Biological Data 2026, 1(1), 20-43.
https://doi.org/????

Review Article

Integrated Computational Frameworks for Multiscale Modeling of Cancer Prevention

Elahe Ghiyabi[1*], Houshang Shaker Nigjeh[2], Sajjad Mortazavi[1]

[1] Department of Biomedical Engineering, Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran
[2] Department of Biomedical Engineering, Faculty of Electrical and Computer Engineering, University of Zanjan, Zanjan, Iran

***** Corresponding Author: Elahe Ghiyabi
Email: Elaheghiyabi1996@gmail.com

CITATION

Ghiyabi, E., Shaker Nigjeh, H., & Mortazavi, S. (2026). Integrated computational frameworks for multiscale modeling of cancer prevention. Biological Data Journal, 1(1), 20–43.
https://doi.org/xxxxxxxxx.

ARTICLE INFO

Received: March 13, 2026
Revised: March 31, 2026
Accepted: April 17, 2026
Available online: April 21, 2026

COPYRIGHT

© The author(s) retain the copyright.
Publisher: University of Kurdistan Press.
This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Abstract: Cancer develops through interconnected molecular, cellular, and tissue-level processes that span multiple spatial and temporal scales. Although multiscale experimental data have expanded opportunities for quantitative analysis, challenges such as limited data availability, biological heterogeneity, and the need for interpretable predictions still restrict progress. A broad range of mathematical and computational approaches has been developed to address these issues. Population-level models based on ordinary differential equations remain central to studies of tumor growth, immune dynamics, and treatment effects, while stochastic versions capture uncertainty in early carcinogenic events and variability in therapeutic outcomes. Spatially resolved methods, including partial differential equations, cellular automata, and multiphase models, incorporate microenvironmental heterogeneity, chemical gradients, and mechanical influences that shape tumor behavior. At the cellular scale, agent-based models represent heterogeneous populations and local interactions, and intracellular network models together with evolutionary game-theoretic approaches clarify regulatory mechanisms and competitive dynamics. Recent efforts increasingly focus on hybrid and multiscale frameworks that integrate these strategies to connect molecular events with emergent tissue-level patterns. This review examines how cancer develops across interconnected molecular, cellular, and tissue-level processes spanning multiple spatial and temporal scales.

Keywords: Cancer Prevention, Agent-Based Model (ABM), Multi-scale Modeling, Computational Oncology, Digital Twins

1. Introduction

Cancer develops through a complex interplay of molecular, cellular, and tissue-level processes that occur across numerous spatial and temporal scales. Advances in experimental methodologies have resulted in an unprecedented amount of multiscale data, ranging from genetic and epigenomic information to cellular imaging and population-level clinical records, opening up new avenues for oncology predictive modeling. Within this context, mathematical oncology has emerged as a field dedicated to connecting biology, medicine, data science, and mathematical modeling. This field aims to better understand tumor initiation, development, and response to therapies. Despite significant advances, fundamental hurdles remain, including data scarcity, inter-patient heterogeneity, nonlinearity of biological systems, and the requirement for reliable and interpretable models (Azizi, 2024; Metzcar et al., 2024; Yankeelov et al., 2016). There are two major modeling paradigms that dominate the field. First, mechanistic (knowledge-driven) models, such as systems of ordinary differential equations (ODEs), use well-established biological and physicochemical principles to characterize system dynamics. These models provide transparency, interpretability, and the capability to simulate hypothetical scenarios or unknown conditions using parameterized governing equations. They are widely used to study tumor-immune interactions, cell proliferation dynamics, signaling pathways, and chemopreventive responses (Li & Lei, 2025; Altrock et al., 2015). Mechanistic techniques have lately been extended beyond pure ODEs to include spatial heterogeneity, microenvironmental elements, and biomechanical phenomena through hybrid frameworks that include ODEs with PDEs or moment-closure reductions of PDEs (Villa et al., 2025; Mirzaei & Shahriyari, 2024). Such models can capture both temporal dynamics and, where necessary, geographic or phenotypic variability, increasing their biological realism while remaining computationally tractable. Second, data-driven methods, such as traditional machine learning and deep learning approaches, use statistical patterns to extract correlations from high-dimensional omics and imaging data. These models excel in predictive accuracy, but they frequently suffer with mechanical interpretability and generalizability outside of the training data distribution (Ghassemi et al., 2020; Zou et al., 2019; Camacho et al., 2018). Recognizing the limitations of either paradigm, recent research increasingly shifts toward mechanistic learning, a hybrid approach that combines mechanistic models with machine-learning-based inference to capitalize on the strengths of both. Mechanistic models provide structure and interpretability, but data-driven methods help with parameter estimates, identifying hidden biological systems, and correcting model flaws. This collaboration is a promising step toward more robust and clinically useful forecasts (Agnello et al., 2024; Brunton & Kutz, 2022; Rackauckas et al., 2020; Rubanova et al., 2019). Despite these advances, a critical gap remains in the systematic application of computational modeling to cancer prevention and early-stage disease. The majority of existing models are primarily designed to study treatment response, therapeutic optimization, and resistance mechanisms in established tumors. In contrast, prevention-oriented processes—such as early mutational events, immune surveillance, microenvironmental adaptation, and long-term exposure to environmental or lifestyle-related risk factors—are often underrepresented or treated indirectly. These processes operate across extended temporal horizons and multiple biological scales, making them difficult to capture using single-paradigm or late-stage–focused modeling approaches. As a result, there is a growing need for integrative computational frameworks that are explicitly tailored to cancer prevention, capable of linking molecular mechanisms, cellular dynamics, spatial microenvironments, and population-level outcomes in a coherent and interpretable manner.From a prevention perspective, computational models play a uniquely important role by enabling the systematic exploration of long-term and low-probability biological processes that are difficult, costly, or ethically infeasible to investigate experimentally. Early carcinogenesis is often driven by rare stochastic events, subtle phenotypic

transitions, and cumulative environmental or lifestyle-related exposures, all of which may unfold over decades before clinical manifestation. Mathematical and computational frameworks provide a controlled in silico environment to investigate how such early perturbations interact with immune surveillance, tissue architecture, and evolutionary selection pressures to either suppress malignant transformation or promote disease initiation. Recent reviews emphasize that without explicitly incorporating prevention-oriented endpoints and early-stage biological constraints, computational oncology risks remaining disproportionately focused on treatment optimization rather than disease avoidance. Consequently, prevention-focused multiscale modeling is increasingly recognized as a necessary methodological extension for translating systems-level biological insight into actionable strategies for risk stratification, early intervention, and precision cancer prevention. Hybrid and multiscale modeling approaches are especially important for cancer prevention, as early-stage cellular dynamics, precancerous changes, immune surveillance, and environmental exposures interact at several biological scales (Altrock et al., 2015; Anderson & Quaranta, 2008). Traditional ODE models provide a coarse-grained perspective of tumor or precancerous progression, but they fail to account for spatial heterogeneity and local cellular interactions. Agent-Based Models (ABMs) may mimic individual cell behavior and interactions with the microenvironment, capturing emergent phenomena such as clonal rivalry, immune cell infiltration, and microarchitectural evolution (Metzcar et al., 2019; Rejniak & Anderson, 2011). However, ABMs are computationally expensive and may lack specific mechanistic foundations (Rejniak & Anderson, 2011). These constraints highlight the need for a unified framework that connects mechanistic ODE models to Agent-Based Models (ABMs), allowing for simultaneous depiction of global system dynamics and localized cell–cell behaviors (Nardini et al., 2021; Rikard et al., 2019). As underlined in recent mathematical oncology reviews, integrative and multiscale modeling frameworks can increase biological realism, interpretability, and provide therapeutically relevant insights for cancer prevention and early-stage tumor progression (Norton et al., 2019; Deisboeck et al., 2011). This article is a review that synthesizes and critically discusses multiscale mathematical and computational modeling strategies relevant to cancer prevention and early-stage progression. Rather than proposing a new modeling framework, we organize existing approaches across population-level, spatial, cellular, and evolutionary scales, and highlight how hybrid and AI-augmented methods can connect mechanistic insight with data-driven inference. We further map these modeling paradigms to commonly used computational platforms, emphasizing practical considerations for prevention-oriented research and future digital-twin development. Table 1 provides a comparative overview of these key methodologies. We will start with fundamental population-level models, go on to spatially aware models, investigate sophisticated models of cellular interaction, and finally examine the critical difficulties of model integration, validation, and clinical translation. The aim of this article is to provide a comprehensive review of mathematical and computational modeling approaches relevant to cancer prevention and early-stage progression. Rather than introducing a new modeling framework, this study systematically synthesizes existing methodologies from a prevention-oriented perspective. Specifically, the contributions of this review are threefold: (i) it organizes population-level, spatial, cellular, and evolutionary models within a unified multiscale framework focused on cancer prevention; (ii) it critically examines hybrid and AI-augmented modeling strategies as emerging tools for integrating mechanistic insight with data-driven inference; and (iii) it connects modeling approaches to practical computational platforms, highlighting their strengths, limitations, and suitability for prevention-focused research and future digital twin applications.

Table 1. Summary and comparison of mathematical modeling approaches in oncology.

