The mathematics of cancer: integrating quantitative models

Philipp M. Altrock[1,2], Lin L. Liu[1]* and Franziska Michor[1]*

Abstract | Mathematical modelling approaches have become increasingly abundant in cancer research. The complexity of cancer is well suited to quantitative approaches as it provides challenges and opportunities for new developments. In turn, mathematical modelling contributes to cancer research by helping to elucidate mechanisms and by providing quantitative predictions that can be validated. The recent expansion of quantitative models addresses many questions regarding tumour initiation, progression and metastases as well as intra-tumour heterogeneity, treatment responses and resistance. Mathematical models can complement experimental and clinical studies, but also challenge current paradigms, redefine our understanding of mechanisms driving tumorigenesis and shape future research in cancer biology.

Mathematical models
Models can describe a system by means of abstraction and mathematical formalism. They enable extrapolation beyond situations originally analysed, quantitative predictions, inferrence of mechanisms, falsification of underlying biological hypotheses and quantitative description of relationships between different components of a system.

[1]Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Department of Biostatistics, Harvard T.H. Chan School of Public Health, 450 Brookline Avenue, Boston, Massachusetts 02115, USA.
[2]Program for Evolutionary Dynamics, Harvard University, 1 Brattle Square, Suite 6, Cambridge, Massachusetts 02138, USA.
**These authors contributed equally to this work.*
Correspondence to F.M.
e-mail:
michor@jimmy.harvard.edu
doi:10.1038/nrc4029

Cancer is a genetic disease fuelled by somatic evolution[1,2]. During somatic evolution, genetic and epigenetic alterations can spread through a population of pre-malignant or cancer cells[3]. As cell populations accumulate progressively more changes over time[1,4], they acquire characteristics that enable them to persist within tissues[5]. These adaptations are characterized by, for example, increased evasion of the immune system[6–8], the selection pressures exerted by therapeutic interventions[9] and the formation of metastases[10,11]. Multiple factors contribute to carcinogenesis, such as stochastic DNA replication errors in cells, interactions between cells and the tissue microenvironment, and environmental exposures such as radiation and diet. Therefore, an understanding of cancer development and progression requires the elucidation of collective properties of cells within a tissue[3] and their interaction with the microenvironment[12,13].

Mathematical models have proved useful for deriving a detailed understanding of mechanisms and processes in cancer[14,15], and have been used to propose new experiments, suggest different treatment modalities and alter risk prognoses[16–27]. Quantitative descriptions of cancer-driving mechanisms at multiple length and timescales lead to new questions that can be addressed with novel experiments and mathematical models that integrate empirical evidence. Such models then systematically evaluate assumptions, investigate alternative mechanisms and make predictions that can be experimentally validated. The power of mathematical modelling lies in its ability to reveal previously unknown or counterintuitive physical principles that might have been overlooked or missed by a qualitative approach to biology. As such, mathematical modelling can test theories on quantitative grounds. At its best, modelling provides indispensable contributions to cancer research, making investigations quantitative and predictive, and hypotheses falsifiable.

In this Review we examine recent topics of importance to basic and clinical cancer research, including methodology to describe cancer at various scales. We begin with models that describe clonal evolution in tumour development and determine the temporal sequence of mutational events. We then discuss mathematical models that describe cancer across multiple scales, such as hybrid models that combine cellular dynamics and microenvironmental factors, followed by modelling of metastasis dynamics and immunotherapy. We close with an outlook on open problems that require quantitative investigation.

Cancer initiation and tissue hierarchy

The dynamics of mutation accumulation. Since the inception of mathematical modelling of cancer, its approaches have sought to explain age-specific incidence curves[28–30] and the dynamics of mutation acquisition[31]. Such approaches allow prediction of the risk of, for example, developing lung cancer, based on a patient's age and smoking history[32]. In this context, a powerful mathematical tool for the study of the probabilistic growth of cell populations is the branching process[33–36] (BOX 1). Branching processes are based on the assumption that cellular events (replication, mutation and death) do not influence each other[37], that is, a cell acts in the same

Hybrid models
A modelling approach that combines several modelling techniques in one. For example, a hybrid model that describes the tumour microenvironment in which stromal cells follow a continuous nonlinear description, whereas tumour cells obey a discrete stochastic process.

Branching process
A stochastic process model of cell division, mutation events and cell death that leads on average to an exponential increase or decrease in the total population size. The branching process is based on the assumption that each individual event occurs at the same rate, independently of, for example, the population size or composition, or the point in time. The branching process is a Markov process; that is, the probability of the next event happening depends only on the current state of the population, and not on its earlier history.

Passenger mutations
Genetic changes that have no obvious or direct effect on cell fitness or cancer development, and may occur and potentially vanish again during any stage of tissue development and homeostasis. According to some definitions, passengers might also be (slightly) deleterious.

Driver mutations
Genetic changes that are causally involved in cancer development, typically conferring a functional change and a somatic evolutionary advantage.

Epistatic interactions
Interactions that occur when the functional effect of one genetic alteration depends on the genetic background of the cell; that is, the state of one or more other genes.

Homeostasis
A property of a system in which variables are regulated so that internal conditions remain stable and relatively constant. An example is the constant tissue size of most organs in the absence of neoplasms.

way irrespective of whether it is alone or one of billions; this choice is made for mathematical simplicity and because of the difficulty of correctly parameterizing more detailed models. As such, at any time, each cell is fully described by cell-intrinsic proliferation, mutation and death rates.

Branching processes have been used to analyse the accumulation of passenger mutations and driver mutations during tumour growth. For instance, a recent contribution[38] predicted the number of cells harbouring a specific driver mutation. The model starts with a single cell harbouring one oncogenic mutation; its clone then accumulates further mutations during subsequent cell divisions. Each mutation reduces the death rate slightly, and thus a new clone with one additional mutation expands more quickly than previous clones. The intrinsic stochasticity of the branching process then generates variability in the times at which subsequent driver mutations arise. This model was later extended to include epistatic interactions between driver mutations, to explain why some lesions carry hallmark genetic changes but do not progress[39].

Selectively neutral passenger mutations may also arise in healthy tissues. To address whether these alterations affect tumorigenesis, a branching process model[40] was developed to study three phases of the somatic evolution of cancer. The first phase describes healthy tissue expansion during development (FIG. 1). In the second phase, the tissue size is constrained by mechanisms of homeostasis. The third phase starts with a single oncogenic driver mutation, but the cell population rapidly

Box 1 | **Branching and Moran processes**

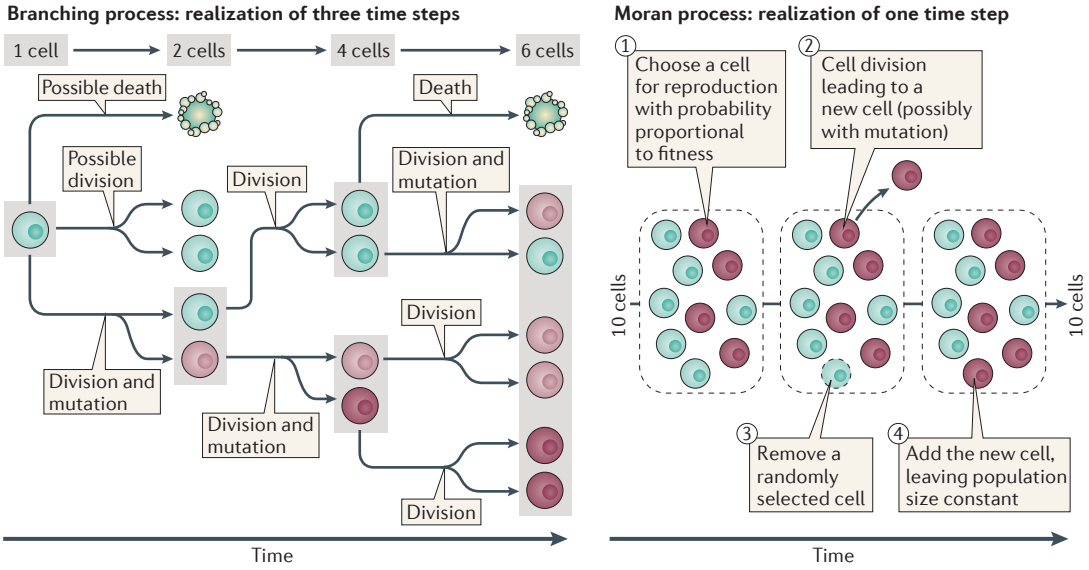
The branching process is commonly used for modelling cancer evolution[36,38,40,202] (see the figure, left panel). It is a Markov process in which every individual cell at time t produces a random number of offspring at a later time $t + \Delta t$. In discrete time, a number of events are possible for each cell: cell division, cell division with mutation, or death. Each event is characterized by a given rate, which is independent of population size and composition. As mutations accumulate in the cell population, each new cell type that emerges may have a new set of rates. Suppose at time t there exist $n_1(t)$ cells harbouring one mutation, whose birth and death rates per cell are λ_1 and μ_1, and $n_2(t)$ cells harbouring two mutations, whose birth and death rates per cell are λ_2 and μ_2. The mutation rate from the first to the second type is u. Then the transition probabilities (P) of those two cell types after a very short time interval Δt are:

$$P(n_1(t+\Delta t)=i+1, n_2(t+\Delta t)=j \mid n_1(t)=i, n_2(t)=j) \approx \lambda_1(1-u)i\Delta t$$
$$P(n_1(t+\Delta t)=i-1, n_2(t+\Delta t)=j \mid n_1(t)=i, n_2(t)=j) \approx \mu_1 i\Delta t$$
$$P(n_1(t+\Delta t)=i, n_2(t+\Delta t)=j+1 \mid n_1(t)=i, n_2(t)=j) \approx (\lambda_2+\lambda_1 u)j\Delta t$$
$$P(n_1(t+\Delta t)=i, n_2(t+\Delta t)=j-1 \mid n_1(t)=i, n_2(t)=j) \approx \mu_2 j\Delta t$$

The Moran process (see the figure, right panel) models stochastic dynamics in a population of constant size. There are n types of individual, $i = 1, 2, \ldots, n$. The numbers of individuals of each type are $N_1, N_2, \ldots, N_n$, which sum to N; this number is constant over time. The types can have different fitness values $f_1, f_2, \ldots, f_n$. During each time step (see the figure; right panel, events 1 to 4), an individual of type i is chosen to reproduce with a probability proportional to f_i, and subsequently, a random individual is chosen to die. This leads to the following probability that type i individuals increase and type j individuals decrease:

$$P(N_i \to N_i+1, N_j \to N_j-1) = \frac{N_i f_i}{N_i f_i + N_j f_j}\frac{N_j}{N}$$

The Moran process can also include random mutations, nonrandom death proportional to 'weakness' (or inverse fitness)[206] or time-dependent fitness[207].



acquires more drivers and passengers as it expands. This model predicts that the number of somatic mutations in tumours is positively correlated with patient age at diagnosis[40], because older tissues have had more time to accumulate alterations. Thus, half or more of the somatic mutations found in a tumour may arise before cancer initiation, and it might be possible to estimate the background mutation rate of a tissue from the number of mutations present in tumour samples of the same histology. This approach leads to the possibility of quantitative interpretation of the contributions of stochasticity and environmental factors to tumorigenesis[41,42], and suggests that cancer aetiology is predominantly the result of error accumulation during stochastic stem cell divisions, which is termed 'bad luck' tumorigenesis. The role of random variation in cancer development, and its consequences for epidemiology and cancer prevention, is a controversial topic[43–46] and requires further investigation[47].

To investigate the effects of slightly deleterious passenger mutations, a probabilistic individual-based model[48,49] was developed in which these mutations accumulate frequently but at random in a tumour with population size at equilibrium[48] (at 'carrying capacity'), interrupted by occasional waves of driver mutations. The driver events are assumed to be rare, but significantly increase the carrying capacity. The model predicts a critical population size below which tumour populations are more likely to go extinct than thrive owing to a gradual decrease in fitness, suggesting potential pharmacological interventions that may speed up the process of tumour extinction[48] by boosting passenger accumulation.

