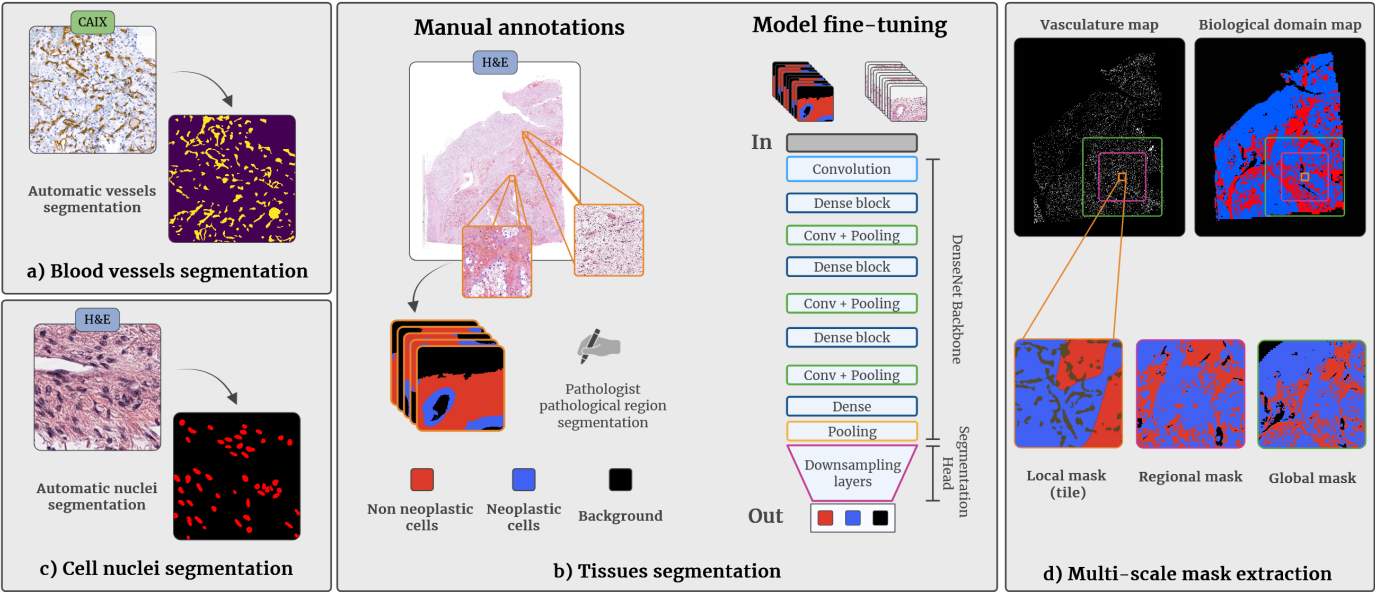
To distinguish neoplastic tissue from surrounding stroma, we trained a pixel-wise segmentation model on H&E tiles manually annotated by a pathologist into three classes: neoplastic, non-neoplastic and background. A DenseNet169 backbone was fine-tuned with a segmentation head composed of four terminal downsampling layers. Training parameters are reported in [Table 1](#). The trained pixel classifier was applied tile-wise across the dataset and the resulting semantic masks were stitched to produce slide-level biological domain maps.

**Table 1:** Tissues segmentation training parameters

|                    |                    |
|--------------------|--------------------|
| Train / Validation | 70 / 30            |
| Batch size         | 8                  |
| Optimizer          | Adam               |
| Loss function      | Cross-entropy      |
| Class weights      | [0.45, 0.24, 0.31] |
| Number of epochs   | 25                 |

124 **Cell nuclei segmentation**

125 Nuclear segmentation was performed on H&E tiles to derive cellular morphometry features. We used StarDist with the pretrained  
126 “2D\_versatile\_he” model (trained on TCGA and TNBC datasets) to obtain instance-level nuclear masks for downstream feature ex-  
127 traction.



**Figure 4:** Segmentation workflow for extracting vascular, tissue and nuclear compartments from WSIs. Tissue segmentation pipeline uses a DenseNet-based pixel classifier; blood-vessel segmentation is performed tile-wise and aggregated to slide level; nuclei segmentation is performed tile-wise for morphometric analysis.

128 **Heterogeneity analysis**

129 To capture the multifaceted intra-tumour heterogeneity relevant to RPT uptake, we computed and aggregated three classes of pa-  
130 rameters: vasculature-derived features, cell nucleiderived features, and a tile-wise CAIX score. The heterogeneity analysis consisted  
131 of three steps: (i) parameter mapping, (ii) parameter selection via spatial correlation with PET uptake, and (iii) spatial cluster identifi-  
132 cation using local Morans Index.

133 **Parametric maps extraction**

134 For each vascular or nuclear metric, local values were computed per tile and averaged to produce tile-wise parameter values. CAIX  
135 was scored per tile by combining staining intensity and the fraction of CAIX-positive cells. Tile-wise values were interpolated into  
136 continuous heatmaps over the extraction mask using Inverse Distance Weighted (IDW) interpolation. IDW estimates values at un-  
137 known locations as a distance-weighted average of neighbouring tile centres, producing continuous parameter maps aligned to  
138 the tile grid for subsequent spatial correlation.

139 **Parametric maps selection**

140 To select robust macroscopic proxies for RPT uptake, we measured the spatial association between each parameter map and the  
141 baseline microPET slice co-registered to the histology section. PET labels were subjected to IDW interpolation and minmax nor-  
142 malisation to [0, 1]. We computed Pearson (linear) and Spearman (monotonic) correlation coefficients between parameter maps



and PET-derived uptake. For vascular- and nuclear-derived metric groups, we selected the parameter that maximised the absolute value of both correlation metrics across slides. CAIX was evaluated as a standalone parameter given its direct mechanistic role in antibody targeting. This selection strategy captures both linear and non-linear spatial relationships between tissue features and tracer distribution<sup>17,18</sup>.