Modeling Approach	Brief Description	Strengths	Limitations	Typical Applications in Oncology
ODEs (Ordinary Differential Equations) (Altrock et al., 2015; Robertson-Tessi et al., 2012)	Models temporal changes in homogeneous, well-mixed populations.	Computationally fast; excellent for systemic PK/PD modeling; easy to analyze.	Lacks spatial detail; cannot capture cell–cell variability (mean-field).	Tumor growth kinetics, treatment response modeling, signaling cascades, PK/PD modeling.
PDEs (Partial Differential Equations) (Azizi, 2025; Byrne, 2010)	Models temporal and spatial changes in cell/molecule densities.	Captures diffusion gradients and spatial heterogeneity.	Computationally expensive; treats populations as a continuum.	Tumor invasion, angiogenesis, drug diffusion, oxygen/nutrient transport.
Cellular Automata (CA) (Gerlee & Anderson, 2007; Kansal et al., 2000)	Discrete grid-based model with rule-based state updates.	Computationally efficient; good for local interactions.	Limited biological realism; rule sets can be rigid.	Morphological tumor growth, necrotic core formation, local interactions.
Agent-Based Models (ABM) (Metzcar et al., 2019; Poleszczuk et al., 2016)	Bottom-up simulation of autonomous agents with unique rules.	High realism; captures heterogeneity and emergent behavior.	Computationally expensive; difficult to validate.	Immune–tumor interactions, heterogeneity-driven resistance, multiscale modeling.
Network Models (Barabasi & Oltvai, 2004; Huang et al., 2009)	Graph-based representation of interactions among biological entities.	Ideal for pathway analysis; identifies regulatory hubs.	Often non-spatial; may miss environmental feedback.	Gene regulatory networks, signaling pathways, drug-target analysis.
Evolutionary Game Theory (EGT) (Gatenby & Vincent, 2003a; Tomlinson & Bodmer, 1997)	Models strategic interactions between different cell populations.	Captures competition/cooperation; predicts resistance.	Abstract; payoff matrices hard to define.	Treatment resistance modeling, adaptive therapy design, clonal evolution.

2. Background and Literature review

Population-level models are the historical cornerstone of Computational oncology models provide tools to simulate profound insights into the temporal dynamics of entire cell populations while reducing spatial complexity. These models regard tissue as a single, well-mixed compartment, with every cell interacting with equal likelihood with every other cell. This abstraction allows for more concentrated research into systemic processes such as the rate at which precancerous lesions grow, the optimization of chemopreventive agent dosing schedules, and the long-term impact of immune-boosting therapy at the population level. While spatial heterogeneity is neglected, these models are nonetheless useful for situations in which systemic dynamics, rather than local architecture, determine results.

2.1. Deterministic ODEs and Pharmacodynamic Modeling

Ordinary Differential Equations (ODEs) constitute the historical foundation of mathematical oncology. Using a mean-field approximation, they treat cell populations as continuous, well-mixed variables, ignoring individual cell locations. This enables researchers to simulate classical tumor–immune–drug interactions, as described in widely used mechanistic models of cancer growth and pharmacodynamics, including interactions between tumor cells (T), effector immune cells (E), and a chemopreventive agent (C):

$$\frac{dT}{dt} = rT\left(1 - \frac{T}{K}\right) - \alpha ET - K_{drug}CT \quad (1)$$

$$\frac{dE}{dt} = s + \frac{gET}{h + T} - dE \quad (2)$$

$$\frac{dC}{dt} = D(t) - \delta C \quad (3)$$

- r: intrinsic tumor growth rate
- K: environmental carrying capacity, i.e., the maximum tumor size sustainable by local resources
- α: immune-mediated tumor cell killing efficacy
- K_drug: drug efficacy parameter
- s, g, h, d: parameters controlling immune population dynamics
- D(t), δ: drug dosing function and clearance rate

These models are commonly used in pharmacokinetics/pharmacodynamics (PK/PD) research to simulate treatment regimens, optimize drug potency, and reduce toxicity. For example, altering D(t) or K_drug in silico can help uncover effective dosage regimens prior to preclinical studies. However, ODE models cannot account for spatial heterogeneity or local cell density effects, which may be important in solid tumors or tissues with organized microenvironments. A schematic representation of these interactions, adapted from established tumor–immune–drug modeling frameworks, is provided in Figure 1, where each regulatory term in equations (1)–(3) is mapped onto the corresponding biological process, including tumor proliferation, immune-mediated killing, drug-induced suppression, immune activation, and drug pharmacokinetics.

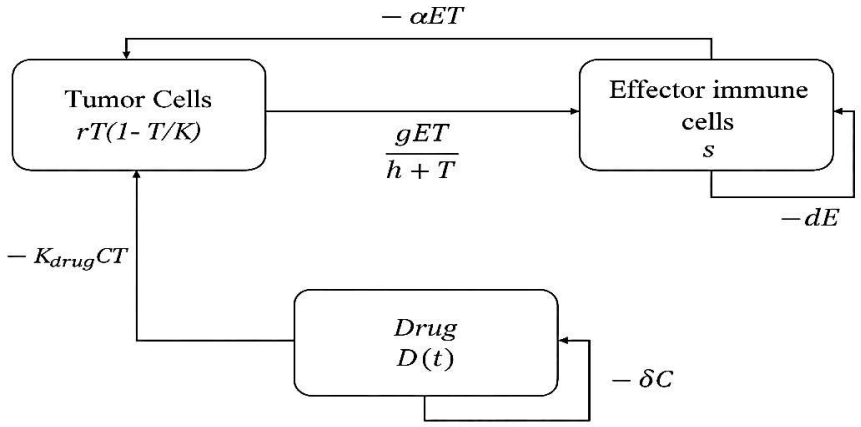


Figure 1. Conceptual diagram illustrating the tumor–immune–drug interactions captured by the ODE system in equations (1)-(3). Tumor cells (T) undergo logistic growth and are suppressed through immune-mediated cytotoxicity and drug-induced effects, while effector immune cells (E) are activated in response to tumor presence and subject to natural decay.

The chemopreventive agent (C) follows standard pharmacokinetic dynamics with external dosing and clearance. The schematic is adapted from classical tumor–immune and pharmacodynamic modeling frameworks reported in the literature (Jarrett et al., 2018; Robertson-Tessi et al., 2012).

2.2. Incorporating Chance: Stochastic Models

Cancer initiation is fundamentally stochastic: the fate of a small number of transformed cells is determined not only by deterministic biological processes but also by probabilistic events that can lead to very different outcomes even under identical conditions. Deterministic ODE models, which describe average population dynamics, cannot capture these fluctuations, especially when cell numbers are low. Stochastic models typically capture two complementary sources of randomness. The first is demographic stochasticity, which arises from inherently random birth–death events at the single-cell level and becomes particularly influential when cell numbers are small; under such conditions, chance alone can determine whether a nascent premalignant clone expands or goes extinct. The second is environmental stochasticity, which reflects fluctuations in external or system-level factors—such as transient changes in resource availability, local microenvironmental conditions, or variable effective treatment intensity—that modulate population behavior even when cell numbers are large. Together, these stochastic components provide a more realistic representation of early carcinogenic dynamics than purely deterministic mean-field models. A simple birth-death process, modeling demographic stochasticity, is described by the master equation:

$$\frac{dP_n(t)}{dt} = \lambda (n-1) P_{n-1}(t) + \mu (n+1) P_{n+1}(t) - (\lambda + \mu) n P_n(t) \quad (4)$$

The master equation provides a probabilistic description of cell population dynamics. Here, $P_n(t)$ is the probability of having n cells at time t, and λ and μ are the per-cell birth and death rates. The terms in the equation account for gains or losses in the population due to individual cell divisions and deaths. using this framework, one can calculate $P_0(t)$, the probability that a small precancerous clone goes extinct, which is a key metric for early preventive strategies. Recent studies use stochastic models to explore therapy resistance, immune evasion, and variability in treatment response, highlighting scenarios where random events dominate population dynamics. Figure 2 illustrates the

stochastic birth–death process underlying early tumor initiation, highlighting how random cell division and death events influence the extinction or expansion of small premalignant clones, as described by the master equation (Eq. 4), where each cell may either divide or die with probabilities determined by the per-cell birth rate λ and death rate μ. This schematic highlights how demographic stochasticity can result in different outcomes for a small population of transformed cells, emphasizing that even under identical conditions, some clones expand while others go extinct.

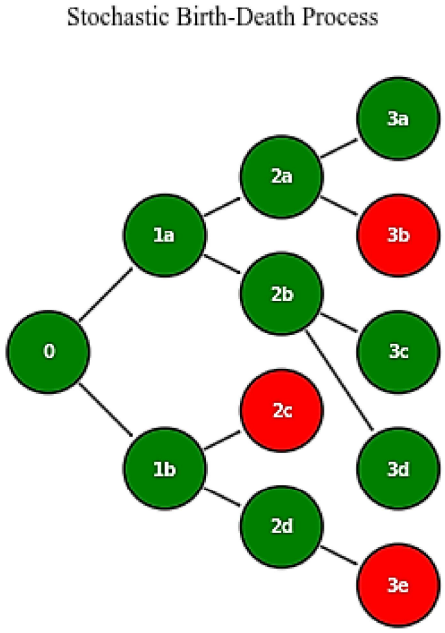


Figure 2. Stochastic birth–death process of early tumor initiation.
Circles represent individual cells (green: alive, red: dead). Arrows indicate possible cell divisions.
This schematic illustrates how the randomness in cell birth and death (demographic stochasticity) leads to different outcomes, with some clones expanding while others go extinct.

2.3. Towards Advanced and Hybrid Approaches

Multiscale and hybrid models: Combining deterministic ODEs with stochastic elements or Agent-Based Models (ABMs) allows researchers to integrate systemic population dynamics with spatial heterogeneity at the cellular level (Gatenby & Vincent, 2003a; Metzcar et al., 2019).
Clinical relevance: These models can simulate tumor-immune-drug interactions more realistically, aiding in the design of personalized therapy schedules (Azizi, 2025).
Future directions: Improved modeling frameworks may include spatially resolved PK/PD, microenvironment feedback, and cellular heterogeneity, bridging the gap between theoretical predictions and clinical outcomes (Rejniak & Anderson, 2011; Byrne, 2010; Anderson & Quaranta, 2008).
In the context of cancer prevention, population-level and hybrid models have also been used to study early carcinogenic dynamics rather than established disease. Such models include stochastic birth–death formulations that estimate the extinction probability of premalignant clones, immune surveillance models that quantify elimination versus escape during early transformation, and long-term exposure models that link cumulative carcinogen dose to cancer initiation risk. These prevention-oriented frameworks are particularly valuable for identifying critical thresholds at which early lesions either regress or progress, providing quantitative insight into the timing and efficacy of

preventive interventions. By focusing on initiation and early progression rather than therapeutic response, these models offer a complementary perspective that aligns more closely with prevention and risk reduction strategies.