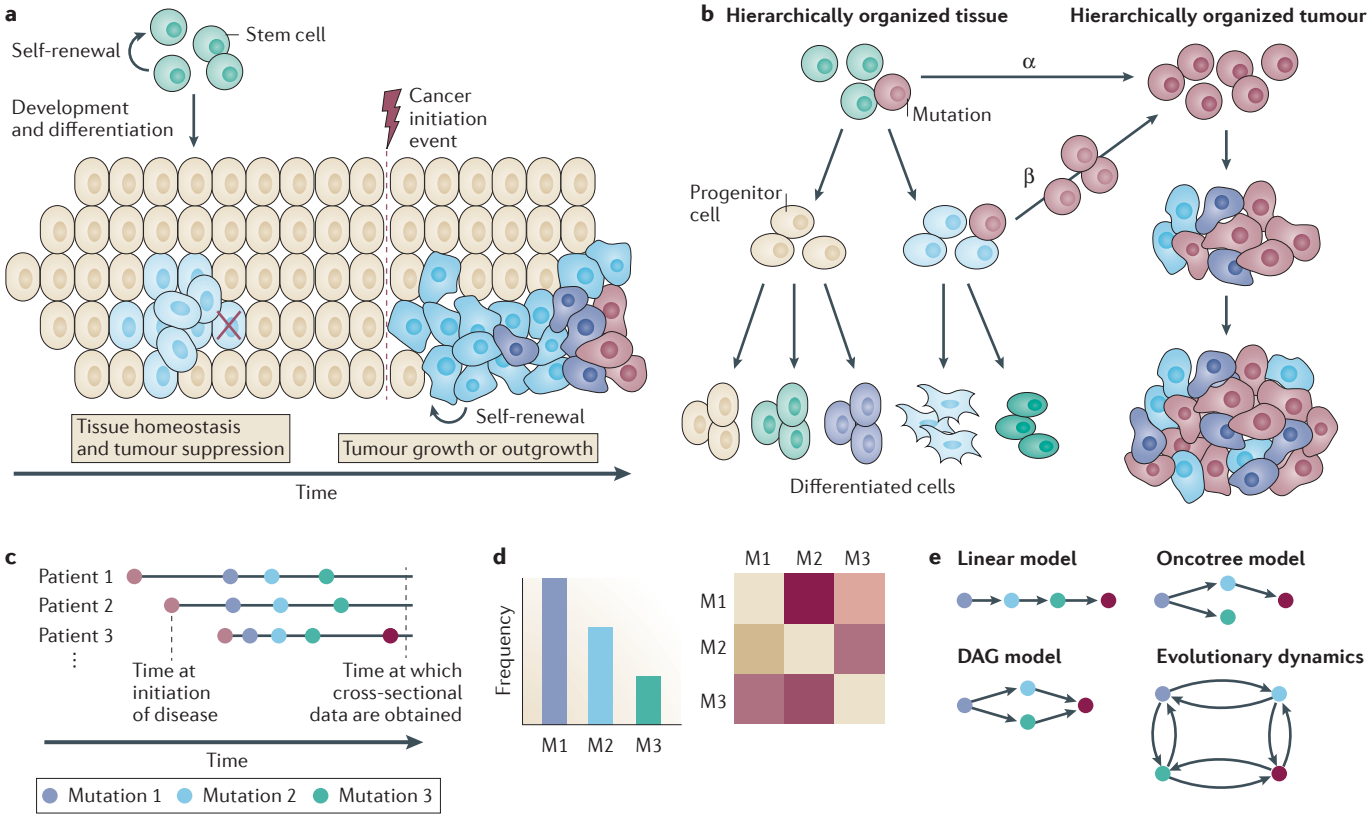


Figure 1 | **Tumour initiation and progression.** **a** | After conception, repeated proliferation, differentiation and selective death lead to the generation of individual tissues. Once development is completed, all tissues are characterized by homeostasis (constant cell numbers over time). Homeostasis breaks down when genetic and other alterations arise that enable individual cell clones to increase in frequency through the process of somatic evolution. **b** | Many healthy tissues are characterized by a hierarchical organization (left panel) encompassing stem cells with self-renewal capacity and low activity, transit-amplifying (also known as progenitor) cells and differentiated cells. Cancer may arise in multiple ways, either through mutation accumulation in the stem cell compartment or within a more differentiated cell population. The first pathway may lead directly to cancer (α) while the second option (β) may include additional mutational changes that enable self-renewal capabilities in the resulting tumour cell populations. Many tumours, like normal tissues, are also organized hierarchically, and in some cases dedifferentiation from mature tumour cells to undifferentiated tumour cells is possible. **c** | Cross-sectional genomic characterization of patient samples after diagnosis of the disease allows for the acquisition of data for many samples, albeit at only one time point per sample. The temporal sequence in which alterations (purple, blue, green and red circles) arise after tumour initiation (pink circles) in individual patients cannot be identified unless modelling approaches are applied. The dashed line indicates the time of sequencing of the patients' tumours and therefore one cannot observe potential future events (red circle). **d** | Such modelling approaches generally take frequencies of mutational events (M1–M3, left panel) and their co-occurrences in a patient cohort[84] into account (right panel; lighter colour indicates higher correlation). **e** | Several modelling approaches have been developed to derive likely temporal orders of events from such data, such as the linear model, the oncotree model, the directed acyclical graph (DAG) model and the evolutionary dynamics model.

Instead of considering a fixed fitness value conferred by each new mutation, another model[50] incorporated a fitness distribution such that a specific mutational event is assigned a randomly drawn fitness effect. The dynamics of mutation accumulation are then modelled using a fixed-size Moran process[51] (BOX 1) until a threshold population fitness is reached that allows disruption of normal homeostasis and clonal expansion (FIG. 1). In the model, the rate of escape from homeostasis 'competes' with the natural ageing of the patient population, as cancer incidence is measured conditional on the patient still being alive at that time. Considerations of both cancer initiation driven by genetic or epigenetic events and human survival processes are useful in understanding the impact of life expectancy and fitness distributions on cancer incidence[50]. The model predicts that a small number of highly advantageous mutations drive cancers with low incidence rates. This approach was later coupled with a branching process phase after tumour initiation. In this model, a cell with sufficiently high fitness expands exponentially while continuing to accumulate mutations whose fitness values are again determined probabilistically[52]. This approach enables the identification of driver and passenger events.

Branching processes are often exactly solvable and are used for their ability to describe clonal evolution in growing tumours. They rarely serve to describe spatial interactions between cells or include environmental factors. Recently, however, a sequence of branching processes[53], characterized by their distance from blood vessels, was used to hypothesize that local tumour microenvironments influence the selection pressure exerted on mutations that cause drug resistance. Indeed, this approach demonstrated that tumour composition and recurrence times depend on tumour-microenvironmental factors, such as nutrient or oxygen gradients[53]. The dynamics of mutation accumulation can also be studied using other stochastic models such as the Wright–Fisher process[54].

The cell of origin of human cancers. To understand the dynamics of mutation accumulation, it is important to describe not only the timing of alterations but the cell type in which they arise. A deterministic model can provide insight into how hierarchical tissue structures affect cancer risk and treatment effects[55–59]. Stochastic modelling, in contrast, has proved useful for determining whether a stem cell, transit-amplifying cell (also known as a progenitor cell) or terminally differentiated cell is more likely to serve as the cell of origin of a particular tumour type. Using such approaches for haematopoietic malignancies[60], a progenitor cell was found to be more likely to initiate tumorigenesis than a stem cell, as the large number of progenitor cells can make up for the need to accumulate a larger number of mutations. The probability of cancer initiation was found to be highest when progenitor cells first acquire an oncogenic mutation and then gain self-renewal capabilities[60].

Other mathematical models have addressed the cell of origin of brain cancers[61,62]. One model[61] described the dynamics of three compartments — stem cells, progenitor cells and differentiated cells — finding that a stem cell mutation is more likely to initiate brain cancer than a similar mutation in the early progenitor pool. Conversely, another model[62] found that progenitor cells are the likely cells of origin of glioblastoma if one of the cancer-initiating mutations imparts self-renewal. These approaches demonstrate that the differentiation hierarchy of a tissue is decisive for the dynamics of mutation accumulation and the kinetics of tumorigenesis (FIG. 1).

The temporal order of events

Our ability to interpret the importance of a specific mutation observed in cancer genomes[31] is hampered by the lack of knowledge of the temporal sequence in which these alterations occur during human tumorigenesis. This temporal order prioritizes the validation of potential drug targets, because early changes may cause a rewiring of the signalling circuitry or confer a state of addiction. A conserved temporal order of events in colorectal cancer was first proposed by Fearon and Vogelstein[63], based on the sequencing of patient samples at different stages of disease progression. This approach assumed a linear genetic model of mutation accumulation that postulates that there exists a single temporal sequence and that events are strictly sequential; simultaneous events are excluded. The order of events could be established through the identification of mutations that correlate with tumour size and stage[63–65], which proved successful in elucidating the dynamics of colorectal cancer. However, other cancer types[66–73] required the development of more complex models.

The oncotree model. The oncotree model is based on a probabilistic phylogenetic tree approach (FIG. 2). It relaxes the assumption of a strict sequential order of the linear genetic model and permits multiple paths to full transformation[74–76]. The temporal order of events is computed as a function of the distance of an event from the root node (that is, the time between initiation and the event). The relative position of each node on the oncotree is then constructed using co-occurrence frequencies of mutations across tumours. The oncotree methodology still has restrictions, as it imposes one single oncotree structure per data set. Therefore, mixture oncotree models were later introduced to combine multiple independent tree structures[77], and were applied to various cancer types, for example, nasopharyngeal carcinoma and oral cancer[78,79]. To overcome the limitation of tree-structured models that do not allow shared ancestors for multiple leaves, directed acyclic graphical models were developed. These models determine the order of somatic alterations from cross-sectional data sets[80–82] at the cost of a larger computational burden owing to increased model complexity. A possible solution to this problem is to decrease modelling resolution, and focus on pathway-level events instead of investigating individual mutations[80,83].

Evolutionary dynamics models. Population genetics approaches coupled with optimization algorithms can explicitly address evolutionary dynamics of cell populations that accumulate cancer-causing changes[54,84,85] (FIG. 2).

Deterministic model
Given a specific initial condition, a deterministic process always yields the same output, and no randomness is involved. Deterministic processes can be chaotic in that a small deviation in the initial condition may yield a large deviation after some time. However, this effect is different from the effect of a stochastic process in which the same initial condition can lead to different results.

Hierarchical tissue structures
Structures according to which most tissues are organized, ranging from slowly proliferating stem and progenitor cells to more quickly proliferating precursors and terminally differentiated cells.

Phylogenetic tree
A branching, tree-structured graph that represents the evolutionary relationships among different (mutational) stages of a tumour cell population, quantified by some measure of distance between individual cells or patient samples.

Graphical models
Mathematical structures that describes pairwise relations (called edges) between objects (called nodes), possibly on several layers. An acyclic graph does not have any cycles. Undirected graphs imply that there is no direction in the relationships along any edge. A tree-like graph has the property that every node can be traced back to a central node, called the root node, while final nodes of a tree are called leaves.

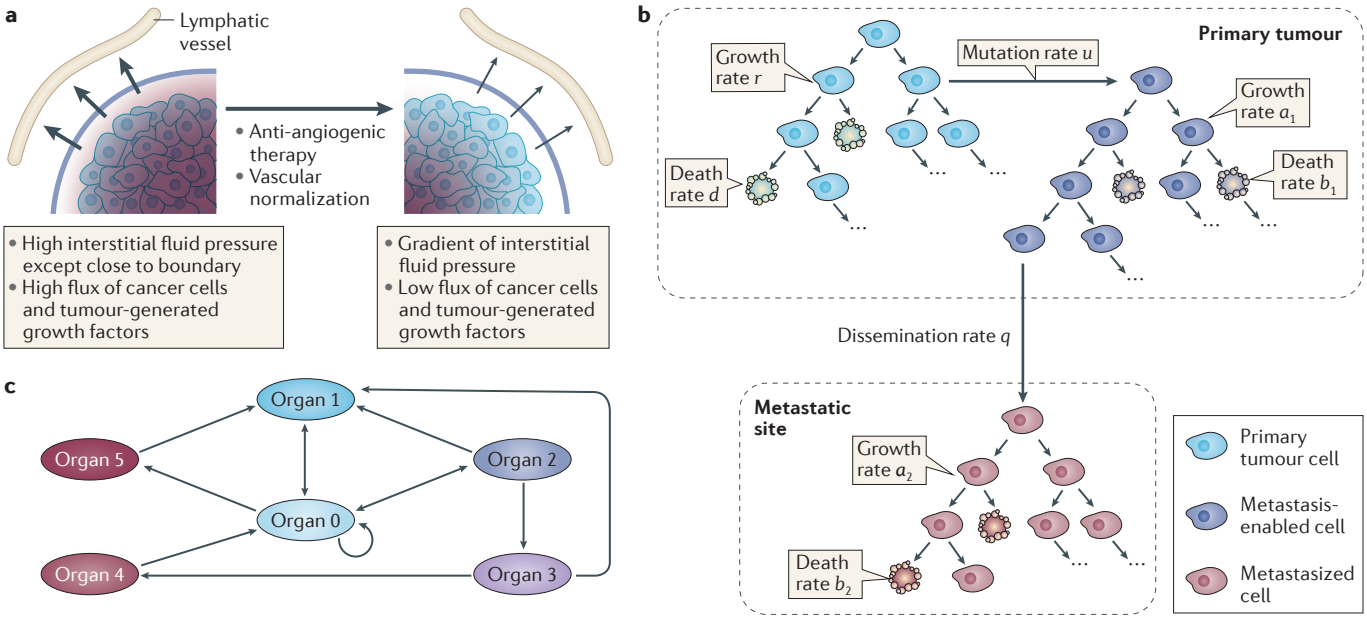


Figure 2 | **Microenvironmental interactions and the emergence of metastases.** **a** | Mathematical modelling predicts the effect of vascular normalization by anti-angiogenic therapy[177]. The model is based on deterministic nonlinear dynamical systems and requires parameterization regarding tumour radius, vascular hydraulic permeability, tissue hydraulic conductivity and vessel density. This approach showed that anti-angiogenic therapy could decrease the interstitial fluid pressure in the tumour by decreasing the tumour size and vascular hydraulic permeability, and potentially also the vascular density. This modelling approach revealed parameter ranges in which interstitial convection within the tumour increases but fluid convection outside decreases, which potentially improves treatment and limits the convection of metastatic cells into lymphatic vessels. **b** | A stochastic model of metastasis formation considers three cell types: primary tumour cells, which have not yet evolved the ability to metastasize and reside in the primary tumour; metastasis-enabled cells, which have evolved the ability to metastasize but still reside in the primary tumour, where they proliferate, die and potentially disseminate to a new metastatic site; and metastasized cells (metastasis-enabled cells become these, once disseminated), which may proliferate and die at different rates[130]. This framework can be used to determine quantities such as the risk of metastatic disease at diagnosis. **c** | A probabilistic model of metastatic spread among organs describes organs as nodes in a graph[171,172]. In this hypothetical example of probabilistic spread between organs, some metastases spread disease back to the primary site. The central node represents the organ harbouring the primary tumour. Arrows represent the probabilistic rates of spread between organs, estimated from patient autopsy data. These do not correspond to actual physical flow, but to the flow of the probability of disease progression. Depending on whether organs are net sources or net sinks for metastatic cells, they are classified as 'spreaders' or 'sponges'. This mathematical method revealed that metastasis formation is a stochastic multistep process. Part **b** is adapted with permission from REF. 169, Elsevier.

