The CellCycler

Systems Forecasting

The CellCycler is a parsimonius model of a growing tumour. The aim is to capture the key dynamics of tumour behaviour so as to simulate and predict the effect of treatment, either *in vitro*, *in vivo*, or in the clinic. The model uses circa 100 ordinary differential equations to simulate cells as they pass through the phases of the cell cycle. However the guiding philosophy of the model is that it should only use parameters that can be observed or reasonably well approximated. The level of analysis is limited to cell state observables such as phase, apoptosis, and damage.

Model details and background

The CellCycler model is divided into three separate components: a cell population model, a PK model, and a tumour growth model.

The cell population model needs to reflect the phase of each cell, growth of the cells as they progress through the cell cycle, and effects such as cell damage and death. It should also allow for a spread of cell cycle times within the population. A number of approaches to modelling a cell population can be found in the literature. One is to model cells individually, using an agent-based model. This technique offers a number of advantages, such as the ability to track individual cells, but can be computationally rather slow. Another method is to use ordinary differential equations, where each equation models cells in a particular state.[1] For example there may be separate equations for cells in G1 phase, or S phase, or damaged cells. However the problem here is that the model may not accurately capture the time-dependent effect of certain drugs, which can be important when exploring phenomena such as synchronisation. For example, suppose a drug damages cells in S phase. Then to determine the effect of the drug on cells, we need to know exactly where cells are in the cycle – if they are just starting S-phase when exposed to drug, then they will get more exposure than cells which are about to exit.

[1] Bridging the gap between in vitro and in vivo: Dose and schedule predictions for the ATR inhibitor AZD6738. Stephen Checkley, Linda MacCallum, James Yates, Paul Jasper, Haobin Luo, John Tolsma, Claus Bendtsen. Scientific Reports.2015;5(3)13545.

The CellCycler model addresses this problem by dividing the cell cycle up into a number of separate compartments of equal length. As discussed further below, typically N=50 compartments have been found to give adequate resolution. In the absence of drug, the equation for the volume V_n of each compartment is very simple:

$$\frac{dV_n}{dt} = k_1 V_{n-1} - k_2 V_n .$$

Here k_1 represents the change in volume due to cells transiting from compartment *n-1* to compartment *n*, and k_2 represents the change in volume due to cells transiting from compartment *n* to compartment *n+1* (note the indices are cyclic).

We assume that cells grow constantly as they progress through the cell cycle. As cells enter from the previous compartment, their volume therefore expands by a factor $2^{\frac{1}{N}}$. So after passing through all N compartments, the cells will have grown by a factor 2 as expected. It follows that the volume change rate k_1 corresponding to cells entering from the previous compartment must be a factor $2^{\frac{1}{N}}$ greater than the corresponding rate k_2 for cells leaving the current compartment, i.e.

$$k_1 = 2^{\frac{1}{N}} k_2 .$$

If we now sum the rate equations for the separate compartments, we obtain the growth rate for the total volume V_T which is

$$\frac{dV_T}{dt} = (k_1 - k_2) V_T .$$

Now, if the average length of the cell cycle is t_d, then the total number of cells is doubling in time t_d, and we can also express the growth rate as an exponential equation

$$\frac{dV_T}{dt} = e^{at}$$

where

$$a = \frac{\log(2)}{t_d} .$$

We therefore have two equations for the two unknowns k_1 and k_2, which can be solved to give

$$k_1 = \frac{\log(2)}{t_d}\frac{2^{\frac{1}{N}}}{2^{\frac{1}{N}}-1},\ k_2 = \frac{\log(2)}{t_d}\frac{1}{2^{\frac{1}{N}}-1}.$$

Using the approximation, valid for large *n*, that

$$\frac{2^{\frac{1}{N}}}{2^{\frac{1}{N}}-1} \cong \frac{1}{\log(2)}$$

we note that the rates are roughly

$$k_1 \cong \frac{(N+\log(2))}{t_d},\ k_2 \cong \frac{N}{t_d}.$$

The rate constants therefore reflect the rate at which cells pass through each stage, with an additional growth term appearing in k_1 which accounts for cell growth.

The initial condition assumes that each compartment has an identical volume of cells. It can be shown that this is consistent with the assumption that cells grow at a constant rate. For example, if mitosis occurs in the last compartment, then there will be twice as many cells in the first compartment, but the same volume.