#### Spatial cluster identification

For deeper characterisation, tile-wise parameter sets were subjected to local spatial autocorrelation analysis via local Morans Index. Tile values were normalised to z-scores and an 8-neighbour connectivity (Moore neighbourhood) defined spatial weights. Local Morans Index identifies clustered hot spots and cold spots as well as high-low and low-high outliers. Statistical significance was assessed under the null hypothesis at  $p = 0.005$ . Each tile was labelled as one of: hot spot, cold spot, high-low outlier or low-high outlier. The three Morans I outcomes (one per selected proxy) were aggregated into a heterogeneity representation for each tile and used as a 3-channel input to the learning model.

#### Compartmental model for [<sup>89</sup>Zr]Zr-girentuximab

Girentuximab is a chimeric monoclonal antibody targeting Carbonic Anhydrase IX (CAIX), expressed in the majority of ccRCC cases<sup>19</sup>. The [<sup>89</sup>Zr]-girentuximab kinetics is represented by a non-linear three-compartment model focused on plasma and cellular dynamics<sup>20</sup>. Upon injection, the antibody circulates in plasma (compartment 1), binds CAIX on cell surfaces forming a bound compartment (compartment 2) and is subsequently internalised into an intracellular compartment (compartment 3). The model assumes negligible dissociation of bound antibody from surface antigen during the time scales considered. Plasma time-activity curves act as the input function to the tumour and the tumour binding rate  $k(2, 1)$  is time-varying and non-linear, capturing antigen-limited binding dynamics. Measured parameters include the association rate constant  $k_a$  and total antigen concentration  $C_{tot,Ag}$  (in vitro). The three-compartment system is governed by:

$$\frac{dC_{free}}{dt} = -k(2, 1) \cdot C_{free} \quad (1)$$

$$\frac{dC_{bound}}{dt} = k(2, 1) \cdot C_{free} - k(3, 2) \cdot C_{bound} \quad (2)$$

$$\frac{dC_{int}}{dt} = k(3, 2) \cdot C_{bound} - k(0, 3) \cdot C_{int} \quad (3)$$

with the non-linear binding rate:

$$k(2, 1) = k_a \cdot (C_{tot, Ag} - C_{bound}) \quad (4)$$

Parameter values and initial conditions used for integration in our experiments (Patient 1) are listed in Table 2. Time-activity curves (TACs) computed from this model are used as temporal inputs to the MOSAIK learning model (see Section 3).

**Table 2:** Compartmental model parameters (Patient 1).

| Parameter        | Unit                  | Value                 |
|------------------|-----------------------|-----------------------|
| $k_a$            | $h^{-1} \cdot M^{-1}$ | $3.5 \times 10^{10}$  |
| $C_{tot, Ag}$    | $\mu M$               | 0.3662                |
| $k(3, 2)$        | $h^{-1}$              | 0.2                   |
| $k(0, 3)$        | $h^{-1}$              | $8.33 \times 10^{-4}$ |
| Molecular weight | $g \cdot mol^{-1}$    | $150 \times 10^3$     |
| Injected dose    | mg                    | 0.1                   |
| $C_{free, 0}$    | $\mu M$               | $7.32 \times 10^{-2}$ |
| $C_{bound, 0}$   | $\mu M$               | 0                     |
| $C_{int, 0}$     | $\mu M$               | 0                     |

#### MOSAIK MODEL: [<sup>89</sup>Zr]ZR-GIRENTUXIMAB UPTAKE PREDICTION

##### Multimodal input construction

The multimodal dataset integrates spatial, heterogeneity, and temporal information extracted from WSIs and the compartmental model (Section 2). Inputs are structured to retain tile-level granularity while incorporating local and global contextual cues.



170 **Spatial inputs:** The biological domain masks generated in Section 2.4 (neoplastic, non-neoplastic, and blood vessel compartments)  
171 were extracted for each tile. To incorporate multi-scale context, two additional masks were generated: a regional mask covering  
172 an 11×11 tile neighbourhood centred on the target tile, and a global mask covering 27×27 tiles. Both regional and global masks  
173 were downsampled to match the target tile size, enabling hierarchical feature extraction reflecting the pathologists multi-scale per-  
174 spective.

175 **Heterogeneity inputs:** Each tile is annotated with the hot/cold or outlier labels derived from local Morans I analysis for the three se-  
176 lected proxies (vascular, nuclear, CAIX). Labels were encoded numerically: Average = 0, Hot spot = 1, Cold spot = 2, High-low outlier  
177 = 3, Low-high outlier = 4. For each tile, 11×11 neighbourhood matrices were created for all three proxies, forming a 3-channel het-  
178 erogeneity input.

179 **Temporal inputs:** Time-Activity Curves (TACs) from the compartmental model were computed using LSODA integration of the sys-  
180 tem of equations (Section 2.6). The free, bound, and internalised antibody curves ( $C_{free}(t)$ ,  $C_{bound}(t)$ ,  $C_{int}(t)$ ) were sampled at patient-  
181 specific timepoints and normalised. These curves provide the temporal dynamics of antibody uptake at the tile level.

## 182 Model architecture

183 The MOSAIK architecture (Multi-modal Spatial Aggregation of Instance Knowledge) is a tile-level, multiple instance learning (MIL)  
184 framework designed to integrate spatial, heterogeneity, and temporal features into robust predictions of RPT uptake (see [Figure 5](#)  
185 for an overview of the model architecture).

