A multiscale model of pulmonary micrometastasis and immune surveillance: towards cancer patient digital twins

Heber L. Rocha[1], Boris Aguilar[2†], Michael Getz[1†], Ilya Shmulevich[2†] and Paul Macklin[1†]

[1]Intelligent Systems Engineering, Indiana University, Bloomington, IN, USA.
[2]Institute for Systems Biology, Seattle, WA, USA.

Contributing authors: hlimadar@iu.edu; boris.aguilar@systemsbiology.org; getzm@iu.edu; ilya.shmulevich@systemsbiology.org; macklinp@iu.edu;
[†]These authors contributed equally to this work.

Abstract

Metastasis is the leading cause of death in patients with cancer. Among the targets of this spread, the lung is one of the most frequent targets of metastasis. In the scientific community, many studies have been developed to study the process of cancer dissemination in the lung. We investigated this process by creating a multiscale mathematical model to study the interactions between the immune system and the progression of micrometastases in the lung. We analyzed the parameter space of the model using high-throughput computing resources to generate over 100,000 virtual patient trajectories. We demonstrated that the model could recapitulate a wide variety of virtual patient trajectories, uncontrolled growth, and partial and complete immune response to tumor growth. We classified the patients and identified key patient parameters that regulate immunosurveillance. We highlight the lessons derived from this analysis and outline the primary challenges in building cancer patient digital twins. These digital twins would enable clinicians to systematically dissect the complexity of cancer in each individual patient, simulate treatment outcomes, and ultimately select the most suitable treatment.

Keywords: micrometastases, multiscale mathematical model, virtual patient, cancer patients digital twins

1 Introduction

Metastasis constitutes the most fatal manifestation of cancer and remains among the least understood aspects of cancer pathophysiology. Early metastasis is often undetected at primary diagnosis, leading to an increased risk of mortality [1, 2]. Identifying occult metastases in early-stage cancer is crucial for optimal therapy and prognosis [3, 4]. The metastasis process is highly inefficient due to its complexity [5–7]. At each step of the metastatic cascade, the host immune system can potentially recognize and kill mutant cancer cells, which may also be immunogenic [5, 8]. Nevertheless, cancer cells still find ways to evade host immune surveillance in distant sites. Understanding the complexity of these processes can lead to the development of novel therapeutic strategies to benefit patients with micrometastasis.

The interactions between metastatic cancer and the immune system (with or without therapeutic modulation) are inherently multi-scale, involving molecular-, cell-, multicellular-, and systems-scale processes at fast and slower time scales, making cancer immunology the quintessential multiscale complex systems problem of our era. Computational models can serve as a virtual laboratory to understand and ultimately control the behavior of such complex systems. When the computational model can be calibrated using individual cancer patient data, predict cancer progression across a range of treatment options, and be refined with new patient measurements, the model evolves into a Cancer Patient Digital Twin (CPDT). This paradigm enables clinicians to systematically analyze the complexity of an individual patient's cancer, simulate treatment outcomes, and ultimately choose the best treatment to achieve the patient's goals [9–13]. Multi-scale models are of special interest for the development of CPDTs as they can integrate several relevant interactions, often from different temporal scales, into one simulation framework. Specifically, multi-scale models can integrate molecular and cellular level interactions between cancer cells and the immune system. In this work, we developed a multiscale model of metastatic cancer cells developing in the lung alveolar epithelium, focusing on the interactions between immune and cancer cells.

Recently, several mathematical models have been developed to help understand the complex crosstalk between the tumor and the immune system [14–18]. In particular, multiscale hybrid models have been frequently applied due to their modularity and versatility [14, 15, 19–21]. Makarya et al. [22] highlight how multiscale mechanistic models can represent different aspects of immune system cells and suggest areas of study in immune-cancer interactions. On the other hand, hybrid multiscale models require estimating parameters that modulate several relevant processes, including cell-cell communication, intracellular gene regulation, cell division, cell death, oxygen consumption, interactions, etc., resulting in highly complex models. Furthermore, the calibration and validation of these models become a challenging problem due to data limitations and the substantial computational cost of their simulations

[23–25]. Consequently, it is essential to integrate machine learning and high-performance computing (HPC) techniques into creating virtual patient models [23, 26].

We built a multiscale agent-based model of pulmonary micrometastases with local and systems-scale immune interactions, including mechanical pressure induced cell death [27–30], secretion of pro-inflammatory cytokines, recruitment, infiltration of immune cells, and other mechanisms. Using HPC based simulations and an ensemble modeling approach, whereby multiple simulations are aggregated into statistically summarized results, we verified that the model is capable of capturing clinically salient outcomes, including uncontrolled growth, partial tumor control, and complete tumor elimination. We also varied and explored ten key immune system parameters on HPC resources to analyze the digital patient trajectories. From our dataset generated from parametric scanning, we selected a subset of virtual patients where cancer immunosurveillance failed and identified crucial parameters from our multiscale model that may boost cancer immunosurveillance.

2 Results

2.1 The model overview of pulmonary micrometastases

The general dynamic of the model results from a series of rules that characterize crucial processes of immune-cancer cell interactions. We assume that immune cells and healthy lung epithelial cells are in homeostasis before the formation of micrometastases. Cancer cells proliferate uncontrollably due to their high reproductive capacity, causing mechanical stress in the region of colonies. This mechanical stress leads to the death of adjacent healthy cells in the region, resulting in tissue damage and stimulating the activation of the immune system. The damaged tissue has a high concentration of cellular debris that stimulates the infiltration of macrophages and dendritic cells into the region. Moreover, macrophages and dendritic cells are activated through phagocytosis and contact, respectively. Macrophages carry out the phagocytosis process by ingesting dead cell waste and releasing tumor necrosis factor (TNF), a cytokine used for cell signaling, which leads to the recruitment of more immune cells to the region. This describes the transition process from the non-activated macrophage phase (M0) to the pro-inflammatory phase (M1). Dendritic cells (DCs) are activated when in contact with dying cells (lung epithelial or cancer cells) and process their antigen material. For simplicity, we consider that all cancer cells present the same antigen. Activated dendritic cells migrate to the lymph node and present the antigen on their surface to T cells, promoting the activation and proliferation of helper and cytotoxic T cells. Consequently, the lymph node sends CD8+ (cytotoxic) and CD4+ (helper) T cells to the tumor microenvironment. CD8+ T cells perform two functions in the tumor microenvironment: (i) they kill cancer cells when they come into contact with them and (ii) inhibit the production of TNF from polarized macrophages, representing the conversion of classically activated macrophages (M1) to the

anti-inflammatory state (M2) [31]. CD4 + T cells further enable the hyperactivation of macrophages, promoting their ability to engulf dead cells and live cancer cells [32] (Figure 1, more details in Methods Sections 4.1 to 4.7).

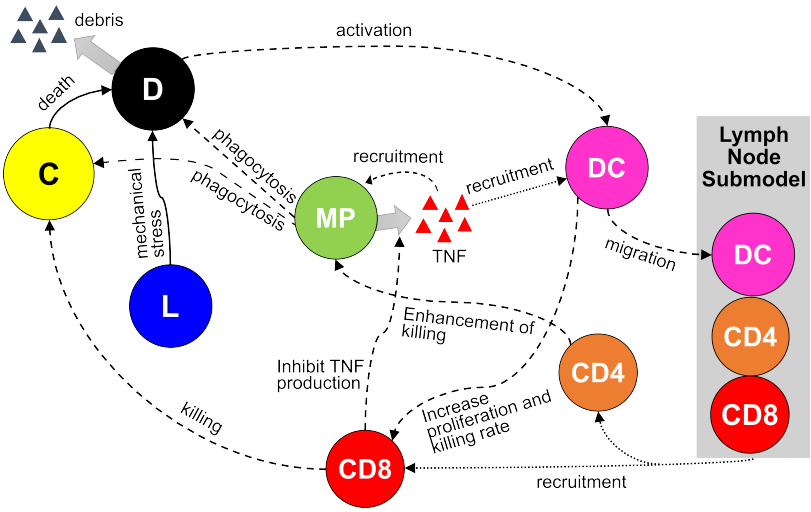

Fig. 1: Schematic diagram of the multiscale model of pulmonary micrometastasis. The contact of dendritic cells with the dying cells and the phagocytosis process by macrophages activate the immune response. Once the immune system is activated, the antigen presentation by DCs in the lymph node results in T Cell recruitment, which may extinguish cancer cells. Macrophages potentialize the immune response by releasing pro-inflammatory cytokines that increase the recruitment of immune cells.