Discretisation effects

Because the CellCycler divides the cell cycle into a fixed number of compartments, one consequence is that the doubling time has a degree of uncertainty. This is illustrated by the figure below, which shows how a perturbation at time zero in one compartment tends to blur out over time, for models with 25, 50, and 100 compartments. and a doubling time of 24 hours. In each case a perturbation is made to compartment 1 at the beginning of the cell cycle (the magnitude is scaled to the number of compartments so the total size of the perturbation is the same in terms of total volume). For the case with 50 compartments, the curve is closely approximated by a normal distribution with standard deviation of 3.5 hours or about 15 percent (not shown).

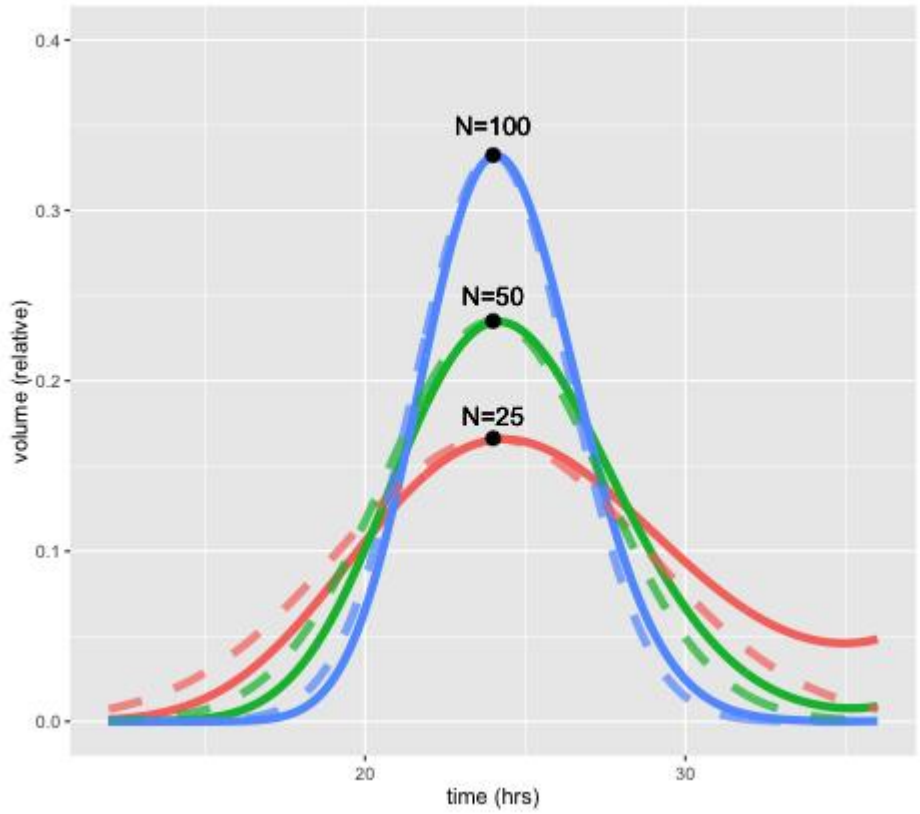

Figure 1. The solid lines show volume in compartment 1 following a perturbation to that compartment alone, after one cell doubling period of 24 hours. The cases shown are with *N*=25, 50, and 100 compartments. The black dots show the estimated peak volumes for the three cases, using the formula developed below. Dashed lines are the corresponding normal distributions.

This blurring effect due to discretisation is a desirable feature (if it didn't exist, we would have to add it) because it is equivalent to saying that the cell population has a variable doubling time, as is the case in growing tumours. One implication is that synchronisation effects caused by drugs reduce over time. While we don't usually have exact data on the spread of doubling times in the growing layer, a choice in the region of 50 compartments gives what appears to be a reasonable degree of spread.