3. Spatially-Aware and Microenvironmental Models

To achieve a higher level of biological realism, it is necessary to move beyond the population averages of non-spatial models and explicitly incorporate the dimension of space. Cancer does not develop in a vacuum; it arises and evolves within a complex, dynamic ecosystem known as the Tumor Microenvironment (TME). This environment, comprising the extracellular matrix, blood vessels, fibroblasts, various immune cells, and signaling molecules, is not a passive backdrop but an active participant that profoundly influences cell behavior through a web of spatial interactions. Crucially, the TME is characterized by spatial gradients of essential molecules. Cells near a blood vessel are exposed to high levels of oxygen and nutrients, promoting proliferation, while distant cells may experience hypoxia, triggering metabolic shifts (such as the Warburg effect) or even cell death. Similarly, gradients of signaling molecules released by other cells can direct cell migration (chemotaxis) and orchestrate collective behavior. Furthermore, direct physical phenomena, such as contact inhibition between neighboring cells and the mechanical barriers presented by the ECM, are fundamentally spatial processes that dictate tissue architecture and the potential for invasion. Spatially-aware models are therefore designed to capture these dynamics, either by treating cell populations as continuous densities that vary in space or by simulating discrete cells on a computational grid (Rejniak & Anderson, 2011; Byrne, 2010; Anderson & Quaranta, 2008). The following sections explore these distinct approaches.

3.1. Continuum Approaches: Partial Differential Equations (PDEs)

PDEs model the spatiotemporal evolution of cell densities and chemical concentrations. Reaction-diffusion equations are a cornerstone for modeling the diffusion of molecules like growth factors or drugs through tissue:

$$\frac{\mathrm{dC}}{\mathrm{dt}} = \mathrm{D}\nabla^2\mathrm{C} + \mathrm{f}(\mathrm{C},\mathrm{x}) \quad (5)$$

In this equation, the term $D\nabla^2 C$ represents the process of diffusion. The second term, *f(C,x)*, represents all local reactions, including production, decay, or consumption. More advanced models can also incorporate terms for advection, representing active transport of cells in a fluid flow. This framework is perfect for modeling the tumor microenvironment (Jarrett et al., 2018; Ramis-Conde et al., 2008).

3.2. Discrete Grid-Based Models: Cellular Automata (CA)

Cellular Automata (CA) offer a computationally efficient, discrete alternative where a grid site's state is determined by its neighbors (Wolfram & Gad-el-Hak, 2003). While simpler than ABMs, they can effectively model emergent spatial phenomena such as the formation of necrotic cores in tumors, the development of complex vascular networks, or the spatial segregation of drug-sensitive and drug-resistant cell populations, providing key insights into the architectural dynamics of precancerous tissues (Kansal et al., 2000).

3.3. Multiphase and Porous-Media Models

Tumor tissue functions as a multiphase medium with interacting solids (cells, ECM), liquids (interstitial fluid), and dissolved chemical species. Multiphase and porous-media models approach these components as overlapping continua, each with its own mass- and momentum-balance equations. This formulation enables researchers to model pressure-driven fluid movement, mechanical compression, matrix degradation, and nutrient redistribution within dense tumors. Importantly, the porous structure of the ECM results in spatially variable permeability, which has a strong influence on drug penetration and the production of invasive protrusions. Such models help to bridge the gap between tiny cell-scale mechanics and macroscopic tumor progression patterns.

3.4. Mechanobiological and ECM-Driven Spatial Models

Mechanical forces, in addition to chemical gradients, play an important role in tumor behavior. Mechanobiological models account for regional differences in ECM stiffness, solid stress, and compressive forces that occur when the tumor grows. These stressors can cause blood vessels to collapse, limit perfusion, and produce hypoxic niches, all of which enhance treatment resistance. Models that combine PDE-based diffusion and elasticity or viscoelasticity equations describe how mechanical cues influence migration, epithelial-mesenchymal transition (EMT), and collective invasion. Incorporating ECM remodeling, particularly collagen reorientation and fiber alignment, improves spatial models' capacity to recreate invasive fronts seen in aggressive malignancies.

3.5. Hybrid Multiscale Models (Coupling PDEs, CA, and ABMs)

Hybrid multiscale models combine several spatial representations—such as PDEs for chemical fields, CA or ABMs for individual cells, and ODEs for intracellular pathways—to form a coherent framework (Norton et al., 2019). This technique allows for the simultaneous investigation of molecular, cellular, and tissue dynamics (Hickey et al., 2023). For example, ODEs can be used to describe intracellular signaling networks that drive proliferation or apoptosis, whereas PDE dynamics govern the surrounding oxygen field and CA rules update cell locations (Ruiz-Martinez et al., 2022). These hybrid systems are capable of reproducing complicated phenomena such as localized clonal growth, adaptive therapy responses, and drug resistance evolution over time (Gonçalves & García-Aznar, 2023; Gong et al., 2017). Their adaptability makes them one of the most effective instruments in current mathematical oncology (Norton et al., 2019).

3.6. Challenges and Future Directions in Spatial Modeling

Spatial models, although powerful, still face several difficulties. Estimating parameters that change from one region to another, including diffusion rates, the permeability of the extracellular matrix, or the mechanical stiffness of tissues, requires very detailed experimental data that are often not available. These models also demand heavy computation, especially when they involve three dimensional simulations or hybrid structures with many interacting agents. Another challenge comes from problems of model identifiability and parameter sensitivity, which can make interpretation and prediction less reliable. Future progress will depend on using patient specific information obtained from spatial transcriptomics, high resolution imaging, and biomechanical measurements. Improvements in GPU based computation and simplified model designs may eventually make it possible to run near real time simulations that support more personalized cancer treatment.

4. Advanced Models of Cellular Interaction and Strategy

This class of models represents a fundamental shift in perspective from the "top-down," population-level views of previous sections to a "bottom-up" approach focused on simulating the actions and interactions of individuals. This individual-based methodology is essential for capturing one of the central challenges in oncology: intra-tumor heterogeneity, as it allows for the assignment of unique properties and behavioral rules to each simulated cell or "agent." To achieve this high level of biological detail, advanced models delve into two key areas explored in this section: the complex intracellular networks that govern an agent's decisions, and the strategic, game-like interactions that emerge when diverse agents compete for resources. These methods provide unparalleled insight into the mechanisms driving cancer progression at the cellular level (Sancho-Araiz et al., 2021; Clarke et al., 2019; Zangooei & Habibi, 2017).

4.1. The Individual at the Center: Agent-Based Models (ABMs)

Agent-Based Models (ABMs) represent a paradigm shift from describing population-level phenomena with equations to simulating the individual actions of a large number of autonomous "agents" (Bora & Emek, 2019). This bottom-up approach is uniquely suited to capturing the cellular and intra-tumor heterogeneity that defines cancer biology. An agent is a discrete software object defined by its internal state and its behavioral rules as illustrated in Figure 3.

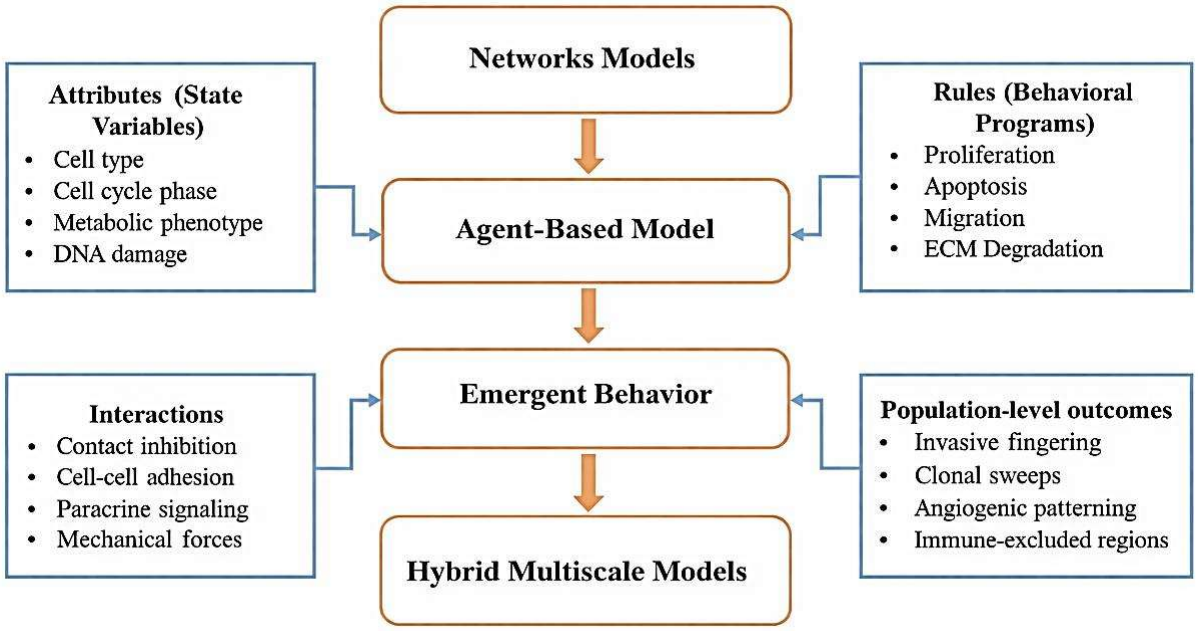


Figure 3. Framework of agent-based modeling in multiscale systems.