Such approaches are based on the Moran process[51,86] (BOX 1), which can be used to calculate transitions between different mutational states. The model parameters are then identified by matching model predictions with empirical observations, using mutational co-occurrence matrices. This methodology was applied to data from patients with glioblastoma, leukaemia or colorectal cancer[84,85]. The analyses found that most cancer types are characterized by multiple evolutionary trajectories that lead to the fully transformed state[84], which suggests a large extent of heterogeneity in the temporal order of cancerous events.

The problem of identifying the order of events is more straightforward if longitudinal data are available. For instance, data on somatic copy number alterations (SCNAs) from longitudinal case–control cohort samples of patients with Barrett's oesophagus can be used to determine the order of genetic alterations[87]. These investigations determined that SCNA enrichment is stable in the non-progressive stage, but more and more diverse SCNAs arise in late-stage cancers. Furthermore, an agent-based simulation model that considers individual cell death and birth events[88] demonstrates that the temporal order inferred from cross-sectional data might be inconsistent with the true mutational order in a cell line and that a phylogenetic tree obtained from a small number of cells within each tumour would better identify tumour genealogy. This inconsistency may be due to tumour heterogeneity and suggests that multiple intra-tumour samples should be used when inferring the order of events[88]. Such approaches have been used to analyse single-cell-based SCNAs[89] and point mutations[90]. Furthermore, exome sequencing of matched normal and tumour samples from patients[91] was used to determine the order of loss of heterozygosity and SCNAs, leading to several follow-up methods[92,93]. Similar approaches can also be used to investigate the timing of the development of advanced tumour stages and metastasis[21].

More recently, an analysis of approximately 350 individual glands from 15 colorectal tumours led to the 'big bang' model of tumour evolution, in which a

Longitudinal data
Repeated observations of the same system or set of systems over time.

Agent-based simulation
A computational approach that models complex systems consisting of interacting discretized items or 'agents'. In cancer modelling, these agents often represent cells, which can mutate into other types, divide into two cells, die or move in space. These simulations can be implemented according to either probabilistic or deterministic laws.

predominant single expansion of tumour cells arises together with early subclonal mixing[94]. Whole-genome SCNA profiles of individual glands helped to identify two phenomena that support the 'big bang' in colorectal cancer: first, the majority of the tumours harboured alterations that existed in all samples, suggesting a monoclonal origin. Second, a large fraction of the carcinomas contained the same SCNA in glands from different locations of the tumour, suggesting that it arose early. These data further revealed similarly high tumour heterogeneity within glands and between glands, rendering selective sweeps unlikely. This study exemplifies how large-scale multi-region genomic data together with evolutionary modelling can lead to novel insights into tumour development.

Modelling the tumour microenvironment

So far, we have discussed approaches that investigate cancer development and somatic evolution, largely without considering the microenvironment. However, cancer progresses as a result of the collective dynamics that emerge from interactions between tumour cells and their microenvironment[95–98]. Mathematical modelling of microenvironmental interactions often requires complex model considerations. For instance, even when considering avascular tumour expansion as uniform spherical growth, temporal and spatial dimensions govern the dynamics, and models based on partial differential equations (BOX 2) are used. In addition to modelling tumour and normal cell populations and their genetic changes, such models require the incorporation of biophysical and environmental properties — namely diffusion of growth factors, hormones, nutrients and oxygen — that affect tumour proliferation and invasion patterns.

Modelling tumour dynamics in one and two dimensions. An early approach[99] improved upon the uniform spherical model by exploiting a non-uniform

Box 2 | **Ordinary and partial differential equations**

Systems that are deterministic (exactly or approximately) can be described by ordinary differential equations (ODEs). Their main characteristic is that they have one independent variable. For dynamic systems, the independent variable is time. Dependent variables can be the volume of a tumour, the fraction of a genetic alteration in a population or the chance of finding a receptor in a certain state at a certain time. ODEs can describe systems of few and many dimensions, and allow chaotic and complex behaviour.

Consider the growth curve of a cell population with a time-dependent growth rate. Let the population size at time t be $x(t)$ and the growth rate decay exponentially as $a(t) = \alpha e^{-bt}$. The ODE for $x(t)$ can be written as:

$$\frac{dx(t)}{dt} = \alpha e^{-bt} x(t)$$

and the solution is:

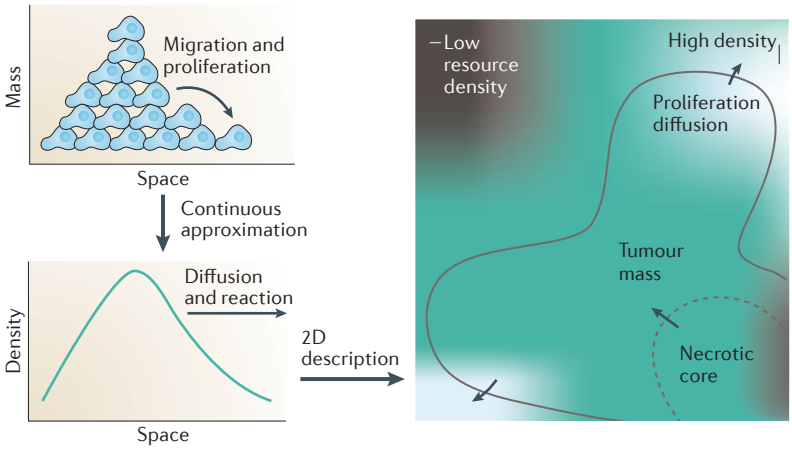
$$x(t) = X_0 e^{\frac{a}{b}(1-e^{-bt})}$$

if the initial size of the population is X_0. This is a special form of sigmoidal growth.

For dynamic systems in which the quantities of interest — such as the concentration of oxygen — depend on more than one independent variable (for example, time and space), partial differential equations (PDEs) are used. This is beneficial especially when descriptions in higher dimensions are needed. For example, the concentration of oxygen in a tissue at time t in position x (for example, the distance to the centre of a blood vessel) can be denoted by $c(x,t)$. Oxygen diffuses with a diffusion constant D, naturally decays with a constant a, is consumed by tumour cells of group size $N(x,t)$ at position x at time t and at a rate g, and is produced by a macromolecule in the extracellular matrix at rate b. Here all rates are independent of time and position. The PDE is then given by:

$$\frac{\partial c(x,t)}{\partial t} = D\frac{\partial^2 c(x,t)}{\partial x^2} - ac(x,t) - g\,N(x,t)c(x,t) + bc(x,t)$$

Such reaction–diffusion equations are studied in many hybrid and multi-scale models[12,105]; their stationary distribution in 2D is exemplified in the figure (right panel). Typically the oxygen concentration is only one component of a system of PDEs, on which the behaviour of tumour cell density depends. Solutions to PDEs are continuous functions, and the figure shows the reasoning behind the approximation of a mass of tumour cells by a continuous density function.



Stochastic process
This describes how a random variable (or set of random variables) changes over time and/or space. A stochastic process ascribes a probability to each event and allows for the prediction of the probability of a certain outcome. In contrast to a deterministic process, the initial condition yields an entire probability distribution over possible events at any later point in time.

Markov process
A memoryless stochastic process in which the conditional probability distribution over all future events depends only on the present state. A Markov chain explicitly addresses stochastic dynamics between discrete states in discrete time, thus allowing for a full characterization using a transition matrix in which the entries describe the probability of transitioning from one state to another.

Biased random walk
The movements of an object or changes in a variable that on average follow a specific pattern or trend.

diffusion of growth factors as a function of tissue permeability. Tumour growth was modelled in one-dimensional space measured by the distance of a cell from the tumour centre[100]. These approaches were later extended to two dimensions, for example, by considering avascular tumour growth on a 2D lattice coupled to a capillary vessel[101]. In this approach, nutrients diffuse deterministically, but cell dynamics obey a stochastic process. Simulation results of this process were consistent with experimental observations of narrow cell layer dynamics and suggest that the competition for nutrients between tumour cells and normal cells is a leading mechanism for the generation of a commonly observed finger-like tumour surface morphology. Early 2D models were also used to study angiogenesis[102], and recapitulated angiogenic features such as growth of capillary sprouts, vascular branching and loop formation[103].

From cells to tumour morphology. Microenvironmental multiscale models usually distinguish between discrete (individual cell-based) and continuum (cell population-based) models. Typically, a continuum deterministic model governs the dynamics of the nutrients, chemical factors and extracellular matrix (ECM). Discrete stochastic models describe cell growth, migration and interactions by defining the probabilistic reaction rates of the respective events in the form of a spatial Markov process (BOX 3). Hybrid approaches bridge several scales and complement fully continuous, highly complex descriptions of tumour dynamics[104].

To address the dichotomy between discrete cellular dynamics and continuous external fields in the microenvironment, detailed mechanistic mathematical models have been developed[105,106]. An early hybrid-modelling approach[107] used a continuum model for the interaction of an endothelial tumour cell population with the ECM. Vessel movement within the tumour was modelled as a biased random walk in space, determined by angiogenic factors and adhesive forces. Vessel networks resulting from angiogenesis were predicted to change owing to the effects of blood flow[108], particularly wall shear stress, which was thus identified as a potential treatment target. A context-adaptive strategy[109] used a continuum model in regions of high cell density and a discrete approach elsewhere. A phenotypic mutation-driven approach[110] then showed that different cell interactions are important at different stages of tumour progression: cell–cell interactions may dominate early while cell–ECM interactions become the dominant determinants of tumour morphology later on. Building on this approach, it was shown[12] that a selectively 'mild' microenvironment can lead to coexisting phenotypes and tumours that are less likely to invade adjacent tissue, whereas a selectively harsh microenvironment can lead to strong selection for a smaller number of more invasive cell phenotypes. Hybrid modelling of such situations arising in bone-metastatic prostate cancer[111] recently incorporated several differentiation levels of osteoblasts and prostate tumour cells as well as tumour-secreted factors. This approach led to the discovery that phasic activity and osteoblast behaviour help us to better understand the treatment response of bone metastases[111].

Box 3 | **Hybrid models in cancer**

Hybrid models combine spatial reaction–diffusion (such as growth–consumption dynamics on a continuous scale) with discrete cellular dynamics (for example, on a lattice), which describe cell growth and motility (see the figure). For instance, in a continuous description of tumour cell density, the number of cells N (in space) is given by the PDE:

$$\frac{\partial N}{\partial t} = D_N \frac{\partial^2}{\partial x^2} N - \frac{\partial}{\partial x}\left(N \frac{\partial}{\partial x} E\right)$$

Extracellular matrix (ECM) cells (represented by E), matrix degrading enzymes (MDE, represented by M) and oxygen concentration (represented by C) obey the dynamics:

$$\frac{\partial E}{\partial t} = -\phi ME$$

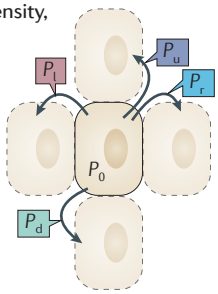
$$\frac{\partial M}{\partial t} = D_M \frac{\partial^2}{\partial x^2} M + \lambda N - \mu M$$

$$\frac{\partial C}{\partial t} = D_C \frac{\partial^2}{\partial x^2} C + fE - nN - cC$$

The diffusion constants are denoted by D_N, D_M and D_C. ECM is degraded proportionally to MDE density, which itself is produced by tumour cells at rate λ and decays neutrally at rate μ. The oxygen decay rate is given by c, oxygen uptake by tumour cells occurs at rate n and oxygen is brought in at rate f, proportional to the ECM density[12,110].

In a hybrid model, the tumour cell population is described as a discrete agent-based system that lives on a lattice or off-lattice[103,208]. On a 2D lattice, stochastic rules for each cell apply.

In the figure, P_u is the probability of moving up; P_d is the probability of moving down; P_r is the probability of moving right; P_l is the probability of moving left and P_0 is the probability of not moving. These probabilities apply to each individual cell and may depend on the state of the environment. In a hybrid model, these movement probabilities (the values of P_x) depend on the concentrations of the continuous variables C, E, M and N. In addition to movement rules, hybrid models also include rules for cell-specific processes such as proliferation, apoptosis and mutation that can be dependent on the environment. Figure reproduced with permission from REF. 12, Elsevier.



Hybrid models are among the most sophisticated approaches in mathematical cancer modelling. Nonetheless, they are still simplified caricatures of tumour growth in that they often do not consider the 3D structure of tissues[112] and other aspects. Recently, a series of investigations[113–115] coupled 3D discrete models to continuum nonlinear tumour growth models, and revealed clonal evolution patterns in 3D[116]. These approaches point towards new methods of drug transport to overcome irregular tumour morphologies and clonal distributions, and also reveal intricate connections between tumour type and cellular phenotypes. Another recent innovation in the study of cellular phenotypes and morphology has been the use of evolutionary game theory (BOX 4).

The glycolytic phenotype. An important aspect of tumour cell behaviour in relation to the microenvironment is the emergence of different metabolic states. The Warburg effect[117] refers to increased and sustained rates of glycolysis in the presence of oxygen in the microenvironment. In the literature on modelling, this phenomenon was first approached as a reaction–diffusion continuous model[118], which yielded a description of interactions between the neoplasm and the surrounding normal stroma during acidic conditions induced by glycolytic tumour cells. In this model, the extent of acidity depends on the density of tumour cells and acid reabsorption rate, which predicts that increased acid levels result in more aggressive phenotypes and an acellular gap at the tumour–stromal interface, and this was experimentally and clinically validated. Later, a 2D hybrid model[119] was used to study whether a few tumour cells can change the microenvironment by producing excess acid, which would facilitate their own growth and negatively affect normal cells. In this context, cell division and cell death are contingent on acidity and on the status of their neighbourhood on a 2D lattice, which demonstrates that even small numbers of cancer cells suffice to change the local microenvironment in favour of tumour development.