It is possible to obtain a very simple analytic expression for the curves shown in the figure. The model equations for growing cells can be written:

$$\frac{d\mathbf{V}}{dt} = (k_1\mathbf{J} - k_2\mathbf{I})\mathbf{V}$$

where **I** is the identity matrix, and **J** is the identity matrix with columns permuted once to the left. For example, for the case with only four compartments, we would have

$$\mathbf{J} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

This equation has as its solution

$$\mathbf{V} = e^{(k_1 t\mathbf{J} - k_2 t\mathbf{I})}\mathbf{V_0} = e^{-k_2 t} e^{k_1 t\mathbf{J}}\mathbf{V_0}.$$

To determine the effect after one doubling time, we use the approximations for k_1 and k_2 above, and set $t = t_d$ to obtain

$$k_1 t_d \cong N + \log(2),\ k_2 t_d \cong N$$

so

$$\mathbf{V} \cong e^{-N} e^{(N+\log(2))\mathbf{J}}\mathbf{V}(0)$$

Taking a Taylor expression, we have

$$e^{(N+\log(2))\mathbf{J}} = \mathbf{I} + (N+\log(2))\mathbf{J} + \frac{1}{2!}(N+\log(2))^2\mathbf{J}^2 + \frac{1}{3!}(N+\log(2))^3\mathbf{J}^3 + \ldots$$

Consider a step perturbation so that the vector **V(0)** is 0 in all elements except for the first, normalised for the total volume being perturbed so that

$$V_1(0) = \frac{N}{t_d}.$$

In this case all powers of the matrix **J**, when multiplied by **V(0)**, are zero except for those powers that are divisible by N, so the expression for the volume V_1 in the first compartment after time t_d reduces to

$$V_1(t_d) \cong e^{-N}\left(1 + \frac{1}{N!}(N+\log(2))^N\right)\frac{N}{t_d}$$

plus higher terms which can be neglected over time scales of a single cycle.

Using Stirling's approximation

$$N! \cong \sqrt{2\pi N}\left(\frac{N}{e}\right)^N$$

then gives

$$V_1(t_d) \cong \frac{1}{\sqrt{2\pi N}}\left(\frac{N+\log(2)}{N}\right)^N \frac{N}{t_d}.$$

The term in brackets converges for large N to 2 (this can be seen by taking the logarithm and using a first-order approximation), so

$$V_1(t_d) \cong \frac{1}{t_d}\sqrt{\frac{2N}{\pi}}\,.$$

The blurring effect due to discretisation can therefore be computed for abitrary values of N. Results are in good agreement with simulations, as shown by the black dots in Figure 1. The peak of the distribution around the doubling time will increase only slowly with the square root of N. It is therefore impractical to obtain a highly peaked distribution; however as mentioned above this would be unrealistic because it would imply a near-perfect degree of cell synchronisation which does not occur in growing tumours.

The distributions in the figure can be approximated by a normal distribution (shown by dashed lines), with peak given by the above formula, and a corresponding standard deviation, normalised to doubling time, of

$$\sigma = \frac{1}{\sqrt{N}}\,.$$

So for example using 25 compartments gives a normalised standard deviation of about 0.02 or 20 percent, while using 100 compartments decreases this to 0.01 or 10 percent. The default value is 50 compartments, which gives a standard deviation in doubling times of about 14 percent.

Effect of drug

When drug is present, some cells will be damaged and possibly repaired, or killed outright. To model these effects, we include in the model an additional N compartments for damaged cells, and one additional variable which represents cells lost to apoptosis. The volume of the damaged cells D_n is then given by

$$\frac{dD_n}{dt} = k_d V_n - k_r$$

where k_d is the drug-dependent rate of damage, and k_r is the repair rate. The rate for the volume A of cells lost to apoptosis is given by

$$\frac{dA}{dt} = k_a V_n\,.$$

The equation for proliferating cells is correspondingly modified to give

$$\frac{dV_n}{dt} = k_1 V_{n-1} - k_2 V_n - k_d V_n + k_r - k_a V_n .$$

Since k_1 and k_2 are determined from the cell doubling time, the only additional parameters required by the model are the drug-dependent properties k_d, k_r and k_a, as well as the allocation of compartments between the diferent phases.

Tumour growth

Tumour growth is caused by the proliferation of dividing cells. For example if cells have a cell cycle length t_d, then the total number of cells will double every t_d hours, so the volume will be given by

$$V = 2^{\frac{t}{t_d}} = e^{at}$$

where

$$a = \frac{\log(2)}{t_d}$$

Because tumours are driven by exponential growth, it has been common to use versions of exponential or quasi-exponential growth curves to model tumour volumes. In 1956 Collins introduced the concept of doubling times to measure tumour growth.[2] In practice, the exponential curve requires some modification, because it was found that tumours seemed to grow more slowly as time goes on. One variant is the