

Computational biology

How mathematical modelling can help cure cancer

Understanding how living cells work is difficult due to the number of varied and complex processes occurring in them. This complexity can be elucidated by breaking these processes down into simpler components and focusing on a particular mechanism. One approach to this study is to use mathematical equations – the basis of computational modelling. Dr Susan Mertins, the founder and CEO of Biosystem Strategies LLC, in the USA, is exploring how ordinary differential equations and machine learning can be applied to cancer data for biomarker discovery and drug development, leading to improvements in personalised medicine.

Biology is the study of living organisms, which can be described as complex systems. Cells can have multiple sophisticated functions such as protein production and photosynthesis, and biological systems are dynamic and convoluted which makes them difficult to analyse. However, specific mechanisms can be interpreted more easily using quantitative metrics. In other words, cell behaviours can be simplified by focusing on one specific aspect and only considering cell elements relevant to the question being asked. The biological phenomenon can then be approximated using mathematical equations. This is the principle underlying computational modelling, where biological mechanisms are represented using laws from physics and chemistry.

Computational modelling enables us to simulate reality and make useful predictions in medicine, from drug development to biomarker discovery. Dr Susan Mertins, founder and CEO of BioSystems Strategies, LLC, is using both computational modelling and machine learning to detect drug targets and biomarkers that will help develop personalised approaches to cancer treatment.

CREATING A COMPUTATIONAL MODEL

The first step in computational modelling is knowing what you want to model. This may seem obvious but being clear about this will determine your choice of model and analysis, as recently emphasised by several researchers in the field (eg, Fogarty et al, 2022, Braakman et al, 2022). This conceptual phase is then followed by building the actual model and writing the code for the simulation. While doing so, one needs to consider how the model can be verified: this step consists of estimating the parameters and calibrating the model (Braakman et al, 2022). As Musuamba et al (2021) describe, verification can be seen as solving the equation in the right way, while the following step – validation – consists of solving the right equation. Model validation is the ability to simulate the conditions of interest with certain sensitivities and uncertainties, which reflect the quality of the model. Finally, one can compare the simulation with comparator data to assess the credibility of the model, for example using the results of a lab experiment.

WHAT TYPES OF MODELS ARE THERE?

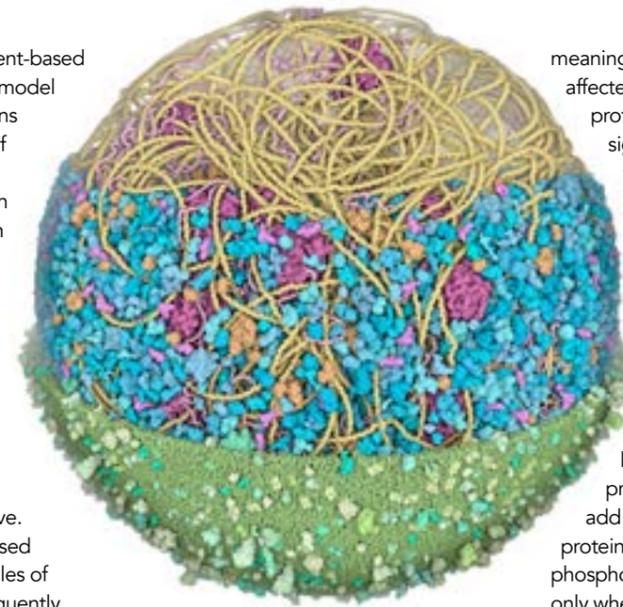
The type of model needed depends on the spatiotemporal scale of the biological process (Dada and Mendes, 2011); for example, a reaction between enzymes will be faster than the development of a whole organ. The intracellular scale can be described with ordinary differential equations (ODE) which involve a single variable and can describe a particle moving through time. Mechanisms such as protein gradients, however, are described by partial differential equations (PDE) which involve multiple variables and can therefore describe a surface changing over time (Carbo et al, 2014). At the cellular scale, there are two main options of models depending on whether the space and time are considered discrete (on-lattice) or continuous (off-lattice)

(Nava-Sedeño et al, 2020). Agent-based models are a typical off-lattice model used to describe the interactions between individual elements of the system. On-lattice models can be separated into two main categories: cellular automata in which agents are represented by a single pixel, and cellular Potts models where agents are a collection of pixels – in other words, it can simulate collective as well as individual cell behaviour. Cellular Potts models can therefore provide a greater resolution but are more computationally expensive. Whole-cell modelling can be used to incorporate the different scales of a biological system and consequently predict how observable traits will arise from genetic mechanisms.