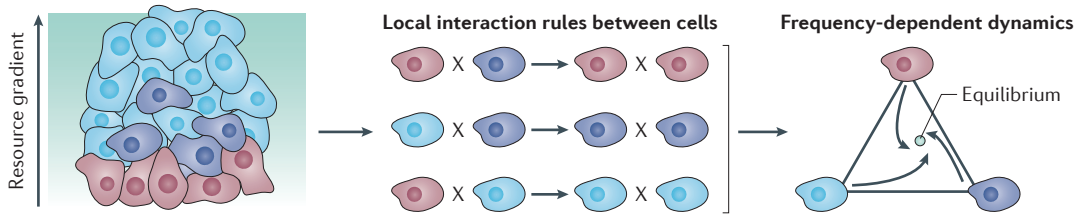
A subsequent hybrid model[120] described individual cellular behaviour with a microenvironmental response network. The network determined the phenotype of the cell based on the local environment, which ensured that cells can behave differently depending on their context in the simulation. The model predicted that the glycolytic phenotype tends to emerge from regions with low oxygen and high ECM density.

Modelling the brain tumour microenvironment. Brain tumours have been particularly well studied in terms of the interplay between the tumour and the microenvironment. Reaction–diffusion modelling[121] — based on the BrainWeb[122] database and patient-specific CT imaging and/or MRI[123]— predicted that glioma growth can be modelled as a travelling wave with a diameter that increases linearly with time. This led to reliable tumour size prediction during the course of disease progression[123] and allowed the extrapolation of tumour dynamics under therapy[121,124]. This work led to a 'therapeutic response index', which is associated with prognosis. Hence, mathematical models can be used as an *in silico* control sample that allows causal reasoning. Another approach[125] further predicted dependencies between survival time, glioma proliferation rates and invasion dynamics,

Box 4 | **Evolutionary game theory in cancer biology**

Evolutionary game theory (EGT) is a subfield in game theory that studies the long-term proliferation of the players in the context of a game. The game is typically formulated as a table that ascribes fitness values (pay-offs) to every pairwise interaction between cell phenotypes (strategies). Based on these interactions, EGT is able to mathematically describe the changes in the relative abundance of phenotypes (frequency-dependent selection), identify equilibria in which phenotypes may coexist and make statements about the stability of equilibria. EGT traditionally describes populations with a predefined set of phenotypes. Only recently has a flexible number of phenotypes[209] or a population that changes in size[210] been considered.

Recent developments show that EGT is an exciting field for mathematical modelling of cancer[131,132,211–214] (see the figure). An early model noted that the exchange of diffusible goods between cells can be captured by an evolutionary game in which cooperator cells provide a biochemical public good that protects aggressive tumour cells[214]. The idea of cooperation between groups of cells[215] led to the realization that the evolutionary game in cancer resembles a public goods game with nonlinear returns (fitness increases quickly with low numbers of cooperators, but returns diminish for high numbers). A recent study[216] introduced an evolutionary game in prostate cancer, describing tumour–stroma interactions, which could explain the three predominant tumour types: clinically indolent, rapidly growing–highly malignant, or environment- and stromal-support-dependent. The idea of 'intratumoural symbiosis' was adapted[217] to confirm the existence of multiple stable states of coexistence between hypoxic and oxygenated cells driven by competition for glucose and lactate. Knowledge of the relative abundance of coexisting cells, and their effect on the behaviour of the tumour[218], will hopefully serve as a stepping stone to a deeper connection between cellular diversity measures and clinical applications.



Linear-quadratic model
A prominent heuristic to describe cell survival under radiation. The number of surviving cells after a certain dose of radiation has been administered takes the form of an exponential function with a linear and a quadratic term in its argument.

Predator–prey models
These models, also known as Lotka–Volterra dynamics, are used to describe the dynamics of ecological species, or types, as a nonlinear deterministic process. They were originally used to describe population dynamics of predators and prey, taking into account abundance, interactions, and population growth and diminution. They can also be used to describe mutualistic and competitive evolutionary dynamics; for example, between cellular types.

thus providing a potentially clinically useful tool for diagnosis and prognosis. A similar model[126] used serial diffusion-weighted MRI to infer diffusion coefficients and proliferation rates of treated and untreated murine brain tumours, which suggested a personalized approach based on parameters derived from individual MRI data[127,128].

Clonal interactions, migration and invasion. Cancer cell populations can thrive on synergistic interactions[129–131]. However, it is largely unknown whether clonal diversity is established probabilistically or deterministically. Furthermore, it is unclear how such interactions drive tumour complexity and subclonal coexistence based on biophysical principles[132]. To quantitatively address the potentially complex interactions between several subclones, an experimental animal model of heterogeneous breast tumours was recently developed[133]. A patient-derived cell line was engineered to express secreted factors, each in a separate subclone. These subclones were then competed against each other and against the parental line *in vivo*, and the resulting monoclonal and polyclonal tumours were analysed for tumour phenotype, size and clonal composition at multiple time points. Only a few clones led to significantly different tumour phenotypes, and only cells secreting interleukin-11 (IL-11) were found to promote growth of the entire tumour population[133]. Interestingly, IL-11-driven tumours did not display increased relative sizes of the IL-11 subclones. An iterative, deterministic modelling approach confirmed that independent growth of different clones could not explain total tumour growth rates, but a linear interaction term proportional to IL-11 frequency could. This finding demonstrated that non-cell-autonomous driving of tumour growth can maintain subclonal diversity.

A cell phenotype model[106] devised an alternative to the hypothesis of mutation-driven tumour invasion. Using glioblastoma as an example, the model demonstrated that mutation accumulation alone could not explain the time to recurrence after surgical removal of the tumour; rather, oxygen levels were found to determine the existence and proliferation of different subclones in a complex way. Thus, oxygen gradients can influence cellular phenotype and the microenvironment determines the rules of cell migration. The model predicted a maximal tumour size after a fixed time that depends on the nature of the invasive phenotype.

A cellular automaton model with hierarchical tissue structure[134] related the discrete cellular states of proliferation, death, migration, ageing and quiescence driven by response to the microenvironment to explain tumour invasiveness. The model predicted that under low migration rates (low rates of occupation of the neighbouring lattice in an agent-based simulation), a single cancer cell generates only a small tumour clone. High migration rates led to seeding of new clones at sites outside the expansion radius of older clones. In the context of this model, regions with low cell density owing to low proliferation capacity or high death rate allow for more rapid expansion, which leads to accelerated tumour growth[134].

Unstable tumour morphology and radiotherapy. Unstable tumour morphology in the form of physical abnormalities and instabilities can have important surgical and therapeutic consequences. To investigate the dynamics of tumour response, taking microvasculature and fluid dynamics into consideration has become increasingly important for simulation models[135]. Detailed descriptions of fluid flow through a tumour with leaky vasculature[136] are helpful for the development of quantitative methods to assess drug delivery and pharmacodynamics[137]. Vascular formation and oxygen deficiency affect normal cells and cancerous cells differently; increased production of key proteins can render cancer cells less affected by hypoxia, and mathematical models are useful for identifying the best approaches for treatments based on these considerations[138]. Similarly, spatial reaction–diffusion models are useful for investigating the efficacy of radiotherapy[139] and survival thereafter[140]. These models make assumptions: for example, that radiotherapy induces cell death following a linear-quadratic model[141], which enables the characterization of growth and invasion under spatial and temporal heterogeneity of the administered dose, and hence determination of the response to radiotherapy[139]. Mathematical models can then be used to predict patient-specific responses to treatment, if they are parameterized with patient data collected at diagnosis[140]. Furthermore, mathematical models of the extent and characteristics of intratumour heterogeneity can be used to identify optimum radiation administration schedules[22].

Interactions between the immune system and tumour cells. Modulating the interplay between immune system components and cancer cells via immunotherapy is one of the most promising goals of modern cancer treatment approaches[142–145]. Using such therapies, immune system cells are enticed to specifically attack tumour cells without harming healthy tissue[146,147]. This field has relied largely on deterministic mathematical models, which exploited the analogy between tumour cell–effector immune cell interactions and ecological predator–prey models[148]. Mathematical modelling approaches to understanding the dynamics of immunotherapy are based on understanding tumour–immune system interactions on multiple scales[149]. At the tissue level, a continuum approach can be used to describe cell densities with a focus on local frequencies of immune cells in the tumour vicinity. At the cellular level, simplified models consider tumour and immune cells in the microenvironment, where cytokines regulate their interactions and dormancy patterns[149]. Another study[150] modelled the competition between tumour and immune system cells formulated as biochemical reactions among tumour cells, antibodies and effector T cells. Simulations showed that the effect of antibody treatment can be either beneficial or detrimental to tumours owing to deleterious effects of the tumour–antibody complex on effector T cells, and recommended caution when artificially manipulating antibodies as cancer therapy[150].

Several recent deterministic modelling approaches[151,152] considered parameters estimated from mouse and human data. For instance, one approach[152] suggested that evaluating patient-specific $CD8^+$ T cell-mediated cytotoxicity could determine the likelihood of beneficial effects of immunotherapy. The dynamics of immunotherapy can also be modelled to identify potentially improved treatment strategies. Another investigation[153,154] considered that immune cells kill tumour cells, the progression of which in turn suppresses the immune system. Various immunotherapy schedules were then studied (for example, constant and periodic schedules), with the goal of identifying better strategies. The model showed that the ability of periodic immunotherapies to remove the tumour depended on its aggressiveness. Another promising avenue of treatment is based on cytokine IL-21 family inhibition, which has stimulatory effects on boosting both the innate and adaptive immune responses[155]. One approach[156] used data from tumour-bearing mice treated with IL-21 to derive a model that was parameterized using data on the effect of IL-21 on natural killer (NK) cells, or $CD8^+$ T cell-mediated tumour lysis. The effect of IL-21 on NK cells was modelled using linear functions in a dose-dependent manner. The effect on $CD8^+$ T cells was incorporated into the T cell memory dynamics in ordinary differential equation (ODE) form. The model showed that using IL-21 in melanoma could result in beneficial outcomes, but not in fibrosarcoma owing to its higher immunogenicity.

A more elaborate agent-based simulation was later developed for melanoma immunotherapy[157]; this approach considered a larger number of cell types of the tumour–immune system. Although these *in silico* simulations led to predictions that were consistent with experimental data, questions remain of whether such modelling approaches can be parameterized adequately to capture the full extent of this complex biological system. This realization points to a common caveat to early mathematical modelling approaches to new systems; future research can only build on quantitative approaches if the findings are robust and the important aspects of biology are known, included and correctly parameterized.

Mathematical modelling of metastases

Metastases are the ultimate reason for cancer mortality[10,158] and are associated with a rapid decrease in treatment prospects[159]. The stochastic nature of metastasis evolution has long been recognized[10]. Early mathematical descriptions of the metastatic process were concerned with competition between cancerous and healthy target tissue[160] and metastatic recurrence[161]. More recent quantitative approaches have tried to shed light on the development and predictors of metastases, such as by modelling dormancy or cell kinetics[162]. Quantitative insights into the mechanisms of metastasis formation and ways to prevent it are relatively sparse, potentially owing to limited availability of data.

Diversity and metastatic efficiency. The diversity between individual cells within a metastasis and between metastases and the primary tumour is important for understanding the mechanisms of the metastatic process, and sheds light on the likelihood of treatment success. To this end, a combined experimental and mathematical modelling approach[163] compared genotypes and phenotypes of single cells from distant metastatic tumours, lymph node lesions and primary tumours derived from patients with breast cancer. These authors found that the difference in intralesion diversity was largest when comparing primary tumours and matched lymph node metastases, and not between two distant metastases, suggesting that the most important bottleneck in metastasis evolution occurs when cells leave the primary site and establish lymph node colonies.

The variety observed in metastatic disease is influenced not only by physical factors but also by heterogeneity in circulating tumour cells (CTCs)[164]. Quantitative modelling of metastases has also incorporated mechanical considerations and blood flow patterns, leading to a 'metastatic efficiency index'[165] that relates lung cancer metastasis incidence in a target organ to blood flow patterns between organs. This work was later extended[166,167] to explicitly describe relative venous and arterial flow between organ pairs, making the approach a more realistic representation of the patient. The aim was to develop a model to test the relative likelihood of tumour self-seeding by either primary site repopulation by CTCs, or secondary seeding when CTCs form metastases and then return to the primary site[166]. This model predicted that secondary seeding may be a common phenomenon, which is important for our understanding of the rapid evolution of resistance that emerges in a sanctuary[168] and migrates back to the tumour.

Markov models of metastasis dynamics. A stochastic branching process model[169,170] (BOX 1) can be used to consider tumour cell growth and death, including mutation accumulation and dissemination events (FIG. 1). This model predicts the probability of metastasis formation before tumour diagnosis, as well as the number and size distribution of metastases, depending on the size of the main lesion. In one study, the model was parameterized using data from patients with pancreatic cancer to identify better therapeutic interventions; it was found that early growth rate reduction may be more effective than tumour resection for maximum patient survival[169].

The network of all organs can be described as a bidirected graph in which nodes represent organs that may receive metastatic cells from any primary tumour or other metastses, and weighted edges describe the probability flow between nodes. Initially, the probability of cancer cells being present is one in the primary site and zero elsewhere. A stochastic Markov chain model can then estimate the transition rates between organs[171], which was learned by iterative updating and by comparing the model predictions with metastatic distributions from large autopsy data sets of untreated deceased patients with primary tumours and metastases. This approach yielded a probability distribution across nodes, by which some nodes have a higher

probability of receiving metastases than others, allowing for the inference of the most likely primary sites, CTC dissemination and back-seeding patterns[171]; that is, the identification of 'spreaders' and 'sponges' of the metastatic process[172]. The location of the primary site and timing of the first metastatic dissemination can determine the course of the disease, guide clinical intervention[173] and inform about lymphatic or haematogenous spread[174].