2.2 Plausible patient outcomes

To investigate immunosurveillance in response to pulmonary micrometastasis growth, we explore a wide range of overall dynamics by adjusting various model parameters. Note that each parameter set can be seen as a distinct virtual patient, illustrating the spectrum of potential outcomes in virtual patient scenarios. In specific parameter configurations (additional information available in section 1 of the supplementary material), we observed varying degrees of immune response during the progression of cancer.

We begin with a common initial state for all virtual patients. In this state, our model represents a region of lung tissue primarily composed of healthy lung epithelial cells, inactive immune cells, and a small number of metastatized cancer cells. These cells are randomly distributed in the microenvironment, as shown in Figure 2A. We then analyze the patient's trajectories over a ten-day period, examining the variability in the responses of the immune system

to micrometastatic development. Figure 2B demonstrates the simulation outcomes for specific virtual patients, highlighting three distinct intensities of immune response over time.

For the patient exhibiting a low immune response (bottom row in Fig. 2B), we observe a typical scenario of tumor escape, marked by the absence of immune cells and the prevalence of cancer cells in the microenvironment. The cancer progression exerts mechanical pressure on the surrounding tissue, resulting in the death of healthy cells and the creation of space for cancer cell proliferation. The stiffness of the tissue acts as a constraint on cancer progression, preventing excessive cellular agglomeration. However, the absence of macrophages precludes TNF production in the microenvironment, thus preventing the immune system activation and the production of adequate T cells required to combat the tumor.

In contrast, in the case of the patient displaying a moderate immune response (middle row in Fig. 2B), we observe a scenario of partial elimination. This is characterized by the presence of active macrophages inducing TNF production and recruiting additional immune cells, leading to substantial immune control. However, the effective immune response occurs late, coinciding with significant damage to lung tissue. This delayed response may be attributed to suboptimal DC activation or T cell production in the lymph nodes.

For the patient with a high immune response (top row in Fig. 2B), we observe a scenario of complete elimination, where the immune system effectively controls tumor growth. Early on, CD8 + T cells dominate the microenvironment, offering significant resistance against cancer cells. Furthermore, due to early immune activation, the tumor is eradicated before the ten-day simulation period concludes.

2.3 Virtual patient classification exposes limitations in conventional stratification

Utilizing virtual patient simulations generated from our parameter exploration (see Methods Section 4.8), we classified all 100k trajectories according to the immune response (Figure 3A, details in Methods Section 4.9). This resulted in 41.8% of simulations being categorized as significant control (SC), 45.1% as no control (NC), and 13% as marginal control (MC). Upon reducing the dimensionality of all trajectories, we noticed two well-defined clusters, with one of them having the three classes (NC, MC, SC) separated into distinct regions and another one with the majority of trajectories classified as no control (Figure 3B). We investigated the differences between these two clusters of NC trajectories and found evidence that they differ in terms of immune system activation. While the cluster with all classes has an activated immune system, the other one is deactivated or in the process of deactivation (more details in section 3 of the supplementary material).

Conflicts arose in the classification of replicates for individual virtual patients. This is because the stochastic nature of immune-tumor interactions introduces substantial uncertainty in the outcomes. Notably, out of the 10,000

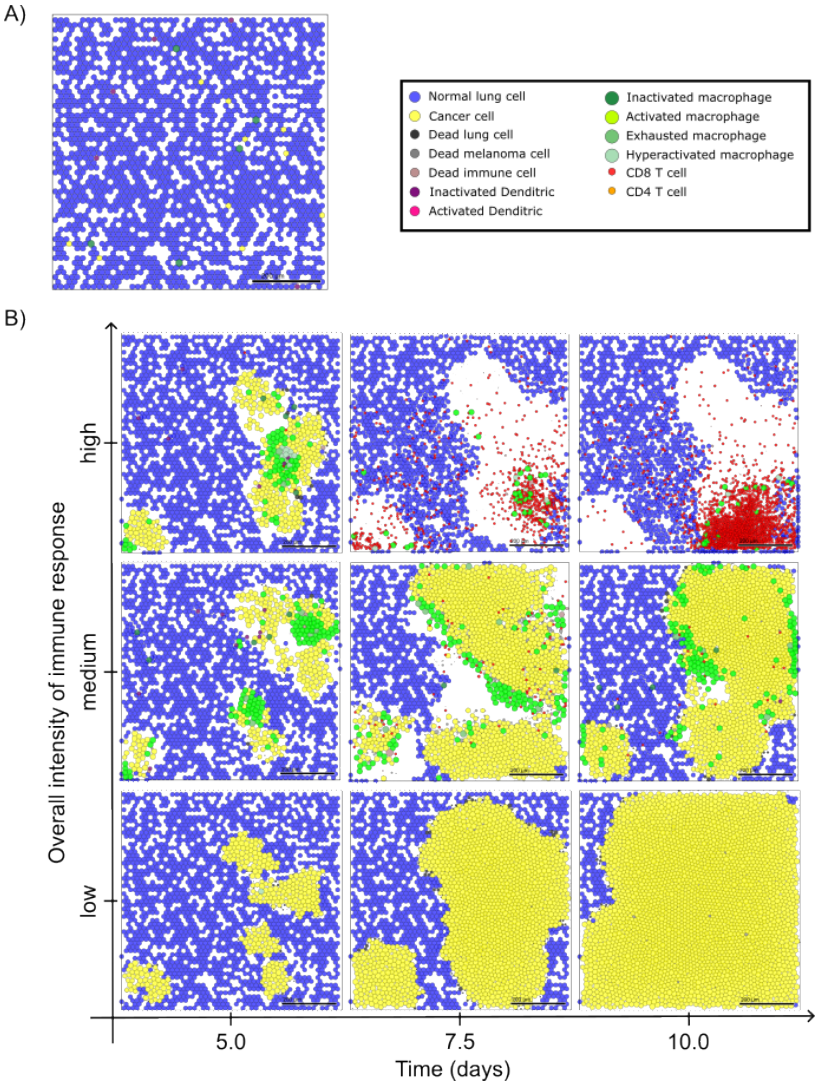

Fig. 2: Multiscale model outcomes. A) Initial condition of multiscale model and legend of cells. B) Simulation snapshots under varying levels of immune response. From left to right, the days 5, 7.5, and 10 of the dynamics.

virtual patients, only 47.1% exhibited consistent classification for all replicates within the same category (Figure 3C). For the remaining patients (52.9%), a unique outcome label was not assigned uniformly across all replicates.