APPLYING MODELS TO CANCER

Cancer is a complex disease which is hard to treat because of its heterogeneity: different patients have tumours with different mutations and even cells from the same tumour can have different mutations. Treating a tumour is therefore challenging, because if some cells are resistant to a treatment the tumour can recur. Patients may have different metabolisms and react to a drug differently. So, understanding the actions of the disease at a cellular level is critical for effective treatment.

Cellular responses are the result of a series of molecular events altogether referred to as signal transduction pathways or 'reaction networks'. The shape of proteins determines the outcome of the pathway, as shape dictates which proteins are attracted/repelled and how proteins interact with each other. These transduction pathways depend on the concentrations of proteins and the speed of their reaction; they control various mechanisms such as gene expression and metabolism, which cancer interferes with. Output from the reactions also depend on feedback loops and are sensitive to input as well. Specifically, cellular responses can be inhibited through negative feedback loops and promoted via positive feedback loops. Under normal conditions, doubling the amount of proteins doubles the reaction rate, but ODEs representing protein



A 3D model of a cell of the pathogenic bacterium *Mycoplasma genitalium* at the beginning of its life cycle.

Acknowledgement: Martina Maritan, Ludovic Autin, David S Goodsell, Scripps Research and RCSB Protein Data Bank. doi: 10.2210/rcsb_pdb/goodsell-gallery-040

meaning that the pathway output is more affected by small changes in the initial protein concentration. Modelling signal transduction with computers therefore can be described as dynamic or kinetic and using these models allows the study of drug mechanism of action, or pharmacodynamics.

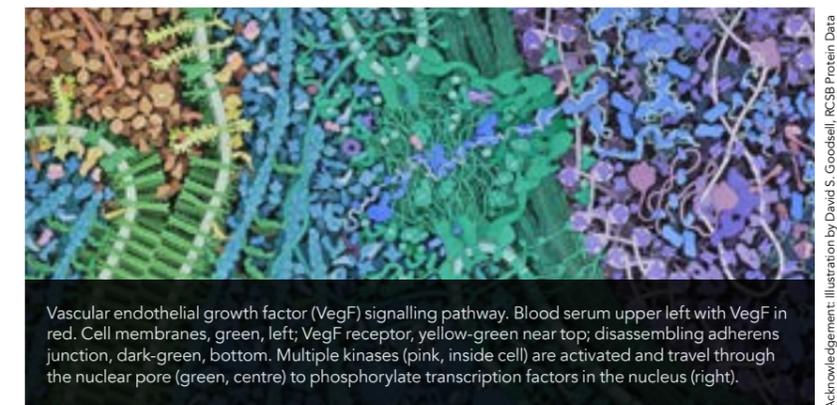
MAPK KINASE SIGNALLING AND DRUG RESISTANCE

The mitogen-activated protein kinase (MAPK) pathway involves proteins called kinases which add phosphate groups to other proteins to modify their function. This phosphorylation mechanism is possible only when the protein is activated, meaning that it is in the right shape. This cascade includes proteins which are mutated in cancer as they control cell proliferation, apoptosis (cell death), cell motility, and differentiation. Rauch et al (2016) explained how computational modelling helps make sense of the MAPK

The type of model needed depends on the spatiotemporal scale of the biological process.

concentrations over time have revealed that chemical responses are not always linear (Goldbeter and Koshland, 1981). For example, in a saturated solution (in which the concentration of a protein has reached a certain value) the reaction rate in the equation changes and the initial protein concentration is not proportional to the output cellular response (motility, or energy production, for example); instead, the protein becomes ultrasensitive,