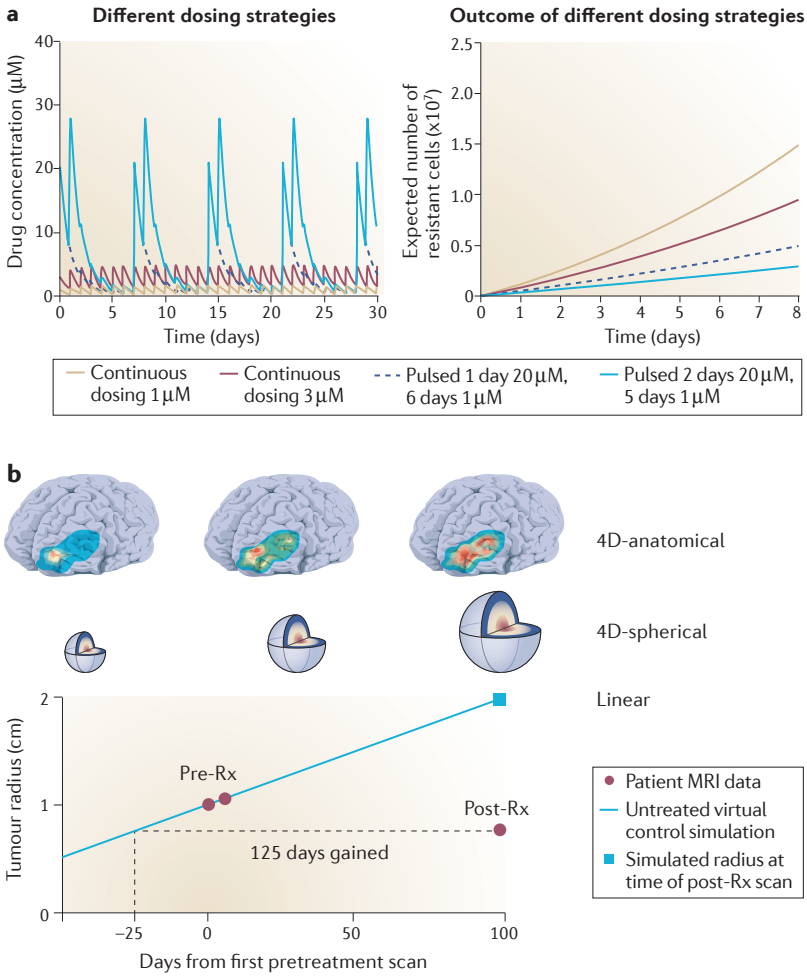


Figure 3 | **Treatment response and dosing strategies.** **a** | Mathematical modelling can be used to identify optimum treatment strategies to delay the emergence of resistance to targeted agents[188]. The left panel displays alternative dosing strategies for the epidermal growth factor receptor (EGFR) inhibitor erlotinib in the treatment of EGFR-mutant non-small-cell lung cancer (concentration over time); the right panel shows the predicted number of resistant cells over time for each of those dosing schedules. **b** | Mathematical modelling of glioma growth and treatment response[191] can be used to evaluate patient-specific responses to therapeutic interventions. Using this approach, 4D-anatomical, 4D-spherical and linear models can simulate glioblastoma growth with different levels of modelling complexity. The metric called 'Days Gained' is defined as time between the post-treatment (post-Rx) measurement and the predicted time to having the same tumour radius if the patient had not been treated. The latter is inferred via modelling and simulation from pretreatment (pre-Rx) measurement. Day 0 in the figure represents the day of the first pretreatment MRI scan. Day −25 represents the day when the tumour radius would be the same as that at the post-treatment MRI scan if the patient had not been treated. Part **a** is reproduced from Chmielecki, J. *et al.* Optimization of dosing for EGFR-mutant non-small cell lung cancer with evolutionary cancer modeling. *Sci. Transl. Med* **3**, 90ra59 (2011). Reprinted with permission from AAAS (REF. 188). Part **b** is reproduced from REF. 191.

Modelling biophysical properties. Several approaches to the modelling of metastases have also incorporated biophysical properties of the ECM. For instance, a recent approach[175] related the cellular force generation in a tumour to the stiffness and geometry of the ECM. This mathematical approach was parameterized using an experimental model that allowed alteration of ECM stiffness. When contractile and protrusive forces are in balance with the amount of drag in the medium, the velocity of a cell can be calculated, and this enabled prediction of when cells may leave a lesion and revealed an optimal velocity for tumour cells through the ECM at intermediate stiffness.

Physical and topological properties also have a role in both drug delivery and tumour staging[176]. The properties of lymphatic patterns led to the observation that changes induced by angiogenic factors affect the potential entry of cancer cells into the lymphatic system[177]. The physical effects of anti-angiogenic agents can be modelled by complex fluid mechanics, which showed that decreased convection of growth factors could limit the extent of angiogenesis in sentinel lymph nodes, and thus reduce lymphatic metastases. Clinical studies confirmed that the use of anti-angiogenic agents normalizes tumour vasculature and increases drug delivery and immune cell abundance at the tumour site[178].

Treatment response and resistance

The 1970s and 1980s brought forth several seminal contributions towards mathematical modelling of cancer treatment responses[179–182]. Norton and Simon[182] studied the growth kinetics of tumours during chemotherapy, and found that the tumour follows a sigmoidal growth curve. They predicted that a dense dosing regime would be superior to the standard approaches of the time, a prediction that was later clinically validated[183]. Goldie and Coldman[179,180] proposed using stochastic processes[37], taking into account probabilistic cell dynamics, to study pre-existing or acquired resistance. Treatment effects were implemented as a reduction in the number of sensitive tumour cells, and the probability of resistance during a regimen of two drugs administered sequentially was shown to depend on the tumour size and mutation rate. Based on their model predictions, the authors suggested that drugs should be given as early as possible and in alternating schedules[184]. These predictions were tested in a trial later on but failed to demonstrate better outcomes in the arm with alternating rather than sequential treatment[185]. These approaches have inspired several groups to investigate optimum administration schedules for various situations[186]. Some approaches[187] were based on branching processes to model sensitive and resistant cancer cell dynamics under general treatment schedules. Such models can be coupled with pharmacokinetic approaches to identify drug administration schedules that maximally delay the emergence of resistance[188,189] (FIG. 3a). ODE-based investigations have also been successful at modelling tumour resistance to treatment[190].

Other approaches have focused on the responses to radiotherapy using the linear-quadratic model. For instance, a recent approach[22] used a mathematical model

for glioblastoma treatment response that considered both differentiation from resistant stem-like glioblastoma cells to sensitive differentiated cells and dedifferentiation in the opposite direction. The identified strategy suggests that enrichment in the resistant stem cell population could prolong survival by increasing the time to recurrence. This strategy was also validated with a randomized mouse trial, which showed a significantly improved survival distribution in the optimized schedule group[22].

Direct mathematical modelling of the prognostic effects of radiotherapy in patients has also been developed recently[191] (FIG. 3b). This approach led to a novel metric called 'Days Gained', which is defined as the difference in time between the post-treatment MRI scan and the predicted time at which the same tumour radius would be reached had the patient not been treated. The latter is estimated based on the pretreatment MRI scan and subsequent computer simulations. The Days Gained score was found to be significantly associated with patient survival. Similarly, a patient-specific optimized radiation strategy[192] was developed, which connected computer simulations with patient-specific parameters.

The idea of estimating patient-specific parameters was also used in modelling androgen ablation therapy for prostate cancer[193], using an ODE-based model that incorporated the dynamics of normal epithelial cells, androgen-dependent cancer cells and cells that are resistant to first-line anti-testosterone therapy. The key parameter in the model was the competitive advantage of androgen-dependent cells; an increase in this advantage predicted benefits of intermittent scheduling of anti-androgen therapy, as compared with continuous scheduling.

A different approach[194] established the idea of 'adaptive therapy', and suggested the maintenance of a stable tumour burden in which treatment-resistant cells are suppressed by treatment-sensitive cells. A key assumption is that the density of cancer cells without treatment follows a power-law growth model, and that after treatment, time-dependent growth constraints govern the density of resistant cells[195]. The adaptive therapy strategy is then given by the optimum between the maximally tolerated dose and the dose administered in the previous period, multiplied by the ratio of tumour size change between these two periods, with the dose increasing as the tumour expands.

The risk of pre-existing resistance. To derive the probability of accumulating resistant cells in exponentially expanding populations, the Luria–Delbrück model[196,197] is widely used. This approach assumes that both sensitive and resistant cells grow exponentially, and that sensitive cells can generate resistant cells during cell division. A two-type birth–death process[196] (BOX 1) can be used to calculate the probability of pre-existing resistance and the expected number of resistant cells before diagnosis. The latter was found to be independent of the mutation rate if mutations are rare, but to increase with the tumour size at detection. The probability of pre-existing resistance increases in proportion to both detection size and mutation rate. This approach has been generalized to two[198] and to many mutations[199]. An extension of this model[200] allowed for the calculation of the probability of acquiring resistance during therapy. This model considered that, before treatment, drug-sensitive cells are the fittest, whereas during treatment, resistant cells become the most fit. It allows for the calculation of probabilities of 'successful intervention' (that is, tumour extinction) and generalization to n mutations necessary to cause resistance. Treatment success thus depends on the number of cancer cells, the fitness landscape (a map from genotype to reproductive ability) and mutation rates[200]. Similar modelling approaches were developed to explicitly consider cell cycle dynamics[201], and were used to show that, when two or more drugs are used, the probability of pre-existing resistance increases with the rate of quiescence. Finally, several approaches also aimed to determine the number of resistant clones[202,203], and suggested that different mutations might be present at subclonal frequency at the time of treatment initiation.

Mutations that confer resistance to multiple treatments were recently explicitly considered[24,204] in the context of probabilistic resistance evolution under combination therapies. In the simplest setting, one mutation can only lead to resistance to one drug. A model predicted that, for combinations of two or more drugs, resistance is most likely to arise before treatment, and that the effect of additional drugs is diminished as the mutation rate increases[24]. Such models can also be used to study ibrutinib resistance dynamics in chronic lymphocytic leukaemia (CLL)[205]. Using clinical and preclinical data enabled the tuning of a birth–death process (BOX 1), and this also suggested that resistant clones are probably present before treatment albeit at low numbers. This approach inferred a slight fitness advantage of resistant cells over sensitive CLL cells even in the absence of treatment. Another recent contribution[20] similarly determined the probability of resistance before inhibition of epidermal growth factor receptor (EGFR) with panitumumab, and showed that cancer cells probably harbour resistance mutations before detection and treatment. Hence, combination therapy targeting at least two different pathways would be advised. A model[18] in which one mutation confers resistance to multiple drugs (cross-resistance) in turn allowed for the calculation of the probability of resistance before combination therapy. Parameters were estimated using 20 patients with melanoma who were receiving the BRAF inhibitor vemurafenib. Modelling approaches have also been designed to investigate the dynamics of treatment response to combination therapies such as in the case of EGFR inhibitors combined with chemotherapy[203].

Conclusions

Although remarkable progress has been made towards the quantitative description of cancer progression, treatment dynamics and resistance, several important questions remain. For example, quantitative approaches that help to explain the treatment response to immunotherapies, and help to identify patients who

Sigmoidal growth curve
An S-shaped growth pattern in which the population size starts from a low density with positive acceleration, then transitions to negative acceleration at high density. An equilibrium population size can be characterized, for instance, by a proliferation–self renewal–death balance, or by a carrying capacity. Examples are Gompertzian growth and logistic growth.

Power-law growth model
A functional relationship between two quantities (for example, time and tumour size), where one quantity varies with the power (that is, exponent) of the other. The exponent can typically be inferred from linear regression analysis of a doubly logarithmic transformation of the data.

Luria–Delbrück model
The Luria-Delbrück experiment investigated whether mutations occur independently from, or owing to, selection. Data from growth experiments in which *Escherichia coli* were challenged with a virus were compared to a stochastic process model used to calculate the probability of having a certain number of resistant mutants. The findings suggested that mutations occurred randomly over time and were not a response to selection.

may respond to such agents or who will experience potentially fatal side effects, are desperately needed. Furthermore, it remains unclear how best to combine a specific immunotherapy agent with chemotherapies, targeted therapies or radiation. Similar questions apply for other combination treatment modalities, in which both the time–concentration profile and most synergistic combination partners need to be identified. Also, quantitative approaches to metastases are still relatively descriptive, mostly owing to a lack of detailed data from multiple sites in many patients. Indeed, as the ability of a model to predict system dynamics depends on its parameterization, it is essential to obtain accurate parameter estimates from *in vitro* and/or *in vivo* model systems or clinical trials. There is still a paucity of such data and closer collaboration between mathematical modellers and experimental as well as clinical researchers is needed.

When a mathematical model is developed, it is often unclear how much detail it should contain. It needs to be sufficiently complex to describe the phenomenon of interest but still enable correct parameterization. The more biological aspects a model includes, the more parameters need to be measured so that it makes accurate predictions. Parameterization is often difficult given the available experimental model systems. Here, mathematical modelling has the potential to point to gaps in our efforts to comprehensively characterize important features of tumours: for example, tumour genomes, cell behaviour, and microenvironmental properties, as well as cancer epidemiological factors. Such quantitative insights might in turn suggest new analyses and modelling avenues, and potentially create novel mathematics or statistics. Furthermore, it is essential that a mathematical model renders predictions that are robust and experimentally verifiable, such that its assumptions and output can be tested and, if necessary, iteratively improved.