Key components of a cell agent in a cancer model are defined by their state attributes, behavioral rules, and interaction protocols. Attributes represent the internal state, such as cell type (e.g., epithelial, fibroblast, T-cell), cell cycle phase, metabolic status, and DNA damage levels. These are governed by rules that dictate actions like proliferation, apoptosis, or migration based on local environmental cues. Furthermore, interaction rules manage how agents connect with their surroundings through cell-cell adhesion, contact inhibition, or paracrine signaling. The true power of ABMs lies in revealing emergent phenomena, where complex, tissue-level behaviors like tumor

invasion or angiogenesis arise from thousands of simple, local interactions (Metzcar et al., 2019). A growing number of specialized software platforms, such as PhysiCell, Chaste, and CompuCell3D, provide powerful frameworks for building these complex cancer models.

4.2. Modeling Internal Machinery: Network Models

The behavior of each cell agent is ultimately controlled by its internal molecular networks. These can be modeled as graphs, with molecules as nodes and interactions as edges (Kreeger & Lauffenburger, 2010). While Boolean networks (where nodes are ON/OFF) are common for modeling signaling logic (Fumia & Martins, 2013), other approaches use systems of ODEs to describe the changing concentrations of each node in the network, allowing for a more dynamic and quantitative analysis of pathway flux. For prevention, these models can identify "choke points" in signaling pathways that are ideal drug targets (Stéphanou et al., 2018).

4.3. Strategic Interactions: Evolutionary Game Theory (EGT)

Cancer evolution is a multi-player game. EGT provides the mathematical language to analyze the strategic interactions between different cell populations (Gatenby & Vincent, 2003b). The replicator equation describes how the fraction (x_i) of a population using strategy i changes over time:

$$\frac{dx_i}{dt} = x_i(f_i - \bar{f}) \quad (6)$$

The fitness of each strategy (f_i) is determined by a payoff matrix that defines the outcome of interactions with other strategies. For example, in a "Public Goods" game, a cooperator cell that produces a growth factor pays a fitness cost, while a defector cell benefits from the growth factor without paying the cost. EGT can be used to design preventive strategies that alter these payoffs to favor less aggressive cell phenotypes (Zahir et al., 2020).

4.4. Hybrid Multiscale Strategy: Integrating ABMs, Networks, and EGT

Hybrid multiscale models integrate intracellular network dynamics, cell-level ABM behavior, and population-level game theory interactions into a single framework. Hybrid multiscale models integrate intracellular network dynamics, cell-level ABM behavior, and population-level evolutionary interactions within a single coherent framework. In practice, this integration enables agents to make decisions based on intracellular ODE/Boolean network states, while competitive dynamics across subpopulations are represented through evolutionary or game-theoretic formulations, and diffusible fields such as oxygen, nutrients, drugs, and cytokines are captured using PDE-based transport models. By explicitly linking these layers, hybrid frameworks create a mechanistic pathway from molecular regulation to collective tissue-scale behavior. Such coupled frameworks reproduce a range of clinically relevant phenomena, including the emergence of resistant subclones under selection pressure, adaptive responses to time-varying interventions, and eco-evolutionary feedback loops driven by microenvironmental constraints. They also capture scenarios in which clonal coexistence is stabilized by spatial structure and resource gradients, and they can generate invasion fronts and spatial patterning that arise from metabolic and signaling heterogeneity. Collectively, these capabilities make hybrid multiscale models particularly well suited for prevention and early intervention questions where initiation, selection, and niche dynamics are tightly intertwined. Hybrid multiscale frameworks are the most advanced models for cancer prevention and early intervention (Wang et al., 2025; Katsaounis et al., 2024; Chamseddine & Rejniak, 2020; Świerniak & Krześlak, 2016).

Figure 4 illustrates a unified workflow for translating these complex models into predictive, patient-specific Digital Twins. The process begins with multi-omics and clinical data acquisition to inform model structure. Appropriate model formalisms such as ODEs for systemic dynamics or ABMs for spatial heterogeneity are then selected. Calibration fits model parameters to training datasets, followed by rigorous validation against independent data. An iterative feedback loop ensures continuous refinement, enabling the model to accurately predict individualized treatment outcomes and support clinical decision-making. Importantly, agent-based and evolutionary modeling frameworks are increasingly being applied to cancer prevention by explicitly simulating early-stage cellular decision-making and evolutionary selection prior to tumor establishment. By representing premalignant cells, immune components, and stromal elements as interacting agents, these models can explore how immune surveillance, microenvironmental constraints, and competitive dynamics suppress or permit malignant transformation. Evolutionary game-theoretic approaches further enable the study of how preventive strategies—such as immune enhancement or microenvironmental modulation—can shift selective pressures to favor benign or less aggressive phenotypes. Such prevention-focused applications extend the utility of ABMs and EGT beyond treatment resistance, positioning them as powerful tools for understanding and interrupting early carcinogenic trajectories.

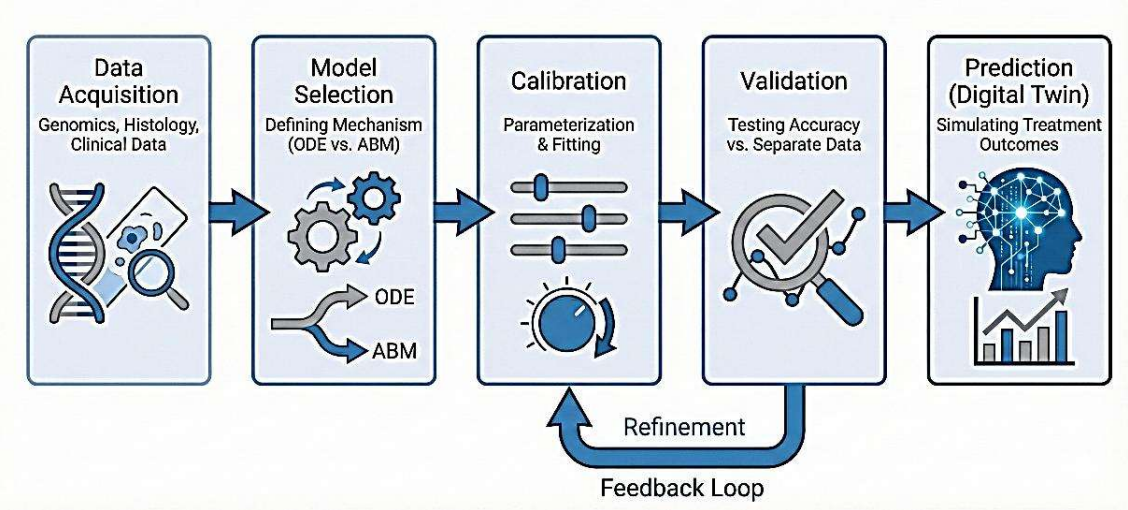


Figure 4. A unified workflow for developing predictive cancer models and patient-specific digital twins. Multi-omics and clinical data inform model structure, followed by selection of an appropriate formalism, parameter calibration, and validation. An iterative feedback loop ensures refinement, enabling accurate prediction of individualized treatment outcomes.

5. Model Integration and Clinical Translation

While individual modeling paradigms provide powerful insights, their ultimate value in oncology is realized only when they are integrated into predictive frameworks and rigorously tested against real-world data for clinical translation. Cancer is an intrinsically multi-scale phenomenon, spanning from molecular interactions within a cell to the dynamics of entire tissues; therefore, developing hybrid models that couple different formalisms (e.g., network models within agent-based systems) is a necessary step to create a holistic, mechanistic understanding (Kreeger & Lauffenburger, 2010). However, the increasing complexity of these integrative models necessitates a correspondingly rigorous process of validation and uncertainty quantification, which serves as the critical bridge from a computational simulation to a trusted tool for medical decision-making (Gross & MacLeod, 2017).

The pinnacle of this integrative and validation effort is the development of the Digital Twin, a patient-specific, dynamic model designed to forecast disease progression and response to therapy on an individual basis (Kaul et al., 2023). This section explores these crucial pillars of translational computational oncology: the methods for model integration, the requirements for establishing predictive credibility, and the visionary goal of the personalized digital twin.

5.1. The Power of Synergy: Hybrid Models

The next major step in this field is the development of hybrid models that combine different modeling approaches across various levels of biological organization. As an example, a detailed model of lung cancer prevention may rely on a partial differential equation to describe how carcinogens from cigarette smoke spread through the airway lining, while an agent based framework represents the behavior of epithelial cells as they mutate, divide, or undergo programmed cell death in reaction to that chemical environment. This allows for a direct link between environmental exposure and cellular-level consequences (Sun & Hu, 2018; Zangooei & Habibi, 2017; Rejniak & Anderson, 2011).

5.2. The Reality Check: Validation and Uncertainty Quantification

A model is only as good as its validation (Gross & MacLeod, 2017). This is a rigorous hierarchy: 1. Verification (checking the code is correct), 2. Calibration (fitting model parameters to specific in vitro data), 3. Validation against different datasets (in vivo, clinical), and 4. Prospective validation (predicting the outcome of a new experiment). Pattern-Oriented Modeling (POM) is a powerful validation technique where a model is assessed on its ability to simultaneously reproduce multiple observed patterns at different biological scales (Grimm et al., 2005). Uncertainty quantification (UQ) assesses how uncertainties in model inputs affect the confidence in its predictions (Seoni et al., 2023).