When classifying patients based on the entirety of replicates and visualizing the average trajectories in a reduced space, we observed that mainly NC and SC patient labels preserve distinct regions in the reduced space. Patients with multiple labels (NC+MC, NC+SC, MC+SC, and NC+MC+SC) and those

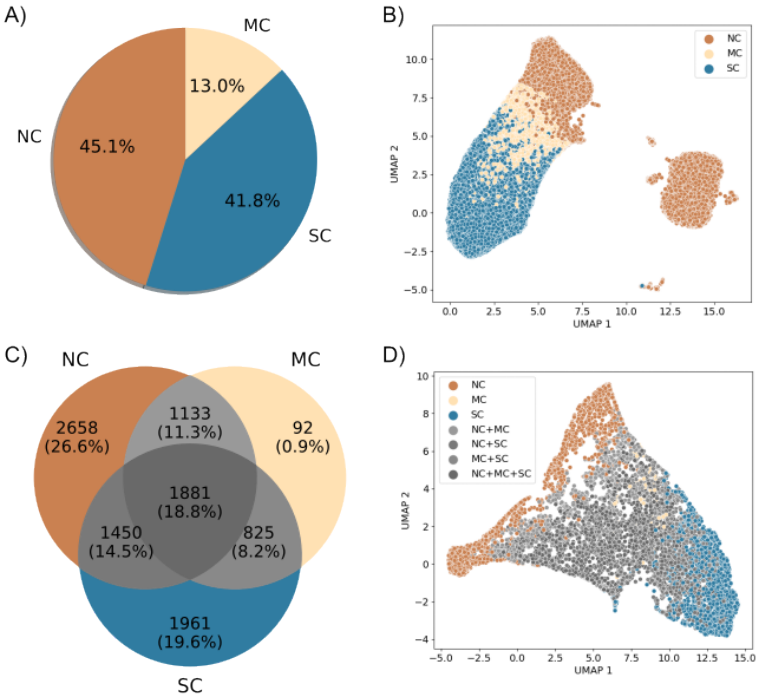

Fig. 3: Virtual patient classification. A) Classification of all 100k simulations according to the thresholds. B) Dimensionality reduction (UMAP) of the spatial/temporal trajectories of all simulations. C) Venn diagram of classification of the 10k virtual patients based on the totality of replicates. D) Dimensionality reduction (UMAP) of the averages of virtual patients trajectories.

with the MC label are intermingled in the reduced space (see Figure 3D). This is because of the lower variances in the immune responses of patients with extreme labels: NC and SC. On the other hand, for some patients, their outcomes have significant variance. For instance, patient 6 (Figure 4) displays sources of instability in the immune response and concurrently exhibits the three labels SC, MC, and NC. The presence of multiple outcomes in this patient can be attributed to substantial variations within the immune cell population, especially during the early and late stages of dendritic cells (DCs) and macrophages activation. This suggests variability in CD8 T cell responses, as illustrated in Figure 4C. Consequently, these variances lead to varying intensities of immune responses, which align with the notable variances observed in the cancer cell population, as depicted in Figure 4A. It is worth noting that the production of T cells in the lymph node does not display substantial variability in this specific patient (Figure 4B).

Considering the variability of patient outcomes, it becomes evident that the usual patient stratification cannot be successfully applied to all virtual

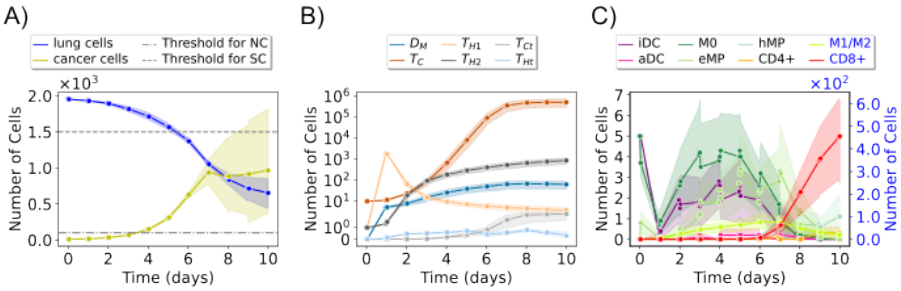

Fig. 4: Cell populations over time of an undetermined virtual patient. A) The cell population in the lymph node. B) The population of cancer cells and tissue lung cells. C) The population of immune cells (iDC: inactive DC, aDC: active DC, eMP: exhausted macrophages, and hMP: hyperactivated macrophages).

patient datasets. This suggests that more information about the patient may be necessary to reduce outcome uncertainty. The fusion of multiple data sources into our model is the target of future investigation.

2.4 Sensitivity analysis reveals key patient characteristics for immunotherapy

In the parameter space of the virtual patients, we identify the locality, spread, and skewness of the ten parameters according to the labeled patient trajectories (see Figure 5A), and perform a statistical hypothesis test to examine the relationship between the classes: No Control (NC) and Significant Control (SC). Considering as null hypothesis that the distribution of each parameter is the same for both classes, we performed the Mann-Whitney U test, according to a confidence level of 95%, we obtain that the parameters r_{leave} (rate of activated DCs leaving the tissue and migrating into the lymph node) and δ_C (natural decay of T cells on lymph node) are not statistically different in either class (p-value ≥ 0.05), while the others parameters are statistically different (p-value < 0.05). However, the parameters $r_{recruit}[D]$ (recruitment rate of dendritic cells) and $r_{recruit}[M]$ (recruitment rate of macrophages) have greater statistical significance in our parametric space. This suggests that the process of recruitment of macrophages and dendritic cells is a crucial component of our model. Their strong associations with patient labels and the outcomes further support our findings regarding a specific undetermined virtual patient, as examined in the previous Section 2.3. Furthermore, we performed a sensitivity analysis of twenty virtual patients in the $r_{recruit}[M] \times r_{recruit}[D]$ space to test if changes in these parameters can induce changes in the outcomes. Ten of these patients were randomly sampled from the 92 patients labeled as MC and ten from the 2658 patients labeled as NC (Figure 3B). For each sampled patient, we performed simulations using the possible combinations of minimum, central, and maximum values of each of the two parameters, defining

the 3x3 mesh. In each partition, we execute ten replicates and label the outcomes according to our classifier (details in Methods Section 4.9), resulting in a total of 90 simulations for each patient (Figure 5B, generated from 1800 new simulations).

In Figure 5B, the results show overall that the higher the recruitment rates, the greater the chance of the patient to have a better response from the immune system, in some cases being able to generate a significant control of tumor growth, for example, patients 1420, 2602, 3081, 4972, 5633, 1810, 2958, 5379, and 7642. The patients 3081 and 4972 outcomes show that the relationship between recruitment rate of DCs and SC label is not entirely linear. On the other hand, for some patients, even varying recruitment rates do not change the outcome, for example, patient 6040. This fact highlights the significance of integrating patient-specific treatment with mechanistic models for treatment efficacy prediction. Mechanistic models empower us to identify patient-specific characteristics and assess treatment alternatives.

3 Discussion

In this study, we have addressed four challenges associated with modeling pulmonary micrometastasis. These challenges include the following: (1) developing an *in silico* model able to represent immune system and micrometastasis interactions, (2) performing a substantial exploration of virtual patient trajectories, (3) detecting three distinct dynamical behaviors that characterize patient outcomes, and (4) identifying the main factors (parameters) that influence patient outcomes and analyzing the patient classification variability in a bi-dimensional parametric space.