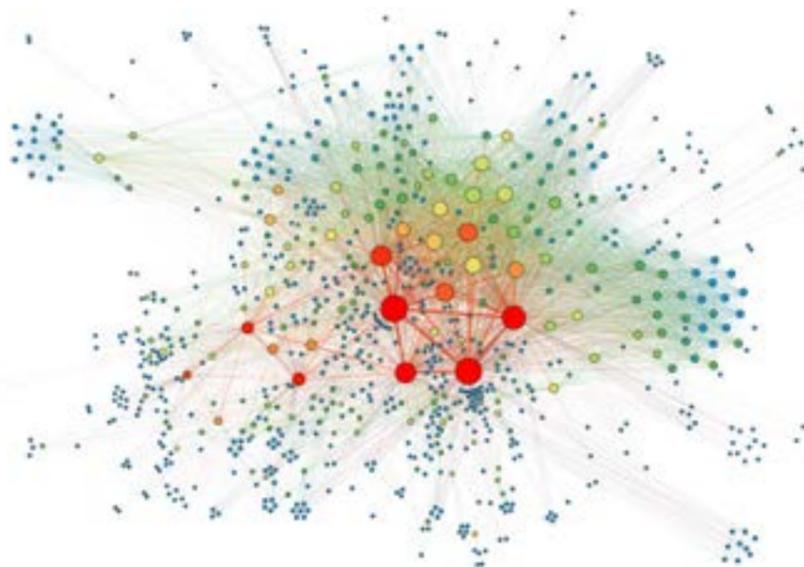
network; for example, when two proteins react and form a dimer (two identical molecules linked together – a process known as dimerization), this can lead to drug resistance. Overall, the efficacy of the drug depends on how the binding of the drug and the protein impacts the shape of the protein. Dimerization tends to be favoured between a drug-bound protein and a drug-free protein, but this conformation activates the



Vascular endothelial growth factor (VegF) signalling pathway. Blood serum upper left with VegF in red. Cell membranes, green, left; VegF receptor, yellow-green near top; disassembling adherens junction, dark-green, bottom. Multiple kinases (pink, inside cell) are activated and travel through the nuclear pore (green, centre) to phosphorylate transcription factors in the nucleus (right).

Acknowledgement: Illustration by David S. Goodsell, RCSB Protein Data Bank. doi: 10.2210/rcsb_pdb/goodsell-gallery-041





Signalling network or node cloud.

drug-free protein instead of inhibiting it. You can compare this dimerization process to a puzzle: if the two pieces are a complementary fit, they have more chances of being stable together. In this case, the drug-bound protein and the drug-free protein bind easily together, leading to the accumulation of such dimers and consequently to drug resistance. Nonetheless, two drugs

ineffective on their own can overcome drug resistance if they are combined.

FINDING BIOMARKERS AND DRUG TARGETS

In her review, Mertins illustrates how computational modelling is used to represent protein concentrations of the MAPK kinase pathway (Mertins, 2022). She explains that an ODE can

be used to describe the amount of protein modification, also known as phosphorylation levels. This simulation requires an input from proteomic (protein) databases to set the initial quantitative parameters. From the ODE output, machine learning can be applied to discover novel biomarkers and targets for new drugs. Indeed, machine learning helps us make connections that are not likely to be made by humans, such as predicting cancer survival or being able to detect cancer early. For instance, the extent of protein modification resulting from the ODE can be investigated to find a correlation with cancer prognosis. Besides, novel drug targets can be found by mimicking mutations which prevent protein production: parameters representing proteins can be removed to simulate the effect of such mutations. By doing so in a systematic way, the importance of proteins can be assessed by analysing the outcome of the simulation. If the absence of a protein results in cancer cell death, it means that this protein is a promising drug target.