5.3. The Future Vision: Towards Digital Twins

The convergence of multi-scale modeling, AI, and vast patient data is paving the way for the Digital Twin—a continuously updated virtual model of an individual (Croatti et al., 2020). The immense data requirements (longitudinal multi-omics, high-resolution imaging, lifestyle data) remain a major bottleneck. However, recent advancements focus on creating actionable digital twins for optimizing cancer therapy, a concept directly transferable to personalizing prevention strategies by simulating an individual's lifelong risk under different scenarios (Chaudhuri et al., 2023).

5.4. Data Integration and Interoperability Challenges

One of the most significant barriers to effective model integration and clinical translation is the heterogeneity of biomedical data, which includes genomics, epigenomics, transcriptomics, proteomics, radiological imaging, histopathology, clinical records, and exposure or lifestyle information (Baião et al., 2025; Ran et al., 2025; Alkim et al., 2023). The datasets differ in structure, dimensionality, temporal resolution, measurement noise, missingness/sampling bias, and cohort-specific variability (Hsu et al., 2026; Ali, 2023). Meaningful interoperability necessitates consistent data forms, strong preparation pipelines, and semantic alignment across biological scales (Farhadizadeh et al., 2025; Acharya & Mukhopadhyay, 2024). Translational modeling relies heavily on ontologies, common data models, FAIR principles, and data management infrastructures (Ali, 2023). Without such harmonization, even the most powerful hybrid multiscale models may be limited by inconsistent or incomplete inputs, reducing their prediction power in clinical and preventative oncology (Baião et al., 2025; Farhadizadeh et al., 2025; Acharya & Mukhopadhyay, 2024).

5.5. The Role of Artificial Intelligence in Model Augmentation

While mechanistic models excel at interpretability and causal reasoning, they frequently suffer with complex parameter spaces and incomplete biological knowledge (Metzcar et al., 2024). AI and machine learning (ML) technologies provide a complementary solution by extracting latent patterns directly from data, improving model scalability and prediction performance (Wysocka et al., 2023). Recent research focuses on AI-augmented hybrid models, which include machine learning components into mechanistic frameworks (Butner et al., 2024). Deep learning can be used to infer patient-specific parameters, surrogate models to accelerate computationally intensive simulations, and reinforcement learning to optimize treatment and preventative efforts (Butner et al., 2024; Gherman et al., 2023). Importantly, when AI is bound by mechanistic priors, it keeps interpretability while reducing overfitting, which is necessary for clinical acceptability (Si et al., 2025; Butner et al., 2024; Gherman et al., 2023).

5.6. Data-Driven Benchmarking and Quantitative Comparison of Prevention-Oriented Models

While qualitative comparison remains essential for understanding the conceptual strengths and limitations of different modeling approaches, the growing availability of large-scale biomedical datasets enables more systematic, data-driven benchmarking of computational models. Resources such as The Cancer Genome Atlas (TCGA), longitudinal cohort studies, and early-phase clinical trials provide multi-omics, imaging, and clinical outcome data that can be leveraged to quantitatively evaluate model performance, particularly in prevention and early-stage disease settings. A prevention-oriented benchmarking framework may assess models across multiple dimensions, including predictive accuracy for early lesion progression or regression, consistency with known biological constraints, computational efficiency, and robustness under data uncertainty.

For example, hybrid mechanistic–AI models can be evaluated by their ability to integrate genomic risk profiles with simulated responses to chemopreventive interventions, comparing predicted trajectories against observed molecular or clinical endpoints. Metrics such as time-to-progression prediction error, stability of inferred parameters, and sensitivity to data sparsity provide quantitative criteria that complement traditional qualitative assessments. Importantly, such benchmarking does not imply a single “best” model but rather enables context-dependent selection of modeling strategies aligned with specific prevention goals. By outlining how real-world datasets can be systematically incorporated into comparative evaluation pipelines, this review aims to bridge the gap between methodological development and evidence-based deployment of computational models in precision cancer prevention. To make such quantitative comparisons actionable, benchmarking should explicitly report (i) the dataset split and external validation setting, (ii) primary prevention-oriented endpoints (e.g., progression within a fixed horizon or time-to-progression), (iii) uncertainty and robustness under missingness or domain shift, and (iv) computational cost (Table 2). In practice, screening trials and longitudinal cohorts are particularly valuable for prevention (risk stratification, overdiagnosis, time-to-progression), while molecular repositories such as TCGA can support omics-informed components and biological plausibility checks for early-stage signatures. Establishing a standardized reporting checklist for these metrics would improve reproducibility and enable fair comparisons across modeling paradigms.

Table 2. Example datasets and quantitative endpoints for prevention-oriented benchmarking.

Data source (example)	Prevention / early-stage question it supports	Suitable model components	Quantitative metrics to report
Screening trials & cohorts (e.g., lung screening trials, longitudinal cohorts)	Risk stratification, screening intensity, overdiagnosis, time-to-progression	Natural-history / population models; decision models; hybrid link to tissue modules	AUC/C-index, calibration, decision-curve/net benefit, time-to-event error, overdiagnosis estimates
Early-phase clinical trials (prevention or interception)	Chemopreventive response, biomarker modulation, safety/benefit tradeoffs	PK/PD + intracellular networks; hybrid mechanistic–AI inference	Effect-size prediction error, responder classification accuracy, calibration, uncertainty intervals, runtime
Multi-omics repositories (e.g., TCGA, where applicable)	Omics-informed signatures, pathway constraints, early-stage biological plausibility	ML/AI modules; mechanistic-learning parameter inference; surrogate modeling	External validation performance, domain-shift robustness, feature stability, uncertainty/robustness
Imaging / pathology datasets (premalignant or early lesions when available)	Early lesion detection/trajectory, microenvironment patterns	Spatial models (PDE/ABM), image-informed priors	Segmentation/detection metrics, progression prediction accuracy, calibration, computational efficiency

5.7. Worked Example: A Quantitative Prevention-Oriented Multiscale Workflow (Illustrative)

To make the prevention focus more tangible, we provide an illustrative end-to-end example of how an integrated multiscale framework can be used to answer a concrete prevention question: how should screening intensity and chemopreventive intervention be personalized for individuals at elevated risk of early lesion progression? In a representative workflow, prevention-relevant covariates (e.g., age, exposure history, family history, and genomic risk proxies) first define an individualized baseline risk. A population/natural-history component then maps these inputs to a probabilistic trajectory of premalignant initiation and progression (e.g., incidence of early lesions and time-to-progression distributions), producing patient-specific priors for downstream mechanistic simulation. At the tissue and cellular scales, a hybrid multiscale module can couple a diffusion process (e.g., PDE fields representing carcinogen exposure or inflammatory mediators) with an agent-based representation of epithelial and immune/stromal cells. Intracellular decision rules (e.g., simplified network logic or ODE surrogates) govern mutation accumulation, proliferation, apoptosis, and immune clearance, allowing the model to simulate how preventive strategies shift early evolutionary selection and niche constraints. Preventive interventions (e.g., altered exposure, immune enhancement, or a chemopreventive dosing schedule) are then represented as parameterized perturbations to the exposure field, cellular fitness, or drug-effect terms. Model outputs can be defined quantitatively in prevention terms, such as predicted probability of lesion regression versus progression, expected time-to-advanced lesion, and comparative risk reduction under alternative prevention scenarios. Crucially, the example can be evaluated using prevention-oriented benchmarking criteria consistent with Section 5.6: discrimination (e.g., AUC for predicting progression within a specified horizon), calibration (agreement between predicted and observed risks), time-to-event prediction error, robustness under data sparsity, and computational efficiency (e.g., wall-clock time per individualized simulation). Such quantitative endpoints translate multiscale simulations into testable claims and provide a practical bridge between methodological development, model credibility, and decision support. Real-world precedents for prevention/screening impact further motivate this approach. For example, NCI-sponsored comparative simulation efforts such as CISNET explicitly use modeling to assess prevention and

screening strategies and to quantify benefits and harms under alternative policies, illustrating how model-based evidence can support guideline-level decisions. While CISNET models operate primarily at the population and decision level, their prevention-oriented validation culture and comparative framework provide a useful template for integrating mechanistic multiscale modules into next-generation prevention digital-twin pipelines.

6. Computational Frameworks and Software Ecosystem

The translation of mathematical constructs into predictive simulations necessitates robust computational platforms. While early cancer models were often implemented using custom scripts in MATLAB or Mathematica, the field has matured towards standardized, open-source simulation environments. These platforms are designed to handle the computational complexity of solving large systems of ODEs or simulating millions of interacting agents, thereby enhancing reproducibility and allowing researchers to focus on biological questions rather than low-level code implementation. Choosing the appropriate software ecosystem is a critical decision that depends on the specific scale of interest, the required computational efficiency, and the learning curve.

6.1. Comparison of Modeling Platforms

Several dedicated platforms have emerged, each optimized for specific modeling paradigms. For continuum models (PDEs) and systems biology (ODEs/Networks), tools like COPASI and VCell provide user-friendly interfaces for simulating intracellular dynamics and reaction-diffusion systems. In contrast, the spatial complexity of the tumor microenvironment is best captured by agent-based platforms. PhysiCell and Chaste represent high-performance frameworks written in C++ that are capable of simulating millions of cells, making them ideal for large-scale tissue simulations. CompuCell3D, based on the Cellular Potts Model (CPM), offers a unique approach to modeling cell shape changes and motility using an energy-minimization formalism. For educational purposes and rapid prototyping, NetLogo remains a popular choice due to its intuitive interface, although it may lack the performance required for clinical-grade simulations. From a prevention and early-stage perspective, these platforms are particularly valuable when they are mapped to endpoints such as premalignant lesion initiation/progression, immune surveillance efficacy, exposure-driven risk trajectories, and simulated chemopreventive responses. Accordingly, we complement the methodological comparison by highlighting, for each platform, representative prevention-oriented use cases that connect computational choices to early carcinogenesis and prevention decision support. Table 3 provides a comparative analysis of these primary computational frameworks, highlighting their underlying methodologies, programming requirements, and typical applications in oncology.