We have met the first challenge by building a multiscale model to represent early micrometastasis growing in the lung tissue. We modeled the abnormal cancer cell proliferation generating a cascade of the effects that trigger the immune system response. The growth of the micrometastasis creates mechanical stress in healthy tissue, inducing pro-inflammatory signals. The signals attract and recruit more immune cells to the metastatic region creating cancer immunosurveillance. The model is able to generate different immune control responses including escape, and partial and complete elimination. Next, we used high performance computing for large scale exploration of ten model parameters, focusing on those that modulate the immune response. We used Latin Hypercube Sampling to select 10,000 virtual patients (each patient is characterized by a unique set of parameter values), and performed 10 replicates for each patient, generating a total of 100,000 simulations. Furthermore, based on thresholds for the cancer cell population after 10 days of the simulation, we classified the outcomes into three distinct behaviors of tumor control: significant control (SC), marginal control (MC), and no control (NC). Our model successfully represents two of the three Es of cancer immunoediting [33], namely elimination and escape, through the simulation of these scenarios. To incorporate the concept of equilibrium, we propose the integration of

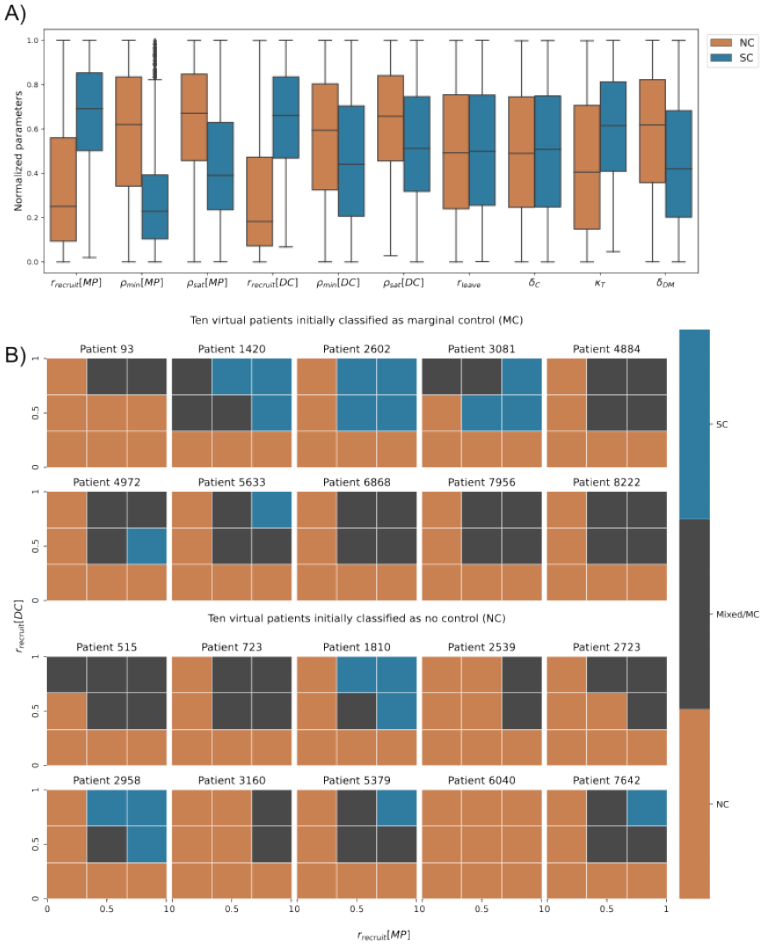

Fig. 5: A) The quartiles of normalized parameters according to each label. The box bounds define interquartile ranges (IQR). The line divides the IQR is the median, and the whiskers length is 1.5 times the interquartile range above the upper quartile and below the lower quartile. B) Analysis of 20 patients on normalized parametric space $r_{recruit}[M] \times r_{recruit}[D]$. The first two rows are 10 patients labeled as MC, and the bottom 2 rows are 10 patients labeled as NC. The mesh of parametric space is regular, and 3×3 for each patient, the color represents the label of mesh partition.

additional phenomena into the model, such as the healing process of damaged tissue and anti-inflammatory responses.

The dimensionality reduction of trajectories in spatial-temporal data supports the distinction of three separate behaviors. Furthermore, the reduced space reveals a distinct set of trajectories suggesting a unique subtype of no-control behavior (NC). We have demonstrated that the differences between

these two NC clusters are associated with immune activation. The reduced space of trajectory averages further underscores the significant uncertainty surrounding certain patient outcomes and effectively separates patients into either NC or SC labels. Moreover, this model reduction process may not represent the optimal approach for spatial-temporal trajectories generated by the model. Additional alternatives should be considered for evaluation in the near future [34–37].

After analyzing the parameter space of virtual patients, we have identified that the recruitment process of macrophages and dendritic cells has the most significant impact on tumor control by the immune system. Further exploration has suggested that altering the values of key parameters in certain patients can lead to changes in outcomes. While this variation in patient labeling is purely computational, in a scenario where the model is validated, this study can be interpreted and linked to potential therapeutic strategies for specific patients. However, we emphasize that our patient classifier is preliminary, and the development of a trajectory classifier that considers space/time variations and the stochasticity of the model is the subject of future investigations [10].

The model developed and implemented in this work focused on early metastatic events and thus only a subset of relevant immune interactions were considered. We are currently working on extending the model to include additional interactions such as the variability of tumor specific antigens, which can affect the capacity of the immune system to detect and kill metastatic cancer cells; this will be used to simulate the efficacy of antigen-specific treatments. Moreover, we plan to extend the model by adding more interactions and cell types, including cells and cytokines that drive an anti-inflammatory response. We will also extend the model to include cells of the innate immune response, such as NK cells, and also cells that are known to greatly affect the microenvironment and the immune response such as cancer-associated fibroblasts (CAFs).

The immediate future path for this work involves using the model to identify digital twin templates [9, 10]. However, during this work, we have encountered challenges along this path. The concept of creating patient templates fails for a considerable number of patients. The stochastic nature of immune-tumor interactions drives substantial uncertainty in the outcomes, even in an ideal scenario *in silico* where there are no measurement errors (perfect patient measurements). Additionally, we have recognized the necessity of obtaining more spatial and temporal information from untreated patients, a challenge often encountered in typical clinical data due to the urgency of initial treatment. To overcome these challenges, our plan is to integrate real-time molecular data and calibrate time-course data acquired from wet labs into a comprehensive agent-based model. This model can then be translated into a more general model to predict patient outcomes. With this novel approach, we can address two crucial aspects: A) Predicting the long-term trajectory of a patient by integrating multiple data sources, instead of relying solely on

initial time data, which can lead to excessively variable outcomes; B) Continuously assimilating subsequent patient data, which is essential for accurately discerning prognoses of patients. Consequently, maintaining an ongoing alignment with new patient measurements will empower the Cancer Patient Digital Twin (CPDT) to anticipate the future disease state and forecast its response to treatment [13, 38]. This transformative strategy equips clinicians to select the most suitable treatment for each patient, revolutionizing cancer care by enhancing patient outcomes while minimizing adverse effects.

4 Methods

4.1 Multiscale model construction

We developed the multiscale model using PhysiCell [39], an open-source framework that allows the development of multicellular models at various spatial/temporal scales. In this tool, the cells are represented by off-lattice agents with independent cell cycle progression, death processes, volume changes, motility, etc. Some of these dynamics may be defined by substrate availability in the environment, which is represented on PhysiCell using an open-source biological diffusion solver, BioFVM [40]. PhysiCell has been used in a wide variety of multicellular problems, such as virus therapy, cancer immunology, tissue mechanics, and drug screening, among others [14, 21, 41–43].

We have previously formulated a model that elucidates the immune response to the dissemination of SARS-CoV-2 in the lungs [21]. This simulator, designed to emulate SARS-CoV-2 interactions within lung tissue, incorporates a collection of mechanisms that are based on domain expertise in immunology, virology, and tissue biology. In the current study, we have adapted this model to investigate the emergence of micrometastases in the lung alveolar epithelium. The cellular components of the tumor microenvironment are formed by healthy epithelial cells from the lung, isolated cancer cells, and immune cells.