Despite these caveats and limitations, we believe that the mathematical modelling field is instrumental in aiding our understanding of cancer biology and treatment. It will become ever more important, as more quantitative measurements become available from preclinical models and patient samples. These approaches will be invaluable for testing hypotheses, falsifying theories and identifying biological mechanisms.

1. Weinberg, R. A. *The Biology of Cancer* (Garland Science, 2013).
2. Vogelstein, B. & Kinzler, K. W. *The Genetic Basis of Human Cancer* (McGraw–Hill, 1998).
3. Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* **100**, 57–70 (2000).
4. Armitage, P. & Doll, R. The age distribution of cancer and a multi-stage theory of carcinogenesis. *Br. J. Cancer* **8**, 1–12 (1954).
5. Nowell, P. C. The clonal evolution of tumor cell populations. *Science* **194**, 23–28 (1976).
6. Garraway, L. A. & Lander, E. S. Lessons from the cancer genome. *Cell* **153**, 17–37 (2013).
7. Dunn, G. P., Bruce, A. T., Ikeda, H., Old, L. J. & Schreiber, R. D. Cancer immunoediting: from immunosurveillance to tumor escape. *Nat. Immunol.* **3**, 991–998 (2002).
8. Pages, F. *et al.* Effector memory T cells, early metastasis, and survival in colorectal cancer. *N. Engl. J. Med.* **353**, 2654–2666 (2005).
9. Allan, J. M. & Travis, L. B. Mechanisms of therapy-related carcinogenesis. *Nat. Rev. Cancer* **5**, 943–955 (2005).
10. Nguyen, D. X., Bos, P. D. & Massague, J. Metastasis: from dissemination to organ-specific colonization. *Nat. Rev. Cancer* **9**, 274–284 (2009).
11. Kim, M. Y. *et al.* Tumor self-seeding by circulating cancer cells. *Cell* **139**, 1315–1326 (2009).
12. Anderson, A. R., Weaver, A. M., Cummings, P. T. & Quaranta, V. Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment. *Cell* **127**, 905–915 (2006).
A multiscale mathematical model that attempts to describe cancer evolution and the dynamics of the microenvironment, and shows that both genetic changes and environmental changes can impact cancer invasiveness.
13. Vaupel, P., Kallinowski, F. & Okunieff, P. Blood flow, oxygen and nutrient supply, and metabolic microenvironment of human tumors: a review. *Cancer Res.* **49**, 6449–6465 (1989).
14. Anderson, A. R. & Quaranta, V. Integrative mathematical oncology. *Nat. Rev. Cancer* **8**, 227–234 (2008).
15. Byrne, H. M. Dissecting cancer through mathematics: from the cell to the animal model. *Nat. Rev. Cancer* **10**, 221–230 (2010).
16. Knudson, A. G. Jr Mutation and cancer: statistical study of retinoblastoma. *Proc. Natl Acad. Sci. USA* **68**, 820–823 (1971).
17. Cavenee, W. K. *et al.* Expression of recessive alleles by chromosomal mechanisms in retinoblastoma. *Nature* **305**, 779–784 (1983).
18. Bozic, I. *et al.* Evolutionary dynamics of cancer in response to targeted combination therapy. *eLife* **2**, e00747 (2013).
A stochastic evolutionary model that identifies probabilities of evolution of resistance to combination therapy.
19. Lenaerts, T., Pacheco, J. M., Traulsen, A. & Dingli, D. Tyrosine kinase inhibitor therapy can cure chronic myeloid leukemia without hitting leukemic stem cells. *Haematologica* **95**, 900–907 (2010).
20. Diaz, L. A. Jr *et al.* The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. *Nature* **486**, 537–540 (2012).
21. Jones, S. *et al.* Comparative lesion sequencing provides insights into tumor evolution. *Proc. Natl Acad. Sci. USA* **105**, 4283–4288 (2008).
22. Leder, K. *et al.* Mathematical modeling of PDGF-driven glioblastoma reveals optimized radiation dosing schedules. *Cell* **156**, 603–616 (2014).
A mathematical model and optimization approach to identify better radiation scheduling in glioblastoma that led to survival improvement in a mouse trial.
23. Sherratt, J. A. & Nowak, M. A. Oncogenes, anti-oncogenes and the immune response to cancer: a mathematical model. *Proc. Biol. Sci.* **248**, 261–271 (1992).
24. Komarova, N. L. & Wodarz, D. Drug resistance in cancer: principles of emergence and prevention. *Proc. Natl Acad. Sci. USA* **102**, 9714–9719 (2005).
A mathematical modelling contribution towards the understanding of resistance that existed prior to chemotherapy and targeted combination therapy.
25. Sanga, S. *et al.* Mathematical modeling of cancer progression and response to chemotherapy. *Expert Rev. Anticancer Ther.* **6**, 1361–1376 (2006).
26. Swanson, K. R., Alvord, E. C. Jr & Murray, J. D. Virtual brain tumours (gliomas) enhance the reality of medical imaging and highlight inadequacies of current therapy. *Br. J. Cancer* **86**, 14–18 (2002).
A hallmark paper that shows how a 'virtual tumour' allows data-driven therapy improvements using mathematical modelling.
27. Frank, S. A., Iwasa, Y. & Nowak, M. A. Patterns of cell division and the risk of cancer. *Genetics* **163**, 1527–1532 (2003).
28. Armitage, P. & Doll, R. A two-stage theory of carcinogenesis in relation to the age distribution of human cancer. *Br. J. Cancer* **11**, 161–169 (1957).
29. Nordling, C. O. A new theory on cancer-inducing mechanism. *Br. J. Cancer* **7**, 68–72 (1953).
30. Fisher, J. & Hollomon, J. A hypothesis for the origin of cancer foci. *Cancer* **4**, 916–918 (1951).
31. Stratton, M. R., Campbell, P. J. & Futreal, P. A. The cancer genome. *Nature* **458**, 719–724 (2009).
32. Foy, M., Spitz, M. R., Kimmel, M. & Gorlova, O. Y. A smoking-based carcinogenesis model for lung cancer risk prediction. *Int. J. Cancer* **129**, 1907–1913 (2011).
33. Kimmel, M. & Axelrod, D. *Branching Processes in Biology. Interdisciplinary Applied Mathematics* (Springer, 2015).
34. Haccou, P. *Branching Processes: Variation, Growth, and Extinction of Populations* (Cambridge Univ. Press, 2005).
35. Durrett, R. *Branching Process Models of Cancer* (Springer, 2015).
36. Antal, T. & Krapivsky, P. L. Exact solution of a two-type branching process: models of tumor progression. *J. Statist. Mechan.-Theory Exper.* (2011).
37. Parzen, E. *Stochastic processes* (SIAM, 1999).
38. Bozic, I. *et al.* Accumulation of driver and passenger mutations during tumor progression. *Proc. Natl Acad. Sci. USA* **107**, 18545–18550 (2010).
A mathematical modelling contribution that studies the accumulation of driver and passenger mutations in cancer.
39. Bauer, B., Siebert, R. & Traulsen, A. Cancer initiation with epistatic interactions between driver and passenger mutations. *J. Theor. Biol.* **358**, 52–60 (2014).
40. Tomasetti, C., Vogelstein, B. & Parmigiani, G. Half or more of the somatic mutations in cancers of self-renewing tissues originate prior to tumor initiation. *Proc. Natl Acad. Sci. USA* **110**, 1999–2004 (2013).
A mathematical model of development and tumorigenesis in which driver and passenger mutations accumulate.
41. Tomasetti, C. & Vogelstein, B. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science* **347**, 78–81 (2015).
42. Wodarz, D. & Zauber, A. G. Cancer: risk factors and random chances. *Nature* **517**, 563–564 (2015).
43. Ashford, N. A. *et al.* Cancer risk: role of environment. *Science* **347**, 727 (2015).
44. O'Callaghan, M. Cancer risk: accuracy of literature. *Science* **347**, 729 (2015).
45. Potter, J. D. & Prentice, R. L. Cancer risk: tumors excluded. *Science* **347**, 727 (2015).
46. Tomasetti, C. & Vogelstein, B. Cancer risk: role of environment-response. *Science* **347**, 729–731 (2015).
47. Noble, R., Kaltz, O. & Hochberg, M. E. Peto's paradox and human cancers. *Phil. Trans. R. Soc. B* **370**, 11 (2015).