Table 3. Comparative analysis of computational platforms for cancer modeling.

Software Platform	Primary Modeling Paradigm	Core Language / Interface	Key Features & Strengths	Typical Application in Oncology
PhysiCell	Agent-Based Model (Off-lattice)	C++ / Python (XML config)	Highly scalable (simulates millions of cells); incorporates BioFVM for fast diffusion solving; open-source.	Large-scale tumor growth; immunotherapy interactions; spatially-dependent drug response; early clonal expansion and immune surveillance in premalignant tissue; exposure/chemoprevention gradient simulations.
CompuCell3D	Cellular Potts Model (Lattice-based)	Python / C++	Models cell shape, adhesion, and membrane fluctuations via energy minimization; highly modular.	Angiogenesis; cell migration mechanisms; tissue morphogenesis; invasion dynamics; early dysplastic pattern formation and premalignant tissue remodeling where cell mechanics/shape are central.
Chaste	Hybrid (ABM, PDE, ODE)	C++	Computational durability; originally for cardiac modeling; rigorous code testing; distinct "node-based" mechanics.	Colorectal crypt dynamics; detailed cell-cycle modeling; mechanical force propagation in tissues; stem-cell niche dynamics and early adenoma initiation/progression relevant to prevention and early detection.
COPASI	ODE / Stochastic / Network	GUI (C++ backend)	User-friendly (no coding required); supports SBML standards; excellent for parameter estimation.	Intracellular signaling pathways (e.g., MAPK, p53); PK/PD modeling; metabolic flux analysis; simulation of chemopreventive pathway modulation and early-stage biomarker-linked dynamics.
NetLogo	Agent-Based Model	NetLogo (Java-based)	Easy to learn; rapid prototyping; extensive model library; excellent for visualization.	Educational models; basic immune–tumor predator–prey dynamics; hypothesis testing; rapid prototyping of prevention scenarios (e.g., screening adherence behaviors, simplified risk-factor accumulation, conceptual immune surveillance).

6.2. Selection Criteria for Researchers

The choice of modeling platform should come directly from the biological question being studied. When the focus is on events inside the cell, such as signaling pathways, or on processes like drug behavior in the body without any spatial structure, tools such as COPASI or ordinary differential equation solvers are usually sufficient. When the main interest is in spatial patterns, for instance the development of low-oxygen regions inside a tumor or the movement of immune cells into cancer tissue, platforms based on spatial agent behavior become essential. PhysiCell is currently the strongest option for studies that require very large numbers of cells, such as solid tumor simulations, while CompuCell3D is better suited for questions that depend on cell mechanics and changes in cell shape. With the growing interest in multiscale modeling, new interfaces have been created that let these tools work together, for example by placing a COPASI signaling network inside a PhysiCell agent. These advances are pushing the field toward more complete and realistic digital twins. From a practical perspective, the selection of an appropriate computational platform can be viewed as a scale-driven decision tree. When the primary research question concerns intracellular signaling, pharmacokinetics, or population-averaged dynamics without explicit spatial structure, equation-based platforms such as COPASI or ODE-focused solvers are generally sufficient. If spatial gradients, diffusion processes, or tissue-level heterogeneity are central to the problem—such as modeling oxygen, nutrient, or drug transport—continuum or hybrid PDE-based environments become more appropriate. For questions that require explicit representation of individual cells, heterogeneity, and local interactions, particularly in early-stage cancer or prevention-oriented scenarios, agent-based platforms are preferable. Among these, PhysiCell is well suited for large-scale tissue simulations involving millions of cells, whereas CompuCell3D is advantageous when cell shape, adhesion, and mechanical interactions play a dominant role. Hybrid frameworks, such as Chaste or coupled environments integrating intracellular networks with ABMs, are most appropriate when multiscale interactions across molecular, cellular, and tissue levels must be captured simultaneously. This conceptual decision pathway complements the comparative overview in Table 2 and provides researchers with a rapid, biologically grounded strategy for selecting modeling platforms aligned with their specific scientific objectives.

7. Conclusion

This review traces the progress of mathematical and computational modeling in cancer prevention, from early population-level ODE models to advanced spatial, multi-scale, and agent-based frameworks that can capture the varied and adaptable nature of tumor ecosystems. The research has progressed from non-spatial formulations to spatially explicit PDEs, cellular automata, and, finally, individual-based ABMs, resulting in models that more realistically mimic biological reality and the milieu in which cancer originates. We also highlight the expanding importance of intracellular network models and evolutionary game theory, which shed light on how molecular decision-making and strategic interactions influence early carcinogenesis and treatment resistance. A major development since 2022 has been the move from isolated modeling approaches toward fully connected predictive systems. Modern methods increasingly combine deep learning with mechanistic modeling, incorporate strong methods for measuring uncertainty, and aim to create tools that can eventually be used in clinical settings. Hybrid models that operate across multiple biological levels reflect this integration by linking environmental exposures directly to cellular responses, while structured validation procedures and pattern-oriented strategies help maintain predictive reliability. All of these advances contribute to the rise of the Digital Twin concept, which represents a personalized and continuously updated computational version of a patient that can

anticipate the course of cancer and test preventive strategies before they are applied. Despite these advancements, several knowledge gaps continue to limit the translation of prevention-oriented models into clinically usable tools. First, prevention requires longitudinal and pre-diagnostic data—capturing exposure history, lifestyle dynamics, immune surveillance, and premalignant lesion trajectories—yet such datasets remain comparatively scarce, fragmented, and inconsistently curated. Second, multimodal data integration is still constrained by weak interoperability across omics, imaging, and clinical records, which complicates coherent calibration across biological scales. Third, as models become more mechanistically detailed and multiscale, challenges of identifiability, parameter sloppiness, and propagation of uncertainty become central barriers to trustworthy inference and decision-making.

Table 4. Summary of recent works based on provided literature.

Reference	Model Type(s)	Application / Biological Problem	Key Finding / Implication
(Hanahan, 2022)	Conceptual Framework	Redefining the "Hallmarks of Cancer" for the new decade.	Introduces new hallmarks like phenotypic plasticity and senescent cells, providing new, critical targets for mathematical models to simulate.
(Kaul et al., 2023)	AI/Digital Twin	Reviewing the role of Artificial Intelligence in developing Digital Twins for cancer care.	Concludes that AI is a critical engine for integrating complex data needed to build predictive and personalized digital twins.
(Seoni et al., 2023)	Uncertainty Quantification & AI	Reviewing methods to quantify uncertainty in AI models for healthcare applications.	Emphasizes that for clinical trust, AI models must report their own confidence, a key feature for future cancer models.
(Chaudhuri et al., 2023)	Predictive Digital Twin	Patient-specific optimization of radiotherapy regimens for high-grade gliomas.	Successfully demonstrated a digital twin that predicts treatment response and helps personalize radiation plans under uncertainty.
(Hallek, 2025)	Clinical & Epidemiological	Providing a 2025 update on Chronic Lymphocytic Leukemia (CLL) therapy and pathogenesis.	Offers the latest clinical data which is essential for parameterizing and validating new mathematical models aimed at CLL.

To move toward operational cancer digital twins for prevention, the field benefits from a stepwise roadmap. (1) Prevention-relevant data foundations: prioritize standardized longitudinal cohorts and harmonized pipelines that enable reproducible linkage among omics, imaging, clinical phenotypes, and exposure variables. (2) Hybrid model architecture: combine mechanistic multiscale structure with data-driven inference (mechanistic learning) to enable personalization while maintaining interpretability and causal plausibility. (3) Credibility and validation: adopt explicit uncertainty quantification, external validation across cohorts, and prospective evaluation against pre-specified endpoints (e.g., risk stratification accuracy, time-to-progression prediction, and predicted benefit of preventive interventions). (4) Computational scalability: use model reduction, surrogate modeling, and efficient simulation strategies so that individualized analyses become feasible in near-realistic clinical time constraints. (5) Clinical deployment principles: frame digital twins as decision-support systems that report calibrated confidence, are auditable and reproducible, and are updated responsibly as new data arrive. Overall, strengthening these pillars can shift digital twins from a conceptual aspiration to an evidence-based, prevention-oriented clinical technology: one that supports individualized risk forecasting, prioritization of screening intensity, and simulation of preventive strategies under uncertainty across the lifespan. The studies summarized in Table 4 illustrate this contemporary approach, showing how recent research is dedicated to establishing both

the conceptual and technical foundations for the next generation of predictive oncology, ranging from artificial intelligence-based frameworks to patient specific digital twins.

Data Availability Statement: No new data were generated or analyzed in this study.

Conflicts of Interest: The authors declare no conflicts of interest.

Use of Artificial Intelligence Tools: Artificial intelligence–based tools were used exclusively to assist with language editing, grammatical refinement, and improvement of stylistic clarity. Specifically, ChatGPT (OpenAI) was employed to support English language polishing and to enhance the coherence and readability of the manuscript. All scientific content, including the conceptualization of the review, selection and critical interpretation of the literature, formulation and explanation of mathematical models, preparation of figures and tables, and the development of conclusions, was performed entirely by the authors. No artificial intelligence system was used to generate original scientific ideas, analytical results, or references. All sources contributing to the manuscript are appropriately cited in the reference list.