4.2 Micrometastasis and lung tissue interaction

We modeled lung epithelial cells in homeostasis in the absence of tumor cells. As the tumor grows, cancer cells create intense mechanical pressure on the tissue, causing extensive damage to the region, eventually leading to the death of healthy cells [44]. Similar to [42], we model elasto-plastic mechanics of lung cells following adhesion-repulsion mechanics [39, 45], and interactions associated with an adjacent extracellular matrix [46]. Specifically, cells are subjected to two forces i) elastic force that resists their displacement and ii) plastic reorganization where the cell adheres to another point of the extracellular matrix, generating a rearrangement process. Specifically, given that cell i with position x_i and radius R_i is attached to the extracellular matrix at $\vec{x}_{i,ECM}$ experiencing the two forces described above and considering $\vec{d_i} = \vec{x}_{i,ECM} - \vec{x}_i$ and

$\hat{d}_i = \|\vec{d}_i\| - R_i$ to be very small, we have

$$\frac{d}{dt}\vec{x}_i(t) = r_E \hat{d}_i \frac{\vec{d}_i}{\|\vec{d}_i\|} \tag{1}$$

$$\frac{d}{dt}\vec{x}_{i,ECM}(t) = -r_P \vec{d}_i \tag{2}$$

where r_E is a coefficient for the magnitude of an elastic restorative force and r_P is a relaxation rate. If $\|\vec{d}_i\| = 0$, the cell is in equilibrium, implying $d\vec{x}_i/dt = 0$ and $d\vec{x}_{i,ECM}/dt = 0$. This mechanical stress generates two effects on the dynamics between lung cells and cancer cells. The lung cell undergoes programmed cell death if it has a deformation induced by a cancer cell greater than a maximum deformation value ($\hat{d}_i > d_{max}$), providing cancer cells with additional space to proliferate. On the other hand, cancer cells suffer a decreased proliferation capacity due to the mechanical pressure sustained by the tumor microenvironment [42]. Based on a cell cycle model (flow cytometry separated) from PhysiCell [39], we define the probability of a cell transitioning from phase G0/G1 to phase S in a time interval Δt depending on the local compressive pressure, given by:

$$Prob(\text{transition to S} \mid \text{cell in G0/G1}) = \overline{r}_{01} \cdot \max\left\{0, \frac{p_{max} - p}{p_{max}}\right\} \Delta t \tag{3}$$

where p is the simple pressure (a dimensionless analog of mechanical pressure [39]), $\overline{r}_{01}$ is the rate in the ideal mechanical situation for a cell to start the proliferation process, and p_{max} is the simple pressure threshold at which cancer cells stop proliferating.

4.3 Motility and infiltration of immune cells

The immune cells migrate into the tumor microenvironment via chemotaxis, guided by the tumor necrosis factor and debris gradients secreted by macrophages and dying cells, respectively. Cell motility is defined by their speed, bias direction, and persistence time in the velocity direction (T_{per}) as described in [39]. Mathematically, we calculate the velocity by

$$\vec{v} = s_{mot} \frac{(1-b)\vec{\xi} + b\vec{b}}{\|(1-b)\vec{\xi} + b\vec{b}\|} \tag{4}$$

where s_{mot} is the speed of chemotaxis, b denotes the migration bias (with a range of 0 to 1), $\vec{\xi}$ is a random unit vector direction, and $\vec{b}$ is the migration bias direction. Velocity undergoes stochastic changes in accordance with a Poisson probability distribution with a rate $1/T_{per}$, as discussed in [39]. In this study, we consider $\vec{b}$ as a combination of tumor necrosis factor and debris gradients,

defined by the equation

$$\vec{b} = \frac{\sigma_{TNF}\nabla\rho_{TNF} + \sigma_{debris}\nabla\rho_{debris}}{\|\sigma_{TNF}\nabla\rho_{TNF} + \sigma_{debris}\nabla\rho_{debris}\|} \tag{5}$$

where σ_{TNF} and σ_{debris} are the sensitivity of chemotaxis along with the TNF or debris gradient, respectively. We assume that macrophages and activated dendritic cells (DCs) exhibit TNF and debris sensitive chemotaxis ($\sigma_{TNF} > 0$ and $\sigma_{debris} > 0$), while non-activated DCs and macrophages rely on the debris gradient for chemotaxis ($\sigma_{TNF} = 0$ and $\sigma_{debris} = 1$). CD8 + and CD4 + T cells migrate exclusively based on the contribution of the TNF gradient ($\sigma_{TNF} = 1$ and $\sigma_{debris} = 0$). Activation of DCs and macrophages reduces motility speed, while cells attached to other cells remain immobile.

Eventually, immune cells infiltrate the tumor microenvironment through the vasculature. In this model, we randomly sampled voxels within the computational domain and designated them as vascularized voxels. These vascularized voxels are assumed to represent approximately 8.8% of the total tissue volume [21, 47]. Immune cells recruited to the tumor microenvironment infiltrate the domain through these vascularized voxels. Macrophages and dendritic cells are recruited to the microenvironment due to high concentration of TNF, in a time interval Δt. Cells are recruited to the system according to the following equation:

$$\#\text{ of recruited cells} = r_{recruit}\int_{\Omega}\min\left(1, \max\left(0, \frac{\rho_{TNF} - \rho_{\min}}{\rho_{sat} - \rho_{\min}}\right)\right) dV\Delta t \tag{6}$$

where $r_{recruit}$ is the recruitment rate, $\rho_{\min}$ and ρ_{sat} are the minimum and maximum signal of TNF to recruit the cells, ρ_{TNF} is the TNF concentration in each partition of the domain Ω, and dV is the volume element of Ω. CD8+ and CD4+ T cells are recruited to the tissue based on the response of the lymph node after dendritic cell migration into the lymph node [21].

4.4 Phagocytosis and macrophage phases

Macrophages play a key role in the dynamics of cancer and the immune system, which we capture in this model [48]. They carry out the cleaning process of the tissue by the process of phagocytosis (engulfment) of dead cells, which triggers the production of pro-inflammatory signals (in our model, TNF release) [49]. Eventually, inactivated macrophages (M0), attracted by debris released by dying cells, phagocytize the dying cell following a conditional probabilistic event within a time interval Δt:

$$Prob(\text{engulf cell} \mid \text{nearby dying cells}) = r_{phag}\Delta t \tag{7}$$

where r_{phag} represents the phagocytosis rate of the macrophages. In this process, macrophages become activated (M1, classically activated, pro-inflammatory) and absorb the entire volume of the dying cell (Figure 6A).

M1 macrophages spend time ingesting the genetic material and are unable to phagocytose during this phase. Moreover, the M1 macrophages reach exhaustion by ingesting a large amount of cell debris, leading to the loss of their ability to phagocytose cells and increasing the likelihood of their own death [50] (Figure 6B). Activated macrophages in contact with CD4+ T cells become hyperactivated, enhancing their potential to phagocytose cells, including the possibility of engulfing live cancer cells [31] (Figure 6C). The CD8 + T cells, when in contact with M1 macrophages, inhibit the production of TNF by the macrophages [31] (Figure 6D).