MODELLING FOR PERSONALISED MEDICINE

There are many advantages to using computational simulations. They are usually faster and cheaper than a lab experiment, enabling more candidate drug targets to be revealed which is a significant help in tailoring medicines to individual patients. Computational modelling also decreases treatment-related risks: digital twins can be made to understand how the cancer would react to a drug depending on the molecular profile of a patient in a simulation. Moreover, it can help predict drug resistance and the evolution of a tumour. Conditions and parameters can be set to answer a specific question, and the time steps can be controlled to focus on a specific aspect of interest.

Of course, simulations are often constrained by assumptions which means they are a simplification of reality, and their results do need to be validated in the real world. Despite these limitations, however, computational modelling enables us to move much more quickly and efficiently towards drug discovery and personalised treatments for complex diseases than we can manage by other means.

Machine learning helps us make connections that are not likely to be made by us, such as predicting cancer survival.



3D model of a critical MAPK pathway kinase bound to a clinical agent.

Photo Credit: Martin Grandjean, CC BY-SA 4.0, via Wikimedia Commons

StudioMolekul/Shutterstock.com



Behind the Research

Dr Susan Mertins

E: smertins@biosystemsstrategies.com **W:** www.biosystemsstrategies.com

Research Objectives

Susan Mertins is using computational modelling resembling virtual cells and machine learning to develop new approaches to treating cancer.

Detail

Address

201 E Patrick St, #3102
Frederick, MD 21701, USA

Bio

Susan D Mertins, PhD has over 30 years of experience in drug discovery and development in oncology. She conducted bench science at the National Institutes of Health, initiated her computational work at Mount St Mary's University, Maryland, and founded BioSystems Strategies, LLC to advance both interests in a consulting capacity.

Funding

- BioSystems Strategies, LLC

References

- Mertins, SD, (2022) Capturing biomarkers and molecular targets in cellular landscapes from dynamic reaction network models and machine learning, *Front Oncol*, 11:805592. doi.org/10.3389/fonc.2021.805592
- Fogarty, L, Ammar, M, Holding, T, Powell, A, Kandler, A, (2022) Ten simple rules for principled simulation modelling, *PLoS Comput Biol*, 18(3). doi.org/10.1371/journal.pcbi.1009917
- Braakman, S, Pathmanathan, P, Moore, H, (2022) Evaluation framework for systems models, *CPT Pharmacometrics Syst Pharmacol*, 11(3), 264–289. doi.org/10.1002/psp4.12755
- Musuamba, FT, et al, (2021) Scientific and regulatory evaluation of mechanistic in silico drug and disease models in drug development: Building model credibility, *CPT Pharmacometrics Syst Pharmacol*, 10(8), 804–825. doi.org/10.1002/psp4.12669
- Nava-Sedeño, JM, et al, (2020) Modelling collective cell motion: are on- and off-lattice models equivalent?, *Philos Trans R Soc Lond B Biol Sci*, 375(1807):20190378. doi.org/10.1098/rstb.2019.0378
- Rauch, N, Rukhlenko, OS, Kolch, W, Kholodenko, BN, (2016) MAPK kinase signalling dynamics regulate cell fate decisions and drug resistance, *Current opinion in structural biology*, 41, 151–158. doi.org/10.1016/j.sbi.2016.07.019
- Carbo, A, et al, (2014) Computational modeling of heterogeneity and function of CD4+ T cells, *Front Cell Dev Biol*, 2:31. doi.org/10.3389/fcell.2014.00031
- Dada, JO, Mendes, P, (2011) Multi-scale modelling and simulation in systems biology, *Integrative biology*, 3(2), 86–96. doi.org/10.1039/c0ib00075b
- Goldbeter, A, Koshland, DE Jr, (1981) An amplified sensitivity arising from covalent modification in biological systems, *Proc Natl Acad Sci USA*, 78(11), 6840–4. doi.org/10.1073/pnas.78.11.6840

Personal Response

Do you think that machine learning will become indispensable for curing cancer?

Machine learning holds great promise for diagnosis and treatment for cancer. We are in the very early stages and I expect combinations of the various modelling algorithms will be needed such as connecting cellular automata, ODE equations, and artificial intelligence.