References

Acharya, D., & Mukhopadhyay, A. (2024). A comprehensive review of machine learning techniques for multi-omics data integration: Challenges and applications in precision oncology. *Briefings in Functional Genomics*, 23(5), 549-560. https://doi.org/10.1093/bfgp/elae013

Agnello, A., Amorisco, N.C., Keats, A., Holt, G.K., Buchanan, J., Pamela, S., Vincent, C., & McArdle, G. (2024). Emulation techniques for scenario and classical control design of tokamak plasmas. *Physics of Plasmas*, 31(4), 1-12. https://doi.org/10.1063/5.0187822

Ali, H. (2023). Artificial intelligence in multi-omics data integration: Advancing precision medicine, biomarker discovery and genomic-driven disease interventions. *International Journal of Science and Research Archive*, 8(1), 1012-1030. https://doi.org/10.30574/ijsra.2023.8.1.0189

Alkim, E., Dowst, H., DiCarlo, J., Dobrolecki, L.E., Hernández-Herrera, A., Hormuth, D.A., Liao, Y., McOwiti, A., Pautler, R., & Rimawi, M. (2023). Toward practical integration of omic and imaging data in co-clinical trials. *Tomography*, 9(2), 810-828. https://doi.org/10.3390/tomography9020066

Altrock, P.M., Liu, L.L., & Michor, F. (2015). The mathematics of cancer: Integrating quantitative models. *Nature Reviews Cancer*, 15(12), 730-745. https://doi.org/10.1038/nrc4029

Anderson, A.R., & Quaranta, V. (2008). Integrative mathematical oncology. *Nature Reviews Cancer*, 8(3), 227-234. https://doi.org/10.1038/nrc2329

Azizi, T. (2024). Mathematical modeling of cancer progression. *AppliedMath*, *4*(3), 1065-1079. https://doi.org/10.3390/appliedmath4030057

Azizi, T. (2025). Mathematical modelling of cancer treatments, resistance, optimization. *AppliedMath*, 5(2), 40. https://doi.org/10.3390/appliedmath5020040

Baião, A., Cai, Z., Poulos, R., Robinson, P., Reddel, R., Zhong, Q., Vinga, S., & Gonçalves, E. (2025). A technical review of multi-omics data integration methods: From classical statistical to deep generative approaches. *Briefings in Bioinformatics, 26*(4), bbaf355. https://doi.org/10.1093/bib/bbaf355

Barabasi, A.L., & Oltvai, Z.N. (2004). Network biology: understanding the cell's functional organization. *Nature Reviews Genetics*, 5(2), 101-113. https://doi.org/10.1038/nrg1272

Bora, Ş., & Emek, S. (2019). Agent-based modeling and simulation of biological systems. *Computer Simulation*, *IntechOpen*. 1-17. https://doi.org/10.5772/intechopen.80070

Brunton, S.L., & Kutz, J.N. (2022). *Data-driven science and engineering:* Machine learning, dynamical systems, and control. *Cambridge University Press*. https://doi.org/10.1017/9781108380690

Butner, J.D., Dogra, P., Chung, C., Koay, E.J., Welsh, J.W., Hong, D.S., Cristini, V., & Wang, Z. (2024). Hybridizing mechanistic modeling and deep learning for personalized survival prediction after immune checkpoint inhibitor immunotherapy. *npj Systems Biology and Applications*, 10, 88. https://doi.org/10.1038/s41540-024-00415-8

Byrne, H.M. (2010). Dissecting cancer through mathematics: from the cell to the animal model. *Nature Reviews Cancer*, *10*(3), 221-230. https://doi.org/10.1038/nrc2808

Camacho, D.M., Collins, K.M., Powers, R.K., Costello, J.C., & Collins, J.J. (2018). Next-generation machine learning for biological networks. *Cell*, 173(7), 1581-1592. https://doi.org/10.1016/j.cell.2018.05.015

Chamseddine, I.M., & Rejniak, K.A. (2020). Hybrid modeling frameworks of tumor development and treatment. *Wiley Interdisciplinary Reviews Systems Biology and Medicine*, 12(1), e1461. https://doi.org/10.1002/wsbm.1461

Chaudhuri, A., Pash, G., Hormuth, D.A., Lorenzo, G., Kapteyn, M., Wu, C., Lima, E.A.B.F., Yankeelov, T.E., & Willcox, K. (2023). Predictive digital twin for optimizing patient-specific radiotherapy regimens under uncertainty in high-grade gliomas. *Frontiers in Artificial Intelligence*, 6, 1222612. https://doi.org/10.3389/frai.2023.1222612

Clarke, R., Tyson, J.J., Tan, M., Baumann, W.T., Jin, L., Xuan, J., & Wang, Y. (2019). Systems biology: Perspectives on multiscale modeling in research on endocrine-related cancers. *Endocrine-Related Cancer*, 26(6), R345-R368. https://doi.org/10.1530/ERC-18-0309

Croatti, A., Gabellini, M., Montagna, S., & Ricci, A. (2020). On the integration of agents and digital twins in healthcare. *Journal of Medical Systems*, 44(9), 161. https://doi.org/10.1007/s10916-020-01623-5

Deisboeck, T.S., Wang, Z., Macklin, P., & Cristini, V. (2011). Multiscale cancer modeling. *Annual Review of Biomedical Engineering*, 13, 127-155. https://doi.org/10.1146/annurev-bioeng-071910-124729

Farhadizadeh, M., Weymann, M., Blaß, M., Kraus, J., Gundler, C., Walter, S., Hempen, N., Binder, H., & Binder, N. (2025). A systematic review of challenges and proposed solutions in modeling multimodal data. *arXiv*, 1-49. https://doi.org/10.48550/arXiv.2505.06945

Fumia, H.F., & Martins, M.L. (2013). Boolean network model for cancer pathways: Predicting carcinogenesis and targeted therapy outcomes. *PloS One*, 8(7), e69008. https://doi.org/10.1371/journal.pone.0069008

Gatenby, R.A., & Vincent, T.L. (2003a). Application of quantitative models from population biology and evolutionary game theory to tumor therapeutic strategies. *Molecular Cancer Therapeutics*, 2(9), 919-927.

Gatenby, R.A., & Vincent, T.L. (2003b). An evolutionary model of carcinogenesis. *Cancer Research*, 63(19), 6212-6220.

Gerlee, P., & Anderson, A.R. (2007). An evolutionary hybrid cellular automaton model of solid tumour growth. *Journal of Theoretical Biology*, 246(4), 583-603. https://doi.org/10.1016/j.jtbi.2007.01.027

Ghassemi, M., Naumann, T., Schulam, P., Beam, A.L., Chen, I.Y., & Ranganath, R. (2020). A review of challenges and opportunities in machine learning for health. *AMIA Summits on Translational Science Proceedings*, *2020*, 191-200.

Gherman, I.M., Abdallah, Z.S., Pang, W., Gorochowski, T.E., Grierson, C.S., & Marucci, L. (2023). Bridging the gap between mechanistic biological models and machine learning surrogates. *PLos Computational Biology*, 19(4), e1010988. https://doi.org/10.1371/journal.pcbi.1010988

Gonçalves, I.G., & García-Aznar, J.M. (2023). Hybrid computational models of multicellular tumour growth considering glucose metabolism. *Computational and Structural Biotechnology Journal*, 21, 1262-1271. https://doi.org/10.1016/j.csbj.2023.01.044

Gong, C., Milberg, O., Wang, B., Vicini, P., Narwal, R., Roskos, L., & Popel, A.S. (2017). A computational multiscale agent-based model for simulating spatio-temporal tumour immune response to PD1 and PDL1 inhibition. *Journal of the Royal Society Interface*, 14(134), 20170320. https://doi.org/10.1098/rsif.2017.0320

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W.M., Railsback, S.F., Thulke, H.H., Weiner, J., Wiegand, T., & DeAngelis, D.L. (2005). Pattern-oriented modeling of agent-based complex systems: Lessons from ecology. *Science*, 310(5750), 987-991. https://doi.org/10.1126/science.1116681

Gross, F., & MacLeod, M. (2017). Prospects and problems for standardizing model validation in systems biology. *Progress in Niophysics and Molecular Biology*, 129, 3-12. https://doi.org/10.1016/j.pbiomolbio.2017.01.003

Hallek, M. (2025). Chronic Lymphocytic Leukemia: 2025 update on the epidemiology, pathogenesis, diagnosis, and therapy. *American Journal of Hematology*, *100*(3), 450–480. https://doi.org/10.1002/ajh.27546

Hanahan, D. (2022). Hallmarks of cancer: New dimensions. *Cancer Discovery*, 12(1), 31-46. https://doi.org/10.1158/2159-8290.CD-21-1059

Hickey, J.W., Agmon, E., Horowitz, N., Lamore, M., Sunwoo, J., Covert, M., & Nolan, G.P. (2023). Integrating multiplexed imaging and multiscale modeling identifies tumor phenotype transformation as a critical component of therapeutic T cell efficacy. *BioRxiv*, 12.06.570168. https://doi.org/10.1101/2023.12.06.570168

Hsu, C.Y., Askar, S., Alshkarchy, S.S., Nayak, P.P., Attabi, K.A., Khan, M.A., Mayan, J.A., Sharma, M., Islomov, S., & Soleimani Samarkhazan, H. (2026). AI-driven multi-omics integration in precision oncology: bridging the data deluge to clinical decisions. *Clinical and Experimental Medicine*, 26, 29. https://doi.org/10.1007/s10238-025-01965-9

Huang, S., Ernberg, I., & Kauffman, S. (2009). Cancer attractors: A systems view of tumors from a gene network dynamics and developmental perspective. *Seminars in Cell & Developmental Biology*, 20(7), 869-876. https://doi.org/10.1016/j.semcdb.2009.07.003

Jarrett, A.M., Lima, E.A.B.F., Hormuth, D.A., McKenna, M.T., Feng, X., Ekrut, D.A., Resende, A.C.M., Brock, A., & Yankeelov, T.E. (2018). Mathematical models of tumor cell proliferation: A review of the literature. *Expert Review of Anticancer Therapy*, 18(12), 1271-1286. https://doi.org/10.1080/14737140.2018.1527689