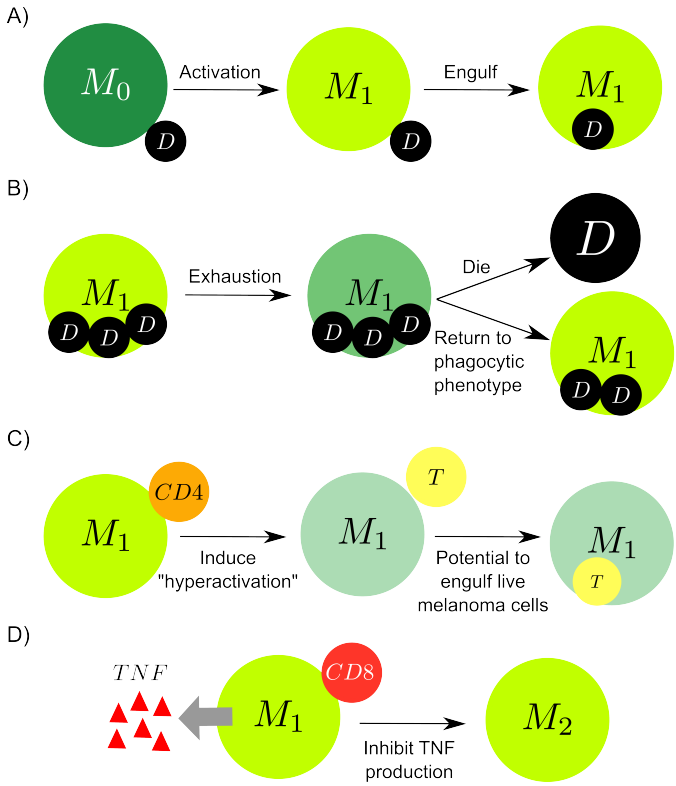

Fig. 6: Phagocytosis and macrophage phases. A) Macrophages in contact with dead cells are activated (M1 phase) and engulf the dead cell. B) M1 macrophages reach exhaustion by engulfing above a certain threshold of debris and stopping phagocytosing, increasing the chance of dying and decreasing the possibility of returning to the phagocytic phenotype. C) The hyperactivation of macrophages induced by CD4 T cells increases the potential to phagocytose live cancer cells. D) Activated macrophage in contact with a CD8 T-cell activates an anti-inflammatory response (M2) and stops secreting TNF.

4.5 Dendritic cell behavior

Initially, all dendritic cells (DCs) are non-activated in the tissue. The DCs attach cancer or a dying epithelial cell and become activated, representing the process of antigen presentation and inflammatory mediators produced by dead lung cells from mechanical stress. The attachment process is modeled in a time interval δt according to the following probabilistic event:

$$Prob(\text{DC attach} \mid \text{nearby dying cancer or lung cells}) = r_{attach} \cdot \Delta t \tag{8}$$

where the r_{attach} is the attachment rate of DCs. Activated DCs migrate to the lymph node according to a probabilistic event with a rate r_{leave}:

$$Prob(\text{DC migrates} \mid \text{activated DC}) = r_{leave} \cdot \Delta t \tag{9}$$

in any time interval $[t, t + \Delta t]$. The migration of DCs to the lymph node spend a time τ_{DC}. In the lymph node, these cells trigger the production, proliferation, and activation of specific T cells (details in Methods Section 4.6). In the tumoral microenvironment, activated DCs enhance the CD8+ T cell proliferation capability and attachment rate (details in Methods Section 4.7).

4.6 Lymph node submodel

We defined the lymph node model from the time course of cell populations. The number of dendritic cells that have migrated from the tumor microenvironment to the lymph node (D_M) upregulates the proliferation and activation of two types of helper T cells (T_{H1} and T_{H2}) and cytotoxic T cells (T_C). Consequently, these cytotoxic T cells activate the production of CD8 + cytotoxic (T_{Ct}) and CD4 + helper (T_{ht}) T cells, which will be trafficked to the tumor microenvironment [21, 51–53]. Altogether, the variation of the cell populations over time is given by:

$$\begin{aligned}
\frac{dD_M}{dt} &= k_D D(t - \tau_{DC}) - \delta_{DM} D_M \\
\frac{dT_{H1}}{dt} &= \sigma_{T_{H1}} \frac{T_{H1}}{(1 + T_{H2})^2} + \pi_{T_{H1}} \frac{D_M {T_{H1}}^2}{(1 + T_{H2})^2} - \delta_{T_{H1}} \frac{D_M {T_{H1}}^3}{(\beta + T_{H2})} - \mu_{T_H} T_{H1} \\
\frac{dT_{H2}}{dt} &= \sigma_{T_{H2}} \frac{T_{H2}}{(1 + T_{H2})} + \pi_{T_{H2}} \frac{(\rho + T_{H1})}{(1 + T_{H2})} \frac{D_M {T_{H2}}^2}{(1 + T_{H1} + T_{H2})} - \mu_{T_H} T_{H2} \\
\frac{dT_C}{dt} &= \pi_T \left(\frac{\phi_T - T_C}{\phi_T} \right) \frac{D_M T_C}{(\beta_{T1} + D_M)} - \delta_T \frac{D_M T_C}{(\beta_{T2} + D_M)} - \delta_C (T_C - \phi_C) \\
\frac{dT_{Ct}}{dt} &= \kappa_T T_C \\
\frac{dT_{ht}}{dt} &= \kappa_T (T_{H1} + T_{H2})
\end{aligned}$$

where k_D and τ_{DC} are the antigen presentation rate by dendritic cells and the time duration by the DCs to migrate into the LN, δ_{DM} is the natural decay of D_M, $\sigma_{T_{H1}}$ and $\sigma_{T_{H2}}$ denote the proliferation rates for type 1 and type 2 helper T cells, $\pi_{T_{H1}}$ is the activation rate of T_{H1} induced by D_M, $\delta_{T_{H1}}$ and β are rates and half maximum deactivation mediated by D_M for type 1 helper T cells, μ_{T_H} is the natural death rate of both cell types T_{H1} and T_{H2}, $\pi_{T_{H2}}$ represents the activation rate of type 2 helper T cells mediated by D_M, ρ is a weight conversion between T_{H1} and T_{H2} activation, π_T and δ_T are activation and deactivation of T_C mediated by D_M, ϕ_T is scaling factor between lymph node and tumor microenvironment, β_{T1} and β_{T2} are half maximum of activation and deactivation of T cells mediated by D_M, δ_C and ϕ_C are the natural decay and population threshold of T cells, κ_T is the recruitment rate of cytotoxic and helper T cells. The CD8 + and CD4 + T cells infiltrate the tumor microenvironment through the vasculature after a time τ_T, according to the cells amount $T_{Ct}(t - \tau_{TC})$ and $T_{ht}(t - \tau_{TC})$, respectively [21, 54, 55]

4.7 Cytotoxic CD8 + T cells killing cancer cells

Given a cancer cell in the neighborhood of CD8 + T cells, the probability of CD8 + T cells attaching to this cell in a time interval Δt is given by:

$$Prob(\text{CD8 + attach} \mid \text{nearby cancer cell}) = r_{attach} \cdot \Delta t \tag{10}$$

where r_{attach} is the rate of attachment [14, 21, 56]. Once the cells are attached, they can eventually detach, according to a probabilistic event that is a function of the duration of attachment T_{attach} as following:

$$Prob(\text{CD8 + detach} \mid \text{CD8 + attached}) = \frac{1}{T_{attach}} \cdot \Delta t \tag{11}$$

in any time interval $[t, t + \Delta t]$. Otherwise, cells remain bound until the time T_{attach}, when the cancer cell dies. Eventually, multiples CD8 + T cell can attach to the same cancer cell, increasing the likelihood of killing.

4.8 Generating virtual patients

Using HPC resources, we conducted a parameter exploration to generate virtual patients employing the multiscale model discussed earlier. Specifically, we identified ten parameters associated with the immune response, out of which six were linked to the recruitment of dendritic cells and macrophages to the microenvironment, three were associated with the production of T cells in the lymph node, and one was related to the migration of activated dendritic cells from the microenvironment to the lymph node. To sample the parameter space, we employed Latin hypercube sampling (LHS) and generated 10,000 parameter sets (see section 1 in supplementary material). To account for the stochastic nature of the model, we created ten replicates for each parameter

set, resulting in a total of 100,000 simulations. Due to the high dimensionality of the data, we compressed the data significantly to facilitate analysis (see section 2 in supplementary material).