186 **Spatial block:** Processes multi-scale masks for each tile via a tri-branch convolutional network. Each branch independently extracts  
187 features from tile, regional, and global masks. Feature maps are concatenated and passed through residual blocks with layer nor-  
188 malization to preserve spatial information.

189 **Heterogeneity block:** Handles the 3-channel heterogeneity matrices representing hot/cold/outlier classifications. A context-aware  
190 attention mechanism prioritises regions with significant intra-tumour variability, effectively making the model heterogeneity-attentive.  
191 Features from vascular, nuclear, and CAIX channels are embedded into a shared latent space and fused via cross-channel self-attention.

192 **Temporal block:** LSTM layers process TACs ( $C_{free}(t)$ ,  $C_{bound}(t)$ ,  $C_{int}(t)$ ) to capture temporal dependencies in uptake. Features from  
193 bound and internalised curves are concatenated and passed through fully connected layers, while  $C_{free}(t)$  is processed separately  
194 before integration.

195 **Feature aggregation and output:** Features from all three blocks are concatenated and passed through a lightweight feed-forward  
196 network. The model outputs two predictions per tile: (i) a classification of tile contribution to slide-level uptake (hot/cold/outlier), and  
197 (ii) regression of the SUV value. This dual-task design leverages multi-task learning to improve generalisation and predictive accu-  
198 racy.

## 199 Training procedure and optimization

200 The MOSAIK model was trained using a combination of classification and regression objectives. The joint loss function is:

$$\mathcal{L} = \alpha \mathcal{L}_{CE} + \lambda \mathcal{L}_{MSE}$$

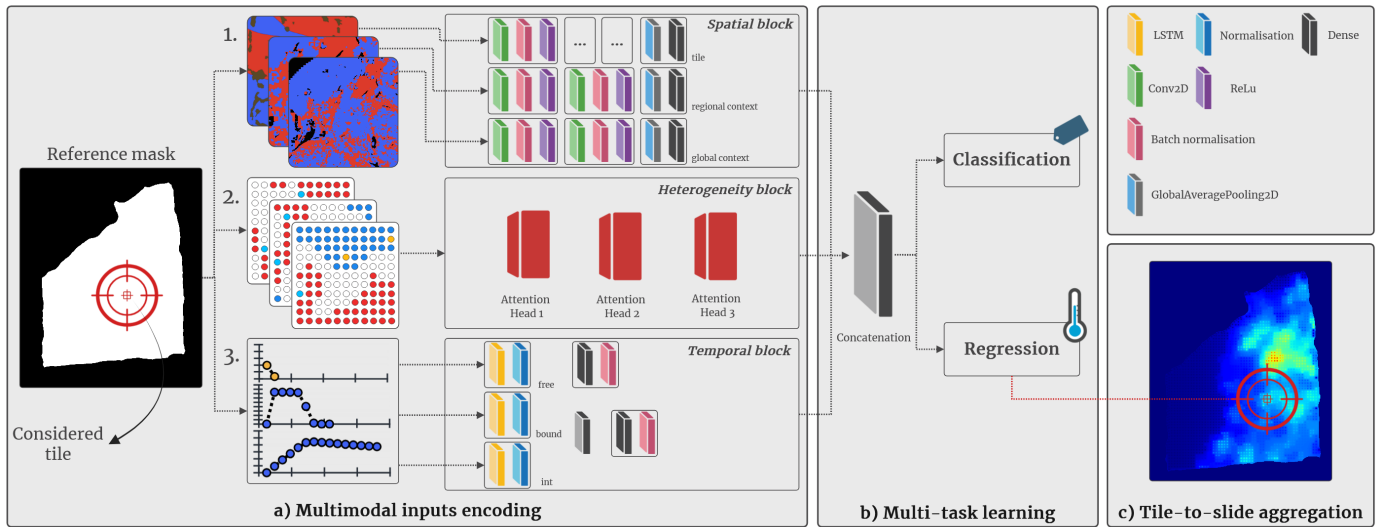
201 where  $\mathcal{L}_{CE}$  is the class-weighted cross-entropy loss for tile classification and  $\mathcal{L}_{MSE}$  is the mean squared error for SUV regression.  $\alpha$   
202 and  $\lambda$  are hyperparameters balancing the contributions of each task.

## 203 Training details:

204 Early stopping and model checkpoints were employed to prevent overfitting. Training used mini-batch stochastic gradient descent  
205 with the Adam optimizer, as specified in [Table 3](#). Input modalities were normalised to zero mean and unit variance to stabilise learn-  
206 ing.

## 207 Output predictions and learning tasks

208 MOSAIK performs two complementary tasks:



**Figure 5:** MOSAIK model architecture: a multi-modal, multi-instance learning framework combining spatial, heterogeneity, and temporal features to predict tile-level radiopharmaceutical uptake.

**Table 3:** MOSAIK model training parameters

|                           |                              |
|---------------------------|------------------------------|
| Train / Validation [%]    | 75.35 / 24.65                |
| Batch size                | 12                           |
| Optimizer                 | Adam                         |
| Learning rate             | $3 \times 10^{-5}$           |
| Loss classification       | categorical cross-entropy    |
| Loss regression           | mse                          |
| Activation classification | Softmax                      |
| Activation regression     | Sigmoid                      |
| $\alpha$                  | 0.3                          |
| $\lambda$                 | 0.7                          |
| Class weights             | [0.632, 0.916, 1.927, 5.336] |
| Number of epochs          | 25                           |

- Tile classification:** Assigns each tile to a class representing its contribution to slide-level uptake (hot, cold, high-low or low-high outlier). This coarse-grained task informs the global spatial heterogeneity of uptake.
- SUV regression:** Predicts the quantitative tracer uptake per tile, enabling fine-grained, spatially-resolved reconstruction of PET SUV maps.

The dual-task framework benefits from shared feature extraction while providing task-specific outputs. Classification guides the model to focus on heterogeneous regions, while regression provides continuous quantitative predictions. Aggregation of tile-level predictions allows reconstruction of slide-level uptake maps, providing interpretable spatial insight into RPT distribution across the tumour microenvironment.

## RESULTS AND DISCUSSION

In this data-driven approach, we developed a comprehensive framework to simulate the spatial distribution of [ $^{89}\text{Zr}$ ]Zr-girentuximab within ccRCC tumours. Our workflow integrates a compartmental model with analyses of H&E, CD34-immunostained, and CAIX-immunostained digital histologies from two patients at different disease stages. This methodology translates extracted tiles into interpretable instances reflecting tissue behavior within tumours. Unlike conventional WSI analyses that directly extract patterns from patches after stain normalization, our approach explicitly considers the compartments of the immunokinetic model governing radiolabelled antibody diffusion and retention over time alongside a local characterization of the biological domain and intra-tumour heterogeneity parameters. Furthermore, the digital twin model was designed as Heterogeneity-Attentive and Context-Aware, enabling the integration of microscopic and macroscopic information. In the following subsections, we present the results collected

226 across the processing pathways, the input construction, and model validation.

227 **Tissue Characterization**

228 Accurate pathological segmentation is pivotal in our framework, as it enables the study of drug interactions with distinct tissue types,  
229 particularly differentiating tumour and non-tumour regions. Preserving spatial patterns at the slide scale requires providing the  
230 model with robust, spatially informative content. The fine-tuned model exhibited strong performance on the validation dataset,  
231 demonstrating its reliability and generalizability to large-scale slides. Metrics on accuracy and precision are summarized in [Table 4](#).

**Table 4:** Compartments segmentation performance

| Per-class metrics | Background | Neoplastic | Non neoplastic |
|-------------------|------------|------------|----------------|
| Accuracy          | 0.968      | 0.948      | 0.958          |
| Precision         | 0.964      | 0.822      | 0.975          |
| Overall metrics   |            |            |                |
| Accuracy          | 0.937      |            |                |
| Precision         | 0.929      |            |                |

232 These metrics demonstrate the models capability to generate accurate tile-level segmentation masks. The aggregated slide-level  
233 map in [Figure 10](#) provides a cohesive representation of the pathological landscape.

234 **Macroscopic Proxies for Heterogeneity**

235 The heterogeneity analysis aimed to identify parameters with the highest spatial correlation to observed drug uptake patterns, high-  
236 lighting factors underlying tissue retention mechanisms.

237 **Parameter Selection** Histological regions (A2A5 for patient 1 and B8 & B11 for patient 2) showed variable correlations, both posi-  
238 tive and negative, with baseline micro PET RPT accumulation. [Figure 7](#) presents a Spearman correlation heatmap for nuclear-derived  
239 parameters and the CAIX score, while [Figure 8](#) shows Pearson correlations for vascular parameters. Blue hues indicate positive cor-  
240 relations, red hues negative correlations. The dual-axis scatterplots in [Figure 9](#) highlight the parameters optimizing both correlation  
241 metrics, revealing spatially relevant features for drug uptake: nuclear difference entropy, vascular equivalent diameter, and the CAIX  
242 score. Notably, CAIX expression showed a high average correlation with baseline uptake, confirming its value as a clinical biomarker  
243 for ccRCC targeting. These findings underscore the importance of capturing spatial heterogeneity across scales, aligning with previ-  
244 ous microscopic studies.

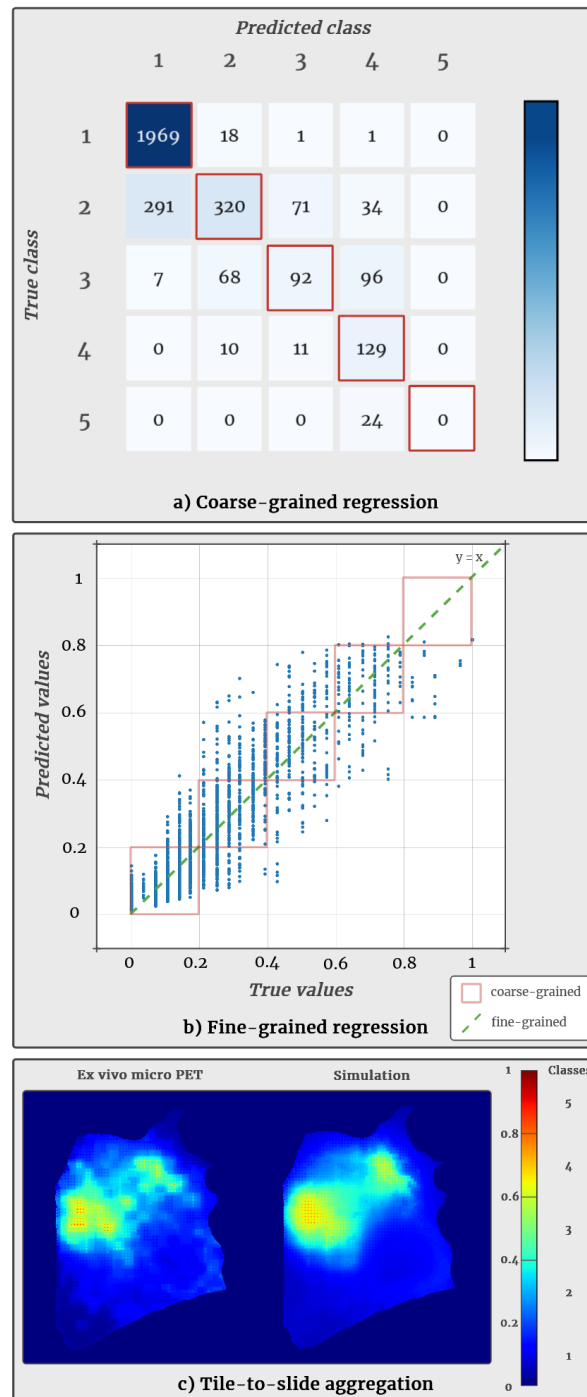
245 **Spatial Cluster Identification** Parameters identified through correlation analyses underwent local Moran's Index evaluation, gen-  
246 erating tile-wise heterogeneity maps with hot and cold clusters, including outliers ([Figure 10](#)). These cluster/outlier labels ensure  
247 the digital twin appropriately considers tumour heterogeneity at multiple spatial scales.

248 **Digital Twin Model Evaluation**

249 Predictions were generated by applying the trained digital twin to tiles from an unseen WSI (slide B11). Tile-level predictions were  
250 aggregated and interpolated using the IDW method to form the predicted RPT uptake map, compared to the ground truth in [Fig-  
251 ure 11](#).

252 The simulation demonstrates the models ability to predict intra-tumour uptake patterns with high fidelity at the slide scale. For SUV  
253 values normalized in [0,1], the MSE for slide B11 was 0.0065 (MAE = 0.0591). Contextual integration from tile-level to broader re-  
254 gions contributed to global consistency, avoiding overprediction in low-uptake areas. However, intermediate and high-variation  
255 regions were occasionally overpredicted. The pervasive nature of heterogeneityoperating locally, regionally, and globally, as well as  
256 inter-patient variabilitylimits the models ability to fully capture complex spatial patterns.

257 Despite the small cohort, these results support the feasibility of using Deep Learning to reveal critical determinants of therapeutic  
258 retention at the tissue level, potentially guiding clinical decision-making. Future work should incorporate dynamic imaging to bet-  
259 ter inform temporal aspects of drug distribution and account for inter-patient variability. Additionally, integrating spatial measure-

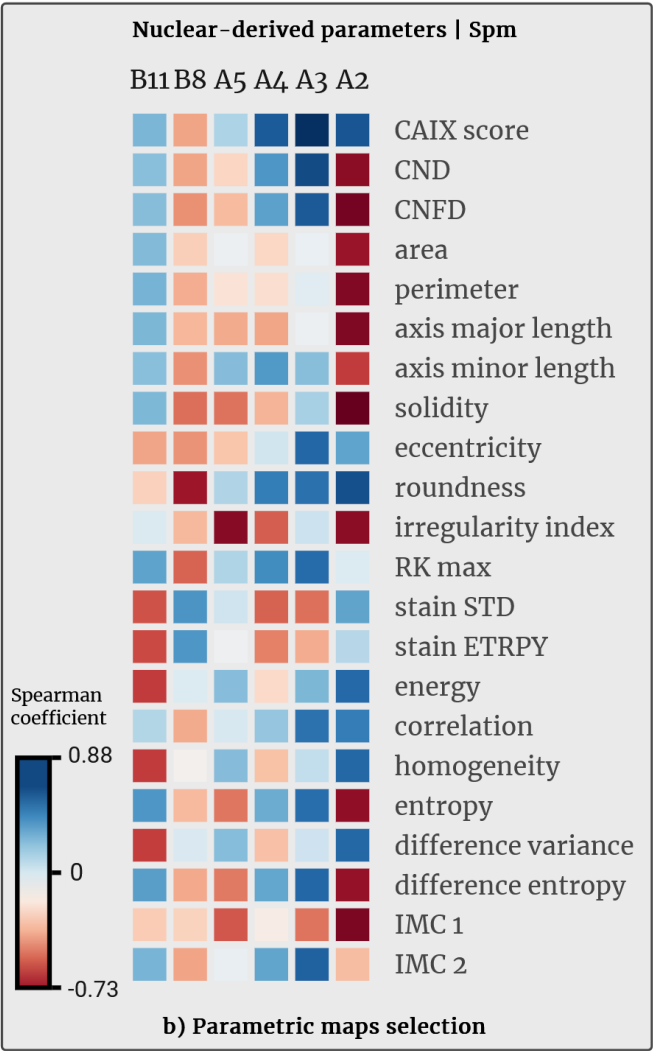


**Figure 6:** Tissue characterization process on slide B11: 1. H&E slide thumbnail, 2. Global tissue segmentation mask.

ments of tumour grade could enhance model interpretability and generalizability, mitigating the risk of learning artifacts and improving reliability for clinical applications.

## CONCLUSIONS

As radiopharmaceuticals advance precision medicine in oncology, data-driven models are increasingly pivotal for predicting therapeutic outcomes. In this work, we implemented a comprehensive framework to account for multiple aspects of radiopharmaceutical absorption within tumours, providing a novel computational pathway to enhance precision oncology. Our theranostic digital twin integrates both spatial and temporal dimensions through a multimodal, multi-input approach, capturing the complex factors that govern drug diffusion in tumour tissues. By explicitly considering the spatial organization of the tumour microenvironment,



**Figure 7:** Spatial correlation heatmap for nuclear-derived parameters and CAIX score across slides.

our model provides insights into tumour heterogeneity a key contributor to therapeutic failure in refractory and metastatic cancers. Furthermore, by incorporating the immunokinetic model of the radiotracer, we account for the mechanisms driving retention kinetics within tumour tissues. Collectively, this work lays the foundation for predictive evaluation of treatment efficacy, enabling exploration of intra-tumour uptake patterns and supporting optimized strategies for improved cancer management.

ACKNOWLEDGMENTS

We acknowledge the departments of Nuclear Medicine and Pathological Anatomy for providing the data and all the personnel who helped in collecting them for this study at IRCCS Ospedale San Raffaele, Milan. L. Cavinato is funded by AdvanCed Technologies for Human-centrEd Medicine grant. L. Cavinato acknowledges the support by MUR, grant Dipartimento di Eccellenza 20232027.

REFERENCES

[1] Raffaele Colombo and Jamie R Rich. The therapeutic window of antibody drug conjugates: a dogma in need of revision. *Cancer Cell*, 40(11):1255–1263, 2022.

[2] U Iyer and VJ Kadambi. Antibody drug conjugate trojan horses in the war on cancer. *Journal of pharmacological and toxicological methods*, 64(3):207–212, 2011.

[3] Daniele Giansanti and Sandra Morelli. Exploring the potential of digital twins in cancer treatment: A narrative review of reviews. *Journal of Clinical Medicine*, 14(10):3574, 2025.

[4] Hengyu Mao. Clinical relevance of mutant-allele tumor heterogeneity and lung adenocarcinoma. *Annals of Translational Medicine*, 7(18):432, 2019.

[5] Nicholas McGranahan and Charles Swanton. Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. *Cancer cell*, 27(1):15–26, 2015.

[6] Aditya Kashyap, Maria Anna Rapsomaniki, Vesna Barros, Anna Fomitcheva-Khartchenko, Adriano Luca Martinelli, Antonio Foncubierta Rodriguez, Maria Gabrani, Michal Rosen-Zvi, and Govind Kaigala. Quantification of tumor heterogeneity: from data acquisition to metric generation. *Trends in Biotechnology*, 40(6):647–676, 2022.

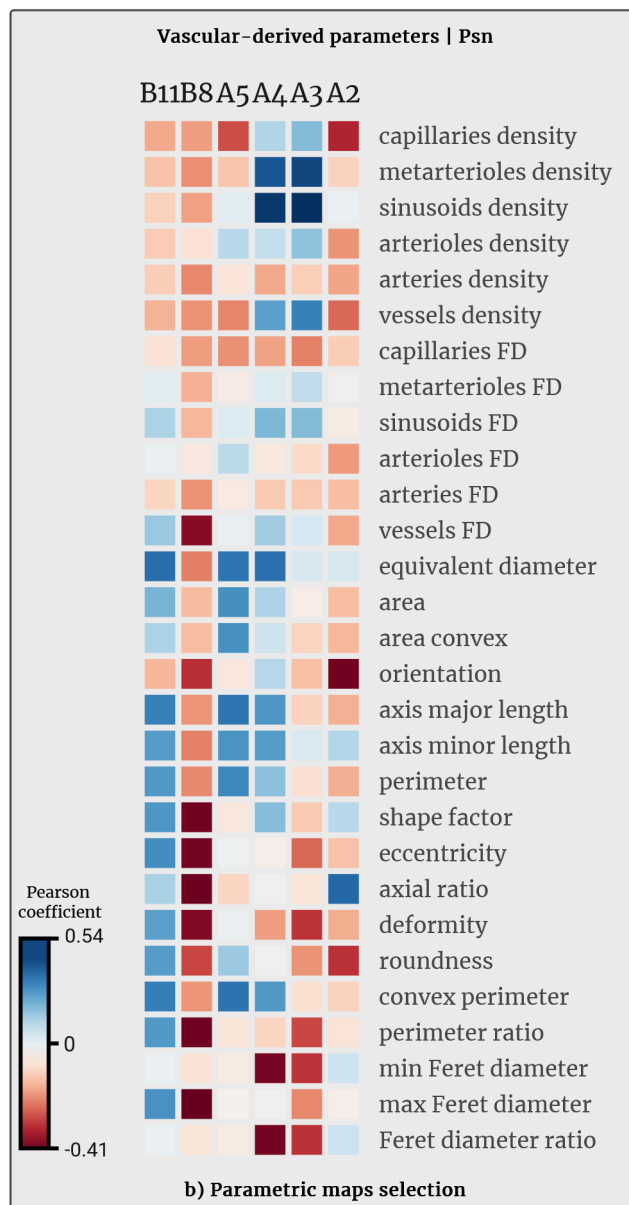
[7] Melissa R Junttila and Frederic J De Sauvage. Influence of tumour micro-environment heterogeneity on therapeutic response. *Nature*, 501(7467):346–354, 2013.

[8] Simon JA van Kuijk, Ala Yaromina, Ruud Houben, Raymon Niemans, Philippe Lambin, and Ludwig J Dubois. Prognostic significance of carbonic anhydrase ix expression in cancer patients: a meta-analysis. *Frontiers in oncology*, 6:69, 2016.

[9] Lauri Louhivuori, Shigeaki Kanatani, and Per Uhlén. Predicting a tumours drug uptake. *Nature biomedical engineering*, 2(10):717–718, 2018.

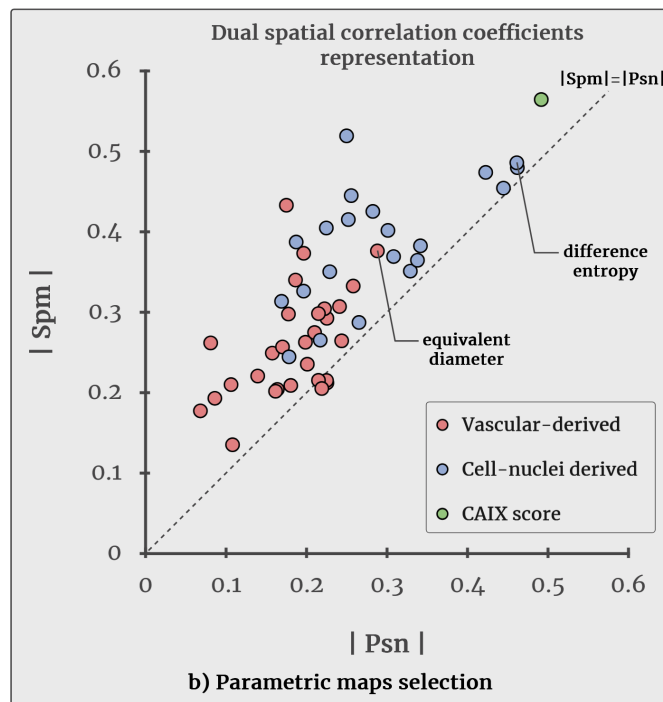
[10] Damien J McHugh, Grazyna Lipowska-Bhalla, Muhammad Babur, Yvonne Watson, Isabel Peset, Hitesh B Mistry, Penny L Hubbard Cristinacce, Josephine H Naish, Jamie Honeychurch, Kaye J Williams, et al. Diffusion model comparison identifies distinct tumor sub-regions and tracks treatment response. *Magnetic Resonance in Medicine*, 84(3):1250–1263, 2020.

[11] Weiqin Zhao, Ziyu Guo, Yinshuang Fan, Yuming Jiang, Maximus CF Yeung, and Lequan Yu. Aligning knowledge concepts to whole slide images for precise histopathology image analysis. *npj Digital Medicine*, 7(1):383, 2024.



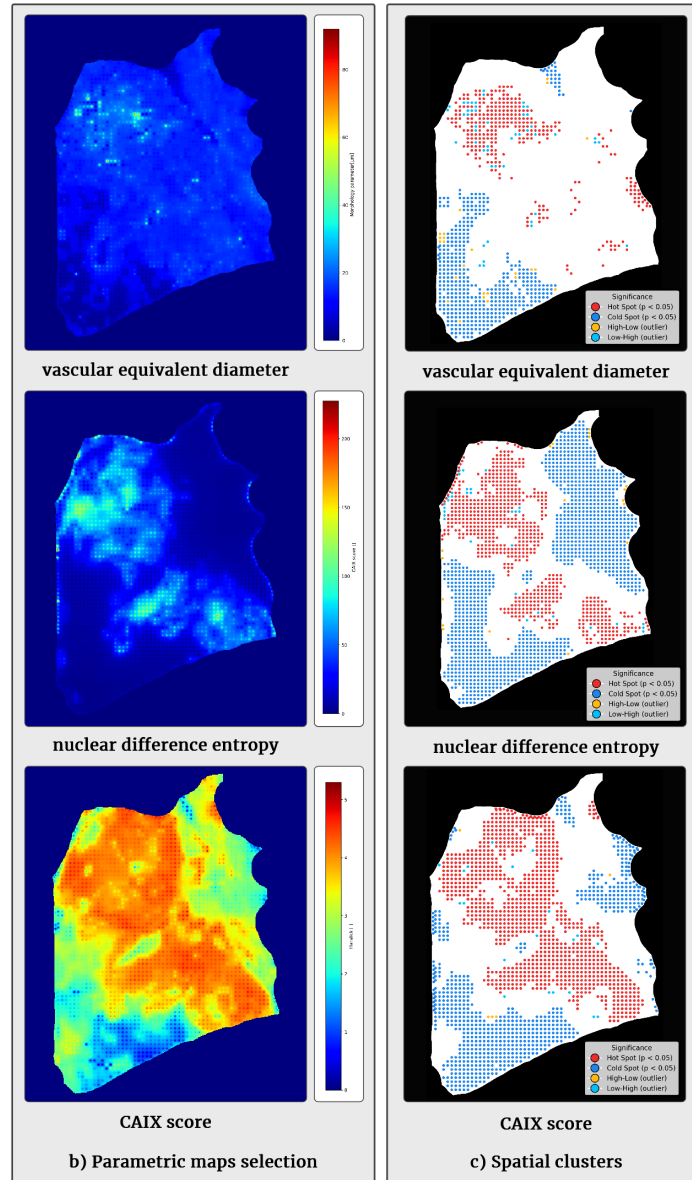
**Figure 8:** Spatial correlation heatmap for vascular-derived parameters across slides.

- [12] Imke H Bartelink, Ella F Jones, Sheerin K Shahidi-Latham, Pei Rong Evelyn Lee, Yanan Zheng, Paolo Vicini, Laura van t Veer, Denise Wolf, Andrei Iagaru, Deanna L Kroetz, et al. Tumor drug penetration measurements could be the neglected piece of the personalized cancer treatment puzzle. *Clinical Pharmacology & Therapeutics*, 106(1):148–163, 2019.
- [13] Hamed Rezayatmand, Mahboobeh Razmkhah, and Iman Razeghian-Jahromi. Drug resistance in cancer therapy: the pandora’s box of cancer stem cells. *Stem cell research & therapy*, 13(1):181, 2022.
- [14] Afshin Derakhshani, Zeinab Rostami, Hossein Safarpour, Mahdi Abdoli Shadbad, Niloufar Sadat Nourbakhsh, Antonella Argentiero, Sina Taefehshokh, Neda Jalili Tabrizi, Omid Kooshkaki, Reza Vaezi Astamal, et al. From oncogenic signaling pathways to single-cell sequencing of immune cells: changing the landscape of cancer immunotherapy. *Molecules*, 26(8):2278, 2021.
- [15] Marc Macenko, Marc Niethammer, James S Marron, David Borland, John T Woosley, Xiaojun Guan, Charles Schmitt, and Nancy E Thomas. A method for normalizing histology slides for quantitative analysis. In *2009 IEEE international symposium on biomedical imaging: from nano to macro*, pages 1107–1110. IEEE, 2009.
- [16] Bronson Flint, Joe M Das, and Carrie A Hall. Body surface area. In *StatPearls [Internet]*. StatPearls Publishing, 2025.
- [17] Seol Ah Park, Yongsoo Kim, Pawel Dlotko, Davide Gurnari, and Jooyoung Hahn. Spatial analysis of malignant-immune cell interactions in the tumor microenvironment using topological data analysis. *bioRxiv*, pages 2025–03, 2025.
- [18] Katarzyna A Rejniak, Veronica Estrella, Tingan Chen, Allison S Cohen, Mark C Lloyd, and David L Morse. The role of tumor tissue architecture in treatment penetration and efficacy: an integrative study. *Frontiers in oncology*, 3:111, 2013.
- [19] Franziska Büschek, Christoph Fraune, Ronald Simon, Martina Kluth, Claudia Hube-Magg, Christina Möller-Koop, Navid Shadanpour, Clemens Bannenberg, Christian Eichelberg, Doris Höflmayer, et al. Aberrant expression of membranous carbonic anhydrase ix (caix) is associated with unfavorable disease course in papillary and clear cell renal cell carcinoma. In *Urologic Oncology: Seminars and Original Investigations*, volume 36, pages 531–e19. Elsevier, 2018.
- [20] Sarah M Cheal, Blesida Punzalan, Michael G Doran, Michael J Evans, Joseph R Osborne, Jason S Lewis, Pat Zanzonico, and Steven M Larson. Pairwise comparison of 89zr- and 124i-labeled cg250 based on positron emission tomography imaging and nonlinear immunokinetic modeling: in vivo carbonic anhydrase ix receptor binding and internalization in mouse xenografts of clear-cell renal cell carcinoma. *European journal of nuclear medicine and molecular imaging*, 41(5):985–994, 2014.

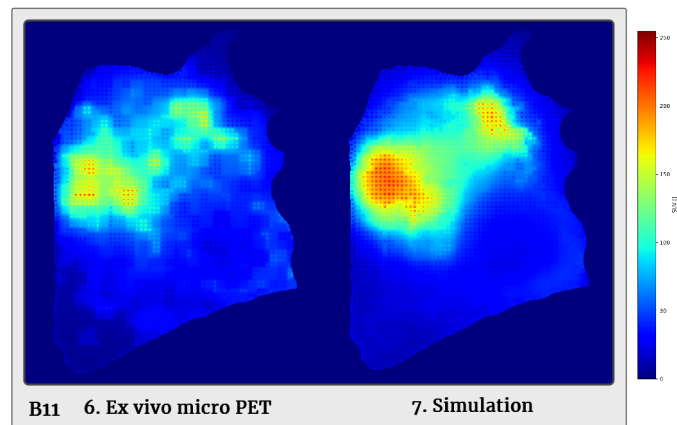


**Figure 9:** Average spatial correlation metrics between parameter mappings and baseline drug uptake maps. The equivalent diameter stands out among vascular-derived parameters, while difference entropy is prominent among nuclear-derived parameters. (Psn: Pearson, Spm: Spearman)





**Figure 10:** Selected parameters from the heterogeneity analysis with tile-wise Moran's I maps: 3. Vascular equivalent diameter mapping and clusters, 4. Nuclear difference entropy mapping and clusters, 5. CAIX score mapping and clusters.



**Figure 11:** Validation of simulated RPT uptake against ground truth: 6. Micro PET image of intra-tumour  $[^{89}\text{Zr}]$ -girentuximab uptake, 7. Aggregated tile-level predictions.

## MOX Technical Reports, last issues

Dipartimento di Matematica  
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

- 64/2025** Celora, S.; Tonini, A.; Regazzoni, F.; Dede', L. Parati, G.; Quarteroni, A.  
*Cardiocirculatory Computational Models for the Study of Hypertension*
- 63/2025** Panzeri, S.; Clemente, A.; Arnone, E.; Mateu, J.; Sangalli, L.M.  
*Spatio-Temporal Intensity Estimation for Inhomogeneous Poisson Point Processes on Linear Networks: A Roughness Penalty Method*
- 62/2025** Rigamonti, V.; Torri, V.; Morris, S. K.; Ieva, F.; Giaquinto, C.; Donà, D.; Di Chiara, C.; Cantarutti, A.; CARICE study group  
*Real-World Effectiveness of Influenza Vaccination in Preventing Influenza and Influenza-Like Illness in Children*
- 61/2025** Coclite, A; Montanelli Eccher, R.; Possenti, L.; Vitullo, P.; Zunino, P.  
*Mathematical modeling and sensitivity analysis of hypoxia-activated drugs*
- 60/2025** Antonietti, P. F.; Caldana, M.; Gentile, L.; Verani M.  
*Deep Learning Accelerated Algebraic Multigrid Methods for Polytopal Discretizations of Second-Order Differential Problems*
- 59/2025** Wolf, F.; Botteghi, N.; Fasel, U.; Manzoni, A.  
*Interpretable and efficient data-driven discovery and control of distributed systems*
- 58/2025** Pivato, C.A.; Cozzi, O.; Fontana, N.; Ieva, F., et al.  
*Clinical outcomes of percutaneous coronary interventions after transcatheter aortic valve replacement*
- 57/2025** Guagliardi, O.; Masci, C.; Breschi, V; Paganoni A, ; Tanelli, M.  
*A Novel DNA-Inspired Framework to Study University Dropout: Insights from Politecnico di Milano*
- 56/2025** Tonini, A.; Bui-Thanh, T.; Regazzoni, F.; Dede', L; Quarteroni, A.  
*Improvements on uncertainty quantification with variational autoencoders*
- Tonini, A.; Bui-Thanh, T.; Regazzoni, F.; Dede', L; Quarteroni, A.  
*Improvements on uncertainty quantification with variational autoencoders*