48. McFarland, C. D., Korolev, K. S., Kryukov, G. V., Sunyaev, S. R. & Mirny, L. A. Impact of deleterious passenger mutations on cancer progression. *Proc. Natl Acad. Sci. USA* **110**, 2910–2915 (2013).
A mathematical model that includes slightly deleterious passenger mutations that accumulate between sweeps caused by oncogenic driver mutations.
49. McFarland, C. D., Mirny, L. A. & Korolev, K. S. Tug-of-war between driver and passenger mutations in cancer and other adaptive processes. *Proc. Natl Acad. Sci. USA* **111**, 15138–15143 (2014).
50. Foo, J., Leder, K. & Michor, F. Stochastic dynamics of cancer initiation. *Phys. Biol.* **8**, 015002 (2011).
51. Moran, P. A. P. Random processes in genetics. *Math. Proc. Cambridge Phil. Soc.* **54**, 60–71 (1958).
52. Foo, J. *et al.* An evolutionary approach for identifying driver mutations in colorectal cancer. *PLoS Comput. Biol.* **11**, e1004350 (2015).
53. Mumenthaler, S. M. *et al.* The impact of microenvironmental heterogeneity on the evolution of drug resistance in cancer cells. *Cancer Inform.* **14**, 19–31 (2015).
54. Beerenwinkel, N. *et al.* Genetic progression and the waiting time to cancer. *PLoS Comput. Biol.* **3**, e225 (2007).
A mathematical model that allows calculation of the expected waiting time to cancer using a Wright–Fisher process and tumour mutation data.
55. Werner, B., Dingli, D. & Traulsen, A. A deterministic model for the occurrence and dynamics of multiple mutations in hierarchically organized tissues. *J. R. Soc. Interface* **10**, 20130349 (2013).
56. Werner, B., Dingli, D., Lenaerts, T., Pacheco, J. M. & Traulsen, A. Dynamics of mutant cells in hierarchical organized tissues. *PLoS Comput. Biol.* **7**, e1002290 (2011).
57. Weekes, S. L. *et al.* A multicompartment mathematical model of cancer stem cell-driven tumor growth dynamics. *Bull. Math. Biol.* **76**, 1762–1782 (2014).
58. Michor, F., Nowak, M. A., Frank, S. A. & Iwasa, Y. Stochastic elimination of cancer cells. *Proc. Biol. Sci.* **270**, 2017–2024 (2003).
59. Roeder, I. *et al.* Dynamic modeling of imatinib-treated chronic myeloid leukemia: functional insights and clinical implications. *Nat. Med.* **12**, 1181–1184 (2006).
60. Haeno, H., Levine, R. L., Gilliland, D. G. & Michor, F. A progenitor cell origin of myeloid malignancies. *Proc. Natl Acad. Sci. USA* **106**, 16616–16621 (2009).
Mathematical modelling of the different evolutionary pathways leading to cancer, which calculates the likelihood of individual cell types serving as the cell of origin.
61. Ganguly, R. & Puri, I. K. Mathematical model for the cancer stem cell hypothesis. *Cell Prolif.* **39**, 3–14 (2006).
62. Hambardzumyan, D., Cheng, Y. K., Haeno, H., Holland, E. C. & Michor, F. The probable cell of origin of NF1- and PDGF-driven glioblastomas. *PLoS ONE* **6**, e24454 (2011).
63. Fearon, E. R. & Vogelstein, B. A genetic model for colorectal tumorigenesis. *Cell* **61**, 759–767 (1990).
64. Vogelstein, B. & Kinzler, K. W. *The Genetic Basis of Human Cancer* (McGraw, 2002).
65. Vogelstein, B. *et al.* Genetic alterations during colorectal-tumor development. *N. Engl. J. Med.* **319**, 525–532 (1988).
66. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330–337 (2012).
67. Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609–615 (2011).
68. Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **489**, 519–525 (2012).
69. Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult *de novo* acute myeloid leukemia. *N. Engl. J. Med.* **368**, 2059–2074 (2013).
70. Forbes, S. A. *et al.* COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* **39**, D945–D950 (2011).
71. International Cancer Genome Consortium *et al.* International network of cancer genome projects. *Nature* **464**, 993–998 (2010).
72. Sjoblom, T. *et al.* The consensus coding sequences of human breast and colorectal cancers. *Science* **314**, 268–274 (2006).
73. Stransky, N. *et al.* The mutational landscape of head and neck squamous cell carcinoma. *Science* **333**, 1157–1160 (2011).
74. Desper, R. *et al.* Inferring tree models for oncogenesis from comparative genome hybridization data. *J. Comput. Biol.* **6**, 37–51 (1999).
75. Desper, R. *et al.* Distance-based reconstruction of tree models for oncogenesis. *J. Comput. Biol.* **7**, 789–803 (2000).
76. Hoglund, M., Frigyesi, A., Sall, T., Gisselsson, D. & Mitelman, F. Statistical behavior of complex cancer karyotypes. *Genes Chromosomes Cancer* **42**, 327–341 (2005).
77. Beerenwinkel, N. *et al.* Learning multiple evolutionary pathways from cross-sectional data. *J. Comput. Biol.* **12**, 584–598 (2005).
The introduction of a mixture model of multiple evolutionary trees that allows detailed characterization of mutations leading to cancer.
78. Huang, Z. *et al.* Construction of tree models for pathogenesis of nasopharyngeal carcinoma. *Genes Chromosomes Cancer* **40**, 307–315 (2004).
79. Pathare, S., Schaffer, A. A., Beerenwinkel, N. & Mahimkar, M. Construction of oncogenetic tree models reveals multiple pathways of oral cancer progression. *Int. J. Cancer* **124**, 2864–2871 (2009).
80. Hjelm, M., Hoglund, M. & Lagergren, J. New probabilistic network models and algorithms for oncogenesis. *J. Comput. Biol.* **13**, 853–865 (2006).
81. Gerstung, M., Baudis, M., Moch, H. & Beerenwinkel, N. Quantifying cancer progression with conjunctive Bayesian networks. *Bioinformatics* **25**, 2809–2815 (2009).
82. Simon, R. *et al.* Chromosome abnormalities in ovarian adenocarcinoma: III. Using breakpoint data to infer and test mathematical models for oncogenesis. *Genes Chromosomes Cancer* **28**, 106–120 (2000).
83. Gerstung, M., Eriksson, N., Lin, J., Vogelstein, B. & Beerenwinkel, N. The temporal order of genetic and pathway alterations in tumorigenesis. *PLoS ONE* **6**, e27136 (2011).
84. Attolini, C. S. *et al.* A mathematical framework to determine the temporal sequence of somatic genetic events in cancer. *Proc. Natl Acad. Sci. USA* **107**, 17604–17609 (2010).
The introduction of an evolutionary modelling approach allowing the identification of the order of mutations fuelling tumour development.
85. Cheng, Y. K. *et al.* A mathematical methodology for determining the temporal order of pathway alterations arising during gliomagenesis. *PLoS Comput. Biol.* **8**, e1002337 (2012).
86. Antal, T. & Scheuring, I. Fixation of strategies for an evolutionary game in finite populations. *Bull. Math. Biol.* **68**, 1923–1944 (2006).
87. Li, X. *et al.* Temporal and spatial evolution of somatic chromosomal alterations: a case-cohort study of Barrett's esophagus. *Cancer Prev. Res. (Phila)* **7**, 114–127 (2014).
88. Sprouffske, K., Pepper, J. W. & Maley, C. C. Accurate reconstruction of the temporal order of mutations in neoplastic progression. *Cancer Prev. Res. (Phila)* **4**, 1135–1144 (2011).
89. Wang, Y. *et al.* Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* **512**, 155–160 (2014).
90. Martins, F. C. *et al.* Evolutionary pathways in BRCA1-associated breast tumors. *Cancer Discov.* **2**, 503–511 (2012).
91. Durinck, S. *et al.* Temporal dissection of tumorigenesis in primary cancers. *Cancer Discov.* **1**, 137–143 (2011).
92. Greenman, C. D. *et al.* Estimation of rearrangement phylogeny for cancer genomes. *Genome Res.* **22**, 346–361 (2012).
93. Purdom, E. *et al.* Methods and challenges in timing chromosomal abnormalities within cancer samples. *Bioinformatics* **29**, 3113–3120 (2013).
94. Sottoriva, A. *et al.* A Big Bang model of human colorectal tumor growth. *Nat. Genet.* **47**, 209–216 (2015).
95. Fidler, I. J. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *Nat. Rev. Cancer* **3**, 453–458 (2003).
96. Liotta, L. A. & Kohn, E. C. The microenvironment of the tumour–host interface. *Nature* **411**, 375–379 (2001).
97. Mueller, M. M. & Fusenig, N. E. Friends or foes – bipolar effects of the tumour stroma in cancer. *Nat. Rev. Cancer* **4**, 839–849 (2004).
98. Paget, S. The distribution of secondary growths in cancer of the breast. 1889. *Cancer Metastasis Rev.* **8**, 98–101 (1989).
99. Byrne, H. M. & Chaplain, M. Growth of necrotic tumors in the presence and absence of inhibitors. *Math. Biosci.* **135**, 187–216 (1996).
100. Sherratt, J. A. & Chaplain, M. A. A new mathematical model for avascular tumour growth. *J. Math. Biol.* **43**, 291–312 (2001).
101. Ferreira, S. C. Jr., Martins, M. L. & Vilela, M. J. Reaction-diffusion model for the growth of avascular tumor. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **65**, 021907 (2002).
102. Orme, M. E. & Chaplain, M. A. Two-dimensional models of tumour angiogenesis and anti-angiogenesis strategies. *IMA J. Math. Appl. Med. Biol.* **14**, 189–205 (1997).
103. Stokes, C. L. & Lauffenburger, D. A. Analysis of the roles of microvessel endothelial cell random motility and chemotaxis in angiogenesis. *J. Theor. Biol.* **152**, 377–403 (1991).
104. Macklin, P. *et al.* Multiscale modelling and nonlinear simulation of vascular tumour growth. *J. Math. Biol.* **58**, 765–798 (2009).
105. Rejniak, K. A. & Anderson, A. R. Hybrid models of tumor growth. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **3**, 115–125 (2011).
106. Hatzikirou, H., Basanta, D., Simon, M., Schaller, K. & Deutsch, A. 'Go or grow': the key to the emergence of invasion in tumour progression? *Math. Med. Biol.* **29**, 49–65 (2012).
107. Anderson, A. R. & Chaplain, M. A. Continuous and discrete mathematical models of tumor-induced angiogenesis. *Bull. Math. Biol.* **60**, 857–899 (1998).
108. McDougall, S. R., Anderson, A. R. & Chaplain, M. A. Mathematical modelling of dynamic adaptive tumour-induced angiogenesis: clinical implications and therapeutic targeting strategies. *J. Theor. Biol.* **241**, 564–589 (2006).
109. Kim, Y., Stolarska, M. A. & Othmer, H. G. A hybrid model for tumor spheroid growth *in vitro* I: theoretical development and early results. *Math. Models Methods Appl. Sci.* **17**, 1773–1798 (2007).
110. Anderson, A. R. A hybrid mathematical model of solid tumour invasion: the importance of cell adhesion. *Math. Med. Biol.* **22**, 163–186 (2005).
One of the first papers to use multiscale hybrid models in cancer.
111. Araujo, A., Cook, L. M., Lynch, C. C. & Basanta, D. An integrated computational model of the bone microenvironment in bone-metastatic prostate cancer. *Cancer Res.* **74**, 2391–2401 (2014).
112. Li, X. *et al.* Nonlinear three-dimensional simulation of solid tumor growth. *Discrete Continuous Dyn. Sys. Ser. B* **7**, 581–604 (2007).
113. Frieboes, H. B. *et al.* Three-dimensional multispecies nonlinear tumor growth-II: tumor invasion and angiogenesis. *J. Theor. Biol.* **264**, 1254–1278 (2010).
114. Wise, S. M., Lowengrub, J. S., Frieboes, H. B. & Cristini, V. Three-dimensional multispecies nonlinear tumor growth—I Model and numerical method. *J. Theor. Biol.* **253**, 524–543 (2008).
115. Frieboes, H. B. *et al.* Prediction of drug response in breast cancer using integrative experimental/computational modeling. *Cancer Res.* **69**, 4484–4492 (2009).
116. Waclaw, B. *et al.* A spatial model predicts that dispersal and cell turnover limit intratumour heterogeneity. *Nature* (2015).
A model that combines spatially explicit evolutionary processes with dynamics of mutation accumulation.
117. Warburg, O., Wind, F. & Negelein, E. The metabolism of tumors in the body. *J. Gen. Physiol.* **8**, 519–530 (1927).
118. Gatenby, R. A. & Gawlinski, E. T. A reaction-diffusion model of cancer invasion. *Cancer Res.* **56**, 5745–5753 (1996).
119. Patel, A. A., Gawlinski, E. T., Lemieux, S. K. & Gatenby, R. A. A cellular automaton model of early tumor growth and invasion. *J. Theor. Biol.* **213**, 315–331 (2001).
120. Gerlee, P. & Anderson, A. R. A hybrid cellular automaton model of clonal evolution in cancer: the emergence of the glycolytic phenotype. *J. Theor. Biol.* **250**, 705–722 (2008).
121. Swanson, K. R., Bridge, C., Murray, J. & Alvord Jr, E. C. Virtual and real brain tumors: using mathematical modeling to quantify glioma growth and invasion. *J. Neurol. Sci.* **216**, 1–10 (2003).