The average wall time for each simulation, utilizing eight threads in shared memory programming with OpenMP, is approximately 11 minutes. If 100,000 simulations are processed in serial, it could take around 764 days. However, by utilizing HPC resources and executing 300 model simulations simultaneously (2400 threads), the entire dataset can be generated in 2.5 days.

4.9 Patient classifier

Based on the cancer cell population for each simulation, we created a classifier to distinguish different immune system responses. We define two threshold values (T_{SC} and T_{NC}) to categorize three types of responses from the cancer cell population on day 10 (P_M). The first type of response occurs when most of the cancer cells have been eliminated by the immune system at the end of the simulation (day 10), describing a significant control (SC) of tumor ($P_M < T_{SC}$) (elimination). The second type is when there is no control (NC) of the tumor, and the immune system is not strong enough to eliminate the tumor ($P_M > T_{NC}$) (escape). And the third type is an intermediate situation, in which there is marginal control (MC) of the tumor ($T_{SC} \leq P_M \leq T_{NC}$) (partial elimination).

Data availability

The dataset publicly is available in [57].

Code availability

The multiscale model and postprocessed dataset publicly are available in https://github.com/heberlr/micromets_lung.

References

[1] Chaffer, C.L., Weinberg, R.A.: A perspective on cancer cell metastasis. science **331**(6024), 1559–1564 (2011)

[2] Martin, T.A., Ye, L., Sanders, A.J., Lane, J., Jiang, W.G.: Cancer invasion and metastasis: molecular and cellular perspective. In: Madame Curie Bioscience Database [Internet]. Landes Bioscience, ??? (2013)

[3] Pantel, K., Cote, R.J., Fodstad, O.: Detection and clinical importance of micrometastatic disease. Journal of the National Cancer Institute **91**(13), 1113–1124 (1999)

[4] Riihimäki, M., Thomsen, H., Hemminki, A., Sundquist, K., Hemminki, K.: Comparison of survival of patients with metastases from known versus

unknown primaries: survival in metastatic cancer. BMC cancer **13**, 1–8 (2013)

[5] Janssen, L.M., Ramsay, E.E., Logsdon, C.D., Overwijk, W.W.: The immune system in cancer metastasis: friend or foe? Journal for immunotherapy of cancer **5**, 1–14 (2017)

[6] Chambers, A.F., Groom, A.C., MacDonald, I.C.: Dissemination and growth of cancer cells in metastatic sites. Nature Reviews Cancer **2**(8), 563–572 (2002)

[7] Mehlen, P., Puisieux, A.: Metastasis: a question of life or death. Nature reviews cancer **6**(6), 449–458 (2006)

[8] Dunn, G.P., Old, L.J., Schreiber, R.D.: The immunobiology of cancer immunosurveillance and immunoediting. Immunity **21**(2), 137–148 (2004)

[9] Hernandez-Boussard, T., Macklin, P., Greenspan, E.J., Gryshuk, A.L., Stahlberg, E., Syeda-Mahmood, T., Shmulevich, I.: Digital twins for predictive oncology will be a paradigm shift for precision cancer care. Nature medicine **27**(12), 2065–2066 (2021)

[10] Stahlberg, E.A., Abdel-Rahman, M., Aguilar, B., Asadpoure, A., Beckman, R.A., Borkon, L.L., Bryan, J.N., Cebulla, C.M., Chang, Y.H., Chatterjee, A., et al.: Exploring approaches for predictive cancer patient digital twins: Opportunities for collaboration and innovation. Frontiers in Digital Health **4** (2022)

[11] Fertig, E.J., Jaffee, E.M., Macklin, P., Stearns, V., Wang, C.: Forecasting cancer: from precision to predictive medicine. Med **2**(9), 1004–1010 (2021)

[12] Madhavan, S., Beckman, R.A., McCoy, M.D., Pishvaian, M.J., Brody, J.R., Macklin, P.: Envisioning the future of precision oncology trials. Nature Cancer **2**(1), 9–11 (2021)

[13] Laubenbacher, R., Niarakis, A., Helikar, T., An, G., Shapiro, B., Malik-Sheriff, R., Sego, T., Knapp, A., Macklin, P., Glazier, J.: Building digital twins of the human immune system: toward a roadmap. npj Digital Medicine **5**(1), 64 (2022)

[14] Ozik, J., Collier, N., Heiland, R., An, G., Macklin, P.: Learning-accelerated discovery of immune-tumour interactions. Molecular systems design & engineering **4**(4), 747–760 (2019)

[15] Norton, K.-A., Gong, C., Jamalian, S., Popel, A.S.: Multiscale agent-based and hybrid modeling of the tumor immune microenvironment. Processes

7(1), 37 (2019)

[16] Mahlbacher, G.E., Reihmer, K.C., Frieboes, H.B.: Mathematical modeling of tumor-immune cell interactions. Journal of Theoretical Biology **469**, 47–60 (2019)

[17] Wilkie, K.P., Hahnfeldt, P.: Mathematical models of immune-induced cancer dormancy and the emergence of immune evasion. Interface Focus **3**(4), 20130010 (2013)

[18] Radunskaya, A., Kim, R., Woods II, T.: Mathematical modeling of tumor immune interactions: a closer look at the role of a pd-l1 inhibitor in cancer immunotherapy. Spora: A Journal of Biomathematics **4**(1), 25–41 (2018)

[19] Pourhasanzade, F., Sabzpoushan, S., Alizadeh, A.M., Esmati, E.: An agent-based model of avascular tumor growth: Immune response tendency to prevent cancer development. Simulation **93**(8), 641–657 (2017)

[20] Cess, C.G., Finley, S.D.: Multi-scale modeling of macrophage—t cell interactions within the tumor microenvironment. PLoS computational biology **16**(12), 1008519 (2020)

[21] Getz, M., Wang, Y., An, G., Asthana, M., Becker, A., Cockrell, C., Collier, N., Craig, M., Davis, C., Faeder, J., et al.: Iterative community-driven development of a SARS-CoV-2 tissue simulator. Preprint at https://www.biorxiv.org/content/10.1101/2020.04.02.019075v4 (2021)

[22] Makaryan, S.Z., Cess, C.G., Finley, S.D.: Modeling immune cell behavior across scales in cancer. Wiley Interdisciplinary Reviews: Systems Biology and Medicine **12**(4), 1484 (2020)

[23] Rocha, H.L., Silva, J.V.d.O., Silva, R.S., Lima, E.A., Almeida, R.C.: Bayesian inference using gaussian process surrogates in cancer modeling. Computer Methods in Applied Mechanics and Engineering **399**, 115412 (2022)

[24] Warne, D.J., Prescott, T.P., Baker, R.E., Simpson, M.J.: Multifidelity multilevel monte carlo to accelerate approximate bayesian parameter inference for partially observed stochastic processes. Journal of Computational Physics **469**, 111543 (2022)

[25] Warne, D.J., Baker, R.E., Simpson, M.J.: Rapid bayesian inference for expensive stochastic models. Journal of Computational and Graphical Statistics **31**(2), 512–528 (2022)

[26] Lintusaari, J., Vuollekoski, H., Kangasraasio, A., Skytén, K., Jarvenpaa, M., Marttinen, P., Gutmann, M.U., Vehtari, A., Corander, J., Kaski, S.:

Elfi: Engine for likelihood-free inference. Journal of Machine Learning Research **19**(16), 1–7 (2018)