Kansal, A.R., Torquato, S., Harsh, G.R., Chiocca, E.A., & Deisboeck, T.S. (2000). Simulated brain tumor growth dynamics using a three-dimensional cellular automaton. *Journal of Theoretical Biology*, 203(4), 367-382. https://doi.org/10.1006/jtbi.2000

Katsaounis, D., Harbour, N., Williams, T., Chaplain, M.A., & Sfakianakis, N. (2024). A genuinely hybrid, multiscale 3d cancer invasion and metastasis modelling framework. *Bulletin of Mathematical Biology*, *86*(6), 64. https://doi.org/10.1007/s11538-024-01286-0

Kaul, R., Ossai, C., Forkan, A.R.M., Jayaraman, P.P., Zelcer, J., Vaughan, S., & Wickramasinghe, N. (2023). The role of AI for developing digital twins in healthcare: The case of cancer care. *WIREs Data Mining and Knowledge Discovery,* 13(1), e1480. https://doi.org/10.1002/widm.1480

Kreeger, P.K., & Lauffenburger, D.A. (2010). Cancer systems biology: a network modeling perspective. *Carcinogenesis*, 31(1), 2-8. https://doi.org/10.1093/carcin/bgp261

Li, C., & Lei, J. (2025). Mathematical modeling of tumor-immune interactions: Methods, applications, and future perspectives. *arXiv,* 1-72. https://doi.org/10.4208/csiam-ls.SO-2024-0008

Metzcar, J., Jutzeler, C.R., Macklin, P., Köhn-Luque, A., & Brüningk, S.C. (2024). A review of mechanistic learning in mathematical oncology. *Frontiers in Immunology*, 15, 1363144. https://doi.org/10.3389/fimmu.2024.1363144

Metzcar, J., Wang, Y., Heiland, R., & Macklin, P. (2019). A review of cell-based computational modeling in cancer biology. *JCO Clinical Cancer Informatics*, 3, 1-13. https://doi.org/10.1200/CCI.18.00069

Mirzaei, N.M., & Shahriyari, L. (2024). Modeling cancer progression: An integrated workflow extending data-driven kinetic models to bio-mechanical PDE models. *Physical Biology*, 21(2), 022001. https://doi.org/10.1088/1478-3975/ad2777

Nardini, J.T., Baker, R.E., Simpson, M.J., & Flores, K.B. (2021). Learning differential equation models from stochastic agent-based model simulations. *Journal of the Royal Society Interface*, 18(176), 20200987. https://doi.org/10.1098/rsif.2020.0987

Norton, K.A., Gong, C., Jamalian, S., & Popel, A.S. (2019). Multiscale agent-based and hybrid modeling of the tumor immune microenvironment. *Processes*, 7(1), 37. https://doi.org/10.3390/pr7010037

Poleszczuk, J., Macklin, P., & Enderling, H. (2016). Agent-based modeling of cancer stem cell driven solid tumor growth. *In Methods in Molecular Biology,* 1516, 335-346. https://doi.org/10.1007/7651-2016-346

Rackauckas, C., Ma, Y., Martensen, J., Warner, C., Zubov, K., Supekar, R., Skinner, D., Ramadhan, A., & Edelman, A. (2020). Universal differential equations for scientific machine learning. *arXiv*, 1-55. https://doi.org/10.48550/arXiv.2001.04385

Ramis-Conde, I., Chaplain, M.A., & Anderson, A.R. (2008). Mathematical modelling of cancer cell invasion of tissue. *Mathematical and Computer Modelling*, 47(5-6), 533-545. https://doi.org/10.1016/j.mcm.2007.02.034

Ran, D., Li, J., Zhao, M., Du, L., Zhang, Y., & Zhu, J. (2025). Artificial intelligence integrates multi-omics data for precision stratification and drug resistance prediction in breast cancer. *Frontiers in Oncology*, 15, 1612474. https://doi.org/10.3389/fonc.2025.1612474

Rejniak, K.A., & Anderson, A.R. (2011). Hybrid models of tumor growth. *Wiley Interdiscip Rev Syst Biol Med*, 3(1), 115-125. https://doi.org/10.1002/wsbm.102

Rejniak, K.A., & Anderson, A.R. (2011). Hybrid models of tumor growth. *WIREs Systems Biology and Medicine, 3*(1), 115–125. https://doi.org/10.1002/wsbm.102

Rikard, S.M., Athey, T.L., Nelson, A.R., Christiansen, S.L., Lee, J.J., Holmes, J.W., Peirce, S.M., & Saucerman, J.J. (2019). Multiscale coupling of an agent-based model of tissue fibrosis and a logic-based model of intracellular signaling. *Frontiers in Physiology*, 10, 1481. https://doi.org/10.3389/fphys.2019.01481

Robertson-Tessi, M., El-Kareh, A., & Goriely, A. (2012). A mathematical model of tumor-immune interactions. *Journal of Theoretical Biology*, 294, 56-73. https://doi.org/10.1016/j.jtbi.2011.10.027

Rubanova, Y., Chen, R.T.Q., & Duvenaud, D. (2019). Latent ordinary differential equations for irregularly-sampled time series. *arXiv*, 1-21. https://doi.org/10.48550/arXiv.1907.03907

Ruiz-Martinez, A., Gong, C., Wang, H., Sove, R.J., Mi, H., Kimko, H., & Popel, A.S. (2022). Simulations of tumor growth and response to immunotherapy by coupling a spatial agent-based model with a whole-patient quantitative systems pharmacology model. *PLos Computational Biology*, 18(7), e1010254. https://doi.org/10.1371/journal.pcbi.1010254

Sancho-Araiz, A., Mangas-Sanjuan, V., & Trocóniz, I.F. (2021). The role of mathematical models in immuno-oncology: challenges and future perspectives. *Pharmaceutics*, 13(7), 1016. https://doi.org/10.3390/pharmaceutics13071016

Seoni, S., Jahmunah, V., Salvi, M., Barua, P.D., Molinari, F., & Acharya, U.R. (2023). Application of uncertainty quantification to artificial intelligence in healthcare: A review of last decade (2013-2023). *Computers in Biology and Medicine*, 165, 107441. https://doi.org/10.1016/j.compbiomed.2023.107441

Si, H., Kumar, S., Lata, S., Ahmad, A., Ghosh, S., Stephansen, K., Nagarkar, D., Zhou, E., & Higgs, B.W. (2025). Mechanistically explainable AI model for predicting synergistic cancer therapy combinations. *Current Oncology*, 32(10), 548. https://doi.org/10.3390/curroncol32100548

Stéphanou, A., Fanchon, E., Innominato, P.F., & Ballesta, A. (2018). Systems biology, systems medicine, systems pharmacology: The what and the why. *Acta Biotheoretica*, 66, 345-365. https://doi.org/10.1007/s10441-018-9330-2

Sun, X., & Hu, B. (2018). Mathematical modeling and computational prediction of cancer drug resistance. *Briefings in Bioinformatics*, 19(6), 1382-1399. https://doi.org/10.1093/bib/bbx065

Świerniak, A., & Krześlak, M. (2016). Cancer heterogeneity and multilayer spatial evolutionary games. *Biol Direct*, 11(1), 53. https://doi.org/10.1186/s13062-016-0156-z

Tomlinson, I., & Bodmer, W. (1997). Modelling the consequences of interactions between tumour cells. *British Journal of Cancer*, 75(2), 157-160. https://doi.org/10.1038/bjc.1997.26

Villa, C., Maini, P.K., Browning, A.P., Jenner, A.L., Hamis, S., & Cassidy, T. (2025). Reducing phenotype-structured partial differential equations models of cancer evolution to systems of ordinary differential equations: A generalised moment dynamics approach. *Journal of Mathematical Biology*, 91(2), 22. https://doi.org/10.1007/s00285-025-02246-5

Wang, L.S., Yu, J., Li, S., & Liu, Z. (2025). Analysis and mean-field limit of a hybrid PDE-ABM modeling angiogenesis-regulated resistance evolution. *Mathematics*, 13(17), 2898. https://doi.org/10.3390/math13172898

Wolfram, S., & Gad-el-Hak, M. (2003). A new kind of science. *Applied Mechanics Reviews*, 56(2), B18-B19. https://doi.org/10.1115/1.1553433

Wysocka, M., Wysocki, O., Zufferey, M., Landers, D., & Freitas, A. (2023). A systematic review of biologically-informed deep learning models for cancer: Fundamental trends for encoding and interpreting oncology data. *BMC Bioinformatics*, 24, 198. https://doi.org/10.1186/s12859-023-05262-8

Yankeelov, T.E., An, G., Saut, O., Luebeck, E.G., Popel, A.S., Ribba, B., Vicini, P., Zhou, X., Weis, J.A., & Ye, K. (2016). Multi-scale modeling in clinical oncology: Opportunities and barriers to success. *Annals of Biomedical Engineering*, 44(9), 2626-2641. https://doi.org/10.1007/s10439-016-1691-6

Zahir, N., Sun, R., Gallahan, D., Gatenby, R.A., & Curtis, C. (2020). Characterizing the ecological and evolutionary dynamics of cancer. *Nature Genetics*, 52(8), 759-767. https://doi.org/10.1038/s41588-020-0668-4

Zangooei, M.H., & Habibi, J. (2017). Hybrid multiscale modeling and prediction of cancer cell behavior. *PloS One*, 12(8), e0183810. https://doi.org/10.1371/journal.pone.0183810

Zou, J., Huss, M., Abid, A., Mohammadi, P., Torkamani, A., & Telenti, A. (2019). A primer on deep learning in genomics. *Nature genetics*, 51(1), 12-18. https://doi.org/10.1038/s41588-018-0295-5