122. Cocosco, C. A., Kollokian, V., Kwan, R. K.-S., Pike, G. B. & Evans, A. C. BrainWeb: online interface to a 3D MRI simulated brain database. *NeuroImage* **5**, S425 (1997).
123. Harpold, H. L., Alvord, E. C. Jr & Swanson, K. R. The evolution of mathematical modeling of glioma proliferation and invasion. *J. Neuropathol. Exp. Neurol.* **66**, 1–9 (2007).
124. Wang, C. H. *et al.* Prognostic significance of growth kinetics in newly diagnosed glioblastomas revealed by combining serial imaging with a novel biomathematical model. *Cancer Res.* **69**, 9133–9140 (2009).
125. Swanson, K. R. *et al.* Quantifying the role of angiogenesis in malignant progression of gliomas: *in silico* modeling integrates imaging and histology. *Cancer Res.* **71**, 7366–7375 (2011).
126. Atuegwu, N. C. *et al.* Incorporation of diffusion-weighted magnetic resonance imaging data into a simple mathematical model of tumor growth. *Phys. Med. Biol.* **57**, 225–240 (2012).
127. Yankeelov, T. E. *et al.* Clinically relevant modeling of tumor growth and treatment response. *Sci. Transl Med.* **5**, 187ps9 (2013).
128. Yankeelov, T. E., Quaranta, V., Evans, K. J. & Rericha, E. C. Toward a science of tumor forecasting for clinical oncology. *Cancer Res.* **75**, 918–923 (2015).
129. Deming, D. A. *et al.* PIK3CA and APC mutations are synergistic in the development of intestinal cancers. *Oncogene* **33**, 2245–2254 (2014).
130. Mahabeleshwar, G. H., Feng, W., Reddy, K., Plow, E. F. & Byzova, T. V. Mechanisms of integrin-vascular endothelial growth factor receptor cross-activation in angiogenesis. *Circ. Res.* **101**, 570–580 (2007).
131. Archetti, M., Ferraro, D. A. & Christofori, G. Heterogeneity for IGF-II production maintained by public goods dynamics in neuroendocrine pancreatic cancer. *Proc. Natl Acad. Sci. USA* **112**, 1833–1838 (2015).
132. Gerlee, P. & Altrock, P. M. Complexity and stability in growing cancer cell populations. *Proc. Natl Acad. Sci. USA* **112**, E2742–E2743 (2015).
133. Marusyk, A. *et al.* Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity. *Nature* **514**, 54–58 (2014).
134. Enderling, H., Hlatky, L. & Hahnfeldt, P. Migration rules: tumours are conglomerates of self-metastases. *Br. J. Cancer* **100**, 1917–1925 (2009).
135. Carmeliet, P. & Jain, R. K. Angiogenesis in cancer and other diseases. *Nature* **407**, 249–257 (2000).
136. Chapman, S. J., Shipley, R. J. & Jawad, R. Multiscale modeling of fluid transport in tumors. *Bull. Math. Biol.* **70**, 2334–2357 (2008).
137. Chaplain, M. A., McDougall, S. R. & Anderson, A. R. Mathematical modeling of tumor-induced angiogenesis. *Annu. Rev. Biomed. Eng.* **8**, 233–257 (2006).
138. Alarcon, T., Byrne, H. M. & Maini, P. K. A multiple scale model for tumor growth. *Multiscale Model. Simul.* **3**, 440–475 (2005).
139. Rockne, R. *et al.* Predicting the efficacy of radiotherapy in individual glioblastoma patients *in vivo*: a mathematical modeling approach. *Phys. Med. Biol.* **55**, 3271–3285 (2010).
140. Jackson, T., Komarova, N. & Swanson, K. R. Mathematical oncology: using mathematics to enable cancer discoveries. *Am. Math. Mon.* **121**, 840–856 (2014).
141. Dale, R. G. The application of the linear-quadratic dose-effect equation to fractionated and protracted radiotherapy. *Br. J. Radiol* **58**, 515–528 (1985).
142. Rosenberg, S. A., Yang, J. C. & Restifo, N. P. Cancer immunotherapy: moving beyond current vaccines. *Nat. Med.* **10**, 909–915 (2004).
143. Kalos, M. *et al.* T cells with chimeric antigen receptors have potent antitumor effects and can establish memory in patients with advanced leukemia. *Sci. Transl Med.* **3**, 95ra73 (2011).
144. Blattman, J. N. & Greenberg, P. D. Cancer immunotherapy: a treatment for the masses. *Science* **305**, 200–205 (2004).
145. Pardoll, D. M. The blockade of immune checkpoints in cancer immunotherapy. *Nat. Rev. Cancer* **12**, 252–264 (2012).
146. Old, L. J. Cancer immunology: the search for specificity—G. H. A. Clowes Memorial lecture. *Cancer Res.* **41**, 361–375 (1981).
147. Abbas, A. K., Lichtman, A. H. & Pillai, S. *Cellular and Molecular Immunology* (Elsevier Health Sciences, 1994).
148. Kuznetsov, V. in *A Survey of Models for Tumor-Immune System Dynamics* (eds Adam, J. A. & Bellomo, N.) 237–294 (Springer, 1997).
149. Bellomo, N. & Preziosi, L. Modelling and mathematical problems related to tumor evolution and its interaction with the immune system. *Math. Computer Modell.* **32**, 413–452 (2000).
150. Kolev, M. Mathematical modelling of the competition between tumors and immune system considering the role of the antibodies. *Math. Computer Modell.* **37**, 1143–1152 (2003).
151. de Pillis, L. G. & Radunskaya, A. A mathematical tumor model with immune resistance and drug therapy: an optimal control approach. *Computat. Math. Methods Med.* **3**, 79–100 (2001).
152. de Pillis, L. G., Radunskaya, A. E. & Wiseman, C. L. A validated mathematical model of cell-mediated immune response to tumor growth. *Cancer Res.* **65**, 7950–7958 (2005).
153. d'Onofrio, A. Metamodeling tumor–immune system interaction, tumor evasion and immunotherapy. *Math. Computer Modell.* **47**, 614–637 (2008).
154. d'Onofrio, A. A general framework for modeling tumor-immune system competition and immunotherapy: mathematical analysis and biomedical inferences. *Phys. D Nonlin. Phenomena* **208**, 220–235 (2005).
155. Sivakumar, P. V., Foster, D. C. & Clegg, C. H. Interleukin-21 is a T-helper cytokine that regulates humoral immunity and cell-mediated anti-tumour responses. *Immunology* **112**, 177–182 (2004).
156. Cappuccio, A., Elishmereni, M. & Agur, Z. Cancer immunotherapy by interleukin-21: potential treatment strategies evaluated in a mathematical model. *Cancer Res.* **66**, 7293–7300 (2006).
157. Pappalardo, F. *et al.* SimB16: modeling induced immune system response against B16-melanoma. *PLoS ONE* **6**, e26523 (2011).
158. Pienta, K. J., Robertson, B. A., Coffey, D. S. & Taichman, R. S. The cancer diaspora: metastasis beyond the seed and soil hypothesis. *Clin. Cancer Res.* **19**, 5849–5855 (2013).
159. Comen, E. & Norton, L. Self-seeding in cancer. *Recent Results Cancer Res.* **195**, 13–23 (2012).
160. Saidel, G. M., Liotta, L. A. & Kleinerman, J. System dynamics of metastatic process from an implanted tumor. *J. Theor. Biol.* **56**, 417–434 (1976).
161. Panetta, J. C. A mathematical model of periodically pulsed chemotherapy: tumor recurrence and metastasis in a competitive environment. *Bull. Math. Biol.* **58**, 425–447 (1996).
162. Enderling, H. *et al.* Paradoxical dependencies of tumor dormancy and progression on basic cell kinetics. *Cancer Res.* **69**, 8814–8821 (2009).
A modelling contribution that combines concepts of a differentiation hierarchy with spatial cell migration and therapy effects.
163. Almendro, V. *et al.* Genetic and phenotypic diversity in breast tumor metastases. *Cancer Res.* **74**, 1338–1348 (2014).
164. Scott, J., Kuhn, P. & Anderson, A. R. Unifying metastasis — integrating intravasation, circulation and end-organ colonization. *Nat. Rev. Cancer* **12**, 445–446 (2012).
165. Weiss, L. Comments on hematogenous metastatic patterns in humans as revealed by autopsy. *Clin. Exp. Metastasis* **10**, 191–199 (1992).
166. Scott, J. G., Basanta, D., Anderson, A. R. & Gerlee, P. A mathematical model of tumour self-seeding reveals secondary metastatic deposits as drivers of primary tumour growth. *J. R. Soc. Interface* **10**, 20130011 (2013).
167. Scott, J. G., Fletcher, A. G., Maini, P. K., Anderson, A. R. & Gerlee, P. A filter-flow perspective of haematogenous metastasis offers a non-genetic paradigm for personalised cancer therapy. *Eur. J. Cancer* **50**, 3068–3075 (2014).
168. Fu, F., Nowak, M. A. & Bonhoeffer, S. Spatial heterogeneity in drug concentrations can facilitate the emergence of resistance to cancer therapy. *PLoS Comput. Biol.* **11**, e1004142 (2015).
169. Haeno, H. *et al.* Computational modeling of pancreatic cancer reveals kinetics of metastasis suggesting optimum treatment strategies. *Cell* **148**, 362–375 (2012).
A stochastic process model that considers tumour growth, death, mutation and dissemination events parameterized using pancreatic cancer patient data to identify improved treatment strategies.
170. Haeno, H. & Michor, F. The evolution of tumor metastases during clonal expansion. *J. Theor. Biol.* **263**, 30–44 (2010).
171. Newton, P. K. *et al.* A stochastic Markov chain model to describe lung cancer growth and metastasis. *PLoS ONE* **7**, e34637 (2012).
172. Newton, P. K. *et al.* Spreaders and sponges define metastasis in lung cancer: a Markov chain Monte Carlo mathematical model. *Cancer Res.* **73**, 2760–2769 (2013).
A mathematical model parameterized using tumour autopsy data to study the metastatic process, which allows different tumour sites to be classified as spreaders or sponges.
173. Bretcha-Boix, P., Rami-Porta, R., Mateu-Navarro, M., Hoyuela-Alonso, C. & Marco-Molina, C. Surgical treatment of lung cancer with adrenal metastasis. *Lung Cancer* **27**, 101–105 (2000).
174. Bazhenova, L. *et al.* Adrenal metastases in lung cancer: clinical implications of a mathematical model. *J. Thorac. Oncol.* **9**, 442–446 (2014).
175. Pathak, A. & Kumar, S. Independent regulation of tumor cell migration by matrix stiffness and confinement. *Proc. Natl Acad. Sci. USA* **109**, 10334–10339 (2012).
176. Michor, F., Liphardt, J., Ferrari, M. & Widom, J. What does physics have to do with cancer? *Nat. Rev. Cancer* **11**, 657–670 (2011).
177. Jain, R. K., Tong, R. T. & Munn, L. L. Effect of vascular normalization by antiangiogenic therapy on interstitial hypertension, peritumor edema, and lymphatic metastasis: insights from a mathematical model. *Cancer Res.* **67**, 2729–2735 (2007).
178. Jain, R. K. Normalizing tumor microenvironment to treat cancer: bench to bedside to biomarkers. *J. Clin. Oncol.* **31**, 2205–2218 (2013).
179. Coldman, A. J. & Goldie, J. H. A stochastic model for the origin and treatment of tumors containing drug-resistant cells. *Bull. Math. Biol.* **48**, 279–292 (1986).
180. Goldie, J. H. & Coldman, A. J. Quantitative model for multiple levels of drug resistance in clinical tumors. *Cancer Treat. Rep.* **67**, 923–931 (1983).
181. Goldie, J. H. & Coldman, A. J. The genetic origin of drug resistance in neoplasms: implications for systemic therapy. *Cancer Res.* **44**, 3643–3653 (1984).
182. Norton, L. & Simon, R. Growth curve of an experimental solid tumor following radiotherapy. *J. Natl Cancer Inst.* **58**, 1735–1741 (1977).
183. Citron, M. L. *et al.* Randomized trial of dose-dense versus conventionally scheduled and sequential versus concurrent combination chemotherapy as postoperative adjuvant treatment of node-positive primary breast cancer: first report of Intergroup Trial C9741/Cancer and Leukemia Group B Trial 9741. *J. Clin. Oncol.* **21**, 1431–1439 (2003).
184. Goldie, J. H. & Coldman, A. J. A mathematic model for relating the drug sensitivity of tumors to their spontaneous mutation rate. *Cancer Treat. Rep.* **63**, 1727–1733 (1979).
185. Bonadonna, G., Zambetti, M., Moliterni, A., Gianni, L. & Valagussa, P. Clinical relevance of different sequencing of doxorubicin and cyclophosphamide, methotrexate, and fluorouracil in operable breast cancer. *J. Clin. Oncol.* **22**, 1614–1620 (2004).
186. Foo, J. & Michor, F. Evolution of acquired resistance to anti-cancer therapy. *J. Theor. Biol.* **355**, 10–20 (2014).
187. Foo, J. & Michor, F. Evolution of resistance to targeted anti-cancer therapies during continuous and pulsed administration strategies. *PLoS Comput. Biol.* **5**, e1000557 (2009).
188. Chmielecki, J. *et al.* Optimization of dosing for EGFR-mutant non-small cell lung cancer with evolutionary cancer modeling. *Sci. Transl Med.* **3**, 90ra59 (2011).
An evolutionary mathematical model that identifies optimum dosing strategies of targeted drugs to delay the emergence of resistance.
189. Foo, J., Chmielecki, J., Pao, W. & Michor, F. Effects of pharmacokinetic processes and varied dosing schedules on the dynamics of acquired resistance to erlotinib in EGFR-mutant lung cancer. *J. Thorac. Oncol.* **7**, 1583–1593 (2012).
190. Michor, F. *et al.* Dynamics of chronic myeloid leukaemia. *Nature* **435**, 1267–1270 (2005).
191. Neal, M. L. *et al.* Response classification based on a minimal model of glioblastoma growth is prognostic for clinical outcomes and distinguishes progression from pseudoprogression. *Cancer Res.* **73**, 2976–2986 (2013).

192. Corwin, D. *et al.* Toward patient-specific, biologically optimized radiation therapy plans for the treatment of glioblastoma. *PLoS ONE* **8**, e79115 (2013).
193. Jain, H. V., Clinton, S. K., Bhinder, A. & Friedman, A. Mathematical modeling of prostate cancer progression in response to androgen ablation therapy. *Proc. Natl Acad. Sci. USA* **108**, 19701–19706 (2011).
194. Gatenby, R. A., Silva, A. S., Gillies, R. J. & Frieden, B. R. Adaptive therapy. *Cancer Res.* **69**, 4894–4903 (2009).
A mathematical model that suggests that outcomes can be improved by maintaining a stable tumour burden in which resistant tumour cells are suppressed by sensitive cells.
195. Gatenby, R. A. & Frieden, B. R. Inducing catastrophe in malignant growth. *Math. Med. Biol.* **25**, 267–283 (2008).
196. Luria, S. E. & Delbruck, M. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **28**, 491–511 (1943).
197. Frank, S. A. Somatic mosaicism and cancer: inference based on a conditional Luria-Delbruck distribution. *J. Theor. Biol.* **223**, 405–412 (2003).
198. Haeno, H., Iwasa, Y. & Michor, F. The evolution of two mutations during clonal expansion. *Genetics* **177**, 2209–2221 (2007).
199. Durrett, R. & Moseley, S. Evolution of resistance and progression to disease during clonal expansion of cancer. *Theor. Popul. Biol.* **77**, 42–48 (2010).
200. Iwasa, Y., Michor, F. & Nowak, M. A. Evolutionary dynamics of escape from biomedical intervention. *Proc. Biol. Sci.* **270**, 2573–2578 (2003).
201. Komarova, N. L. & Wodarz, D. Stochastic modeling of cellular colonies with quiescence: an application to drug resistance in cancer. *Theor. Popul. Biol.* **72**, 523–538 (2007).
202. Bozic, I. & Nowak, M. A. Timing and heterogeneity of mutations associated with drug resistance in metastatic cancers. *Proc. Natl Acad. Sci. USA* **111**, 15964–15968 (2014).
203. Mumenthaler, S. M. *et al.* Evolutionary modeling of combination treatment strategies to overcome resistance to tyrosine kinase inhibitors in non-small cell lung cancer. *Mol. Pharm.* **8**, 2069–2079 (2011).
204. Komarova, N. Stochastic modeling of drug resistance in cancer. *J. Theor. Biol.* **239**, 351–366 (2006).
205. Komarova, N. L., Burger, J. A. & Wodarz, D. Evolution of ibrutinib resistance in chronic lymphocytic leukemia (CLL). *Proc. Natl Acad. Sci. USA* **111**, 13906–13911 (2014).
206. Altrock, P. M. & Traulsen, A. Deterministic evolutionary game dynamics in finite populations. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **80**, 011909 (2009).
207. Werner, B., Lutz, D., Brummendorf, T. H., Traulsen, A. & Balabanov, S. Dynamics of resistance development to imatinib under increasing selection pressure: a combination of mathematical models and *in vitro* data. *PLoS ONE* **6**, e28955 (2011).
208. Gallaher, J. & Anderson, A. R. Evolution of intratumoral phenotypic heterogeneity: the role of trait inheritance. *Interface Focus* **3**, 20130016 (2013).
209. Huang, W., Haubold, B., Hauert, C. & Traulsen, A. Emergence of stable polymorphisms driven by evolutionary games between mutants. *Nat. Commun.* **3**, 919 (2012).
210. Huang, W., Hauert, C. & Traulsen, A. Stochastic game dynamics under demographic fluctuations. *Proc. Natl Acad. Sci. USA* **112**, 9064–9069 (2015).
211. Archetti, M., Ferraro, D. A. & Christofori, G. Reply to Gerlee and Altrock: diffusion and population size in game theory models of cancer. *Proc. Natl Acad. Sci. USA* **112**, E2744 (2015).
212. Tabassum, D. P. & Polyak, K. Tumorigenesis: it takes a village. *Nat. Rev. Cancer* **15**, 473–483 (2015).
213. Pacheco, J. M., Santos, F. C. & Dingli, D. The ecology of cancer from an evolutionary game theory perspective. *Interface Focus* **4**, 20140019 (2014).
214. Axelrod, R., Axelrod, D. E. & Pienta, K. J. Evolution of cooperation among tumor cells. *Proc. Natl Acad. Sci. USA* **103**, 13474–13479 (2006).
215. Archetti, M. Evolutionary game theory of growth factor production: implications for tumour heterogeneity and resistance to therapies. *Br. J. Cancer* **109**, 1056–1062 (2013).
216. Basanta, D. *et al.* Investigating prostate cancer tumour-stroma interactions: clinical and biological insights from an evolutionary game. *Br. J. Cancer* **106**, 174–181 (2012).
217. Kianercy, A., Veltri, R. & Pienta, K. J. Critical transitions in a game theoretic model of tumour metabolism. *Interface Focus* **4**, 20140014 (2014).
218. Basanta, D., Simon, M., Hatzikirou, H. & Deutsch, A. Evolutionary game theory elucidates the role of glycolysis in glioma progression and invasion. *Cell Prolif.* **41**, 980–987 (2008).

Competing interests statement
The authors declare no competing interests.