[27] Purkayastha, P., Jaiswal, M.K., Lele, T.P.: Molecular cancer cell responses to solid compressive stress and interstitial fluid pressure. Cytoskeleton **78**(6), 312–322 (2021)

[28] Takao, S., Taya, M., Chiew, C.: Mechanical stress-induced cell death in breast cancer cells. Biology open **8**(8), 043133 (2019)

[29] Nagata, R., Igaki, T.: Cell competition: Emerging mechanisms to eliminate neighbors. Development, growth & differentiation **60**(9), 522–530 (2018)

[30] Vincent, J.-P., Fletcher, A.G., Baena-Lopez, L.A.: Mechanisms and mechanics of cell competition in epithelia. Nature reviews Molecular cell biology **14**(9), 581–591 (2013)

[31] Yunna, C., Mengru, H., Lei, W., Weidong, C.: Macrophage m1/m2 polarization. European journal of pharmacology **877**, 173090 (2020)

[32] Lai, Y.-P., Jeng, C.-J., Chen, S.-C.: The roles of cd4+ t cells in tumor immunity. ISRN Immunology **2011**, 1–6 (2011)

[33] Dunn, G.P., Old, L.J., Schreiber, R.D.: The three es of cancer immunoediting. Annu. Rev. Immunol. **22**, 329–360 (2004)

[34] Isensee, J., Datseris, G., Parlitz, U.: Predicting spatio-temporal time series using dimension reduced local states. Journal of Nonlinear Science **30**(3), 713–735 (2020)

[35] Steadman, L., Griffiths, N., Jarvis, S., Bell, M., Helman, S., Wallbank, C.: k d-str: A method for spatio-temporal data reduction and modelling. ACM/IMS Transactions on Data Science **2**(3), 1–31 (2021)

[36] Chazal, F., Michel, B.: An introduction to topological data analysis: fundamental and practical aspects for data scientists. Frontiers in artificial intelligence **4**, 108 (2021)

[37] Gagliardi, P.A., Grädel, B., Jacques, M.-A., Hinderling, L., Ender, P., Cohen, A.R., Kastberger, G., Pertz, O., Dobrzyński, M.: Automatic detection of spatio-temporal signalling patterns in cell collectives. Preprint at https://www.biorxiv.org/content/early/2022/07/12/2022.07.12.499734 (2022)

[38] An, G.: Specialty grand challenge: What it will take to cross the valley of death: Translational systems biology,"true" precision medicine, medical

digital twins, artificial intelligence and in silico clinical trials. Frontiers in Systems Biology **2**, 901159 (2022)

[39] Ghaffarizadeh, A., Heiland, R., Friedman, S.H., Mumenthaler, S.M., Macklin, P.: Physicell: An open source physics-based cell simulator for 3-d multicellular systems. PLoS computational biology **14**(2), 1005991 (2018)

[40] Ghaffarizadeh, A., Friedman, S.H., Macklin, P.: Biofvm: an efficient, parallelized diffusive transport solver for 3-d biological simulations. Bioinformatics **32**(8), 1256–1258 (2016)

[41] Rocha, H.L., Godet, I., Kurtoglu, F., Metzcar, J., Konstantinopoulos, K., Bhoyar, S., Gilkes, D.M., Macklin, P.: A persistent invasive phenotype in post-hypoxic tumor cells is revealed by fate mapping and computational modeling. Iscience **24**(9), 102935 (2021)

[42] Wang, Y., Brodin, E., Nishii, K., Frieboes, H.B., Mumenthaler, S.M., Sparks, J.L., Macklin, P.: Impact of tumor-parenchyma biomechanics on liver metastatic progression: a multi-model approach. Scientific reports **11**(1), 1–20 (2021)

[43] Craig, M., Jenner, A.L., Namgung, B., Lee, L.P., Goldman, A.: Engineering in medicine to address the challenge of cancer drug resistance: from micro-and nanotechnologies to computational and mathematical modeling. Chemical Reviews **121**(6), 3352–3389 (2020)

[44] Brás, M.M., Radmacher, M., Sousa, S.R., Granja, P.L.: Melanoma in the eyes of mechanobiology. Frontiers in Cell and Developmental Biology **8**, 54 (2020)

[45] Macklin, P., Edgerton, M.E., Thompson, A.M., Cristini, V.: Patient-calibrated agent-based modelling of ductal carcinoma in situ (DCIS): From microscopic measurements to macroscopic predictions of clinical progression. Journal of Theoretical Biology **301**, 122–140 (2012). https://doi.org/10.1016/j.jtbi.2012.02.002

[46] D'Antonio, G., Macklin, P., Preziosi, L.: An agent-based model for elasto-plastic mechanical interactions between cells, basement membrane and extracellular matrix. Mathematical biosciences and engineering: MBE **10**(1), 75 (2013)

[47] Pienaar, E., Cilfone, N.A., Lin, P.L., Dartois, V., Mattila, J.T., Butler, J.R., Flynn, J.L., Kirschner, D.E., Linderman, J.J.: A computational tool integrating host immunity with antibiotic dynamics to study tuberculosis treatment. Journal of theoretical biology **367**, 166–179 (2015)

[48] Atri, C., Guerfali, F.Z., Laouini, D.: Role of human macrophage polarization in inflammation during infectious diseases. International journal of molecular sciences **19**(6), 1801 (2018)

[49] Zhang, J.-M., An, J.: Cytokines, inflammation and pain. International anesthesiology clinics **45**(2), 27 (2007)

[50] Zent, C.S., Elliott, M.R.: Maxed out macs: physiologic cell clearance as a function of macrophage phagocytic capacity. The FEBS journal **284**(7), 1021–1039 (2017)

[51] Tal, O., Lim, H.Y., Gurevich, I., Milo, I., Shipony, Z., Ng, L.G., Angeli, V., Shakhar, G.: Dc mobilization from the skin requires docking to immobilized ccl21 on lymphatic endothelium and intralymphatic crawling. Journal of Experimental Medicine **208**(10), 2141–2153 (2011)

[52] Okada, T., Cyster, J.G.: Cc chemokine receptor 7 contributes to gi-dependent t cell motility in the lymph node. The Journal of Immunology **178**(5), 2973–2978 (2007)

[53] Toapanta, F.R., Ross, T.M.: Impaired immune responses in the lungs of aged mice following influenza infection. Respiratory research **10**, 1–19 (2009)

[54] Yates, A., Bergmann, C., Van Hemmen, J.L., Stark, J., Callard, R.: Cytokine-modulated regulation of helper t cell populations. Journal of theoretical biology **206**(4), 539–560 (2000)

[55] Lee, H.Y., Topham, D.J., Park, S.Y., Hollenbaugh, J., Treanor, J., Mosmann, T.R., Jin, X., Ward, B.M., Miao, H., Holden-Wiltse, J., *et al.*: Simulation and prediction of the adaptive immune response to influenza a virus infection. Journal of virology **83**(14), 7151–7165 (2009)

[56] Ozik, J., Collier, N., Wozniak, J.M., Macal, C., Cockrell, C., Friedman, S.H., Ghaffarizadeh, A., Heiland, R., An, G., Macklin, P.: High-throughput cancer hypothesis testing with an integrated physicell-emews workflow. BMC bioinformatics **19**(18), 81–97 (2018)

[57] Rocha, H.L.: Pulmonary micrometastasis dataset (2023). https://doi.org/10.6084/m9.figshare.21514116.v3

Acknowledgments

We thank Jeffrey Bryan (University of Missouri) and Tina Hernandez-Boussard (Stanford University) for their helpful discussions. The computational simulations of this research were performed on high-throughput computing clusters at Indiana University, which were supported in part by Lilly

Endowment, Inc., through its support for the Indiana University Pervasive Technology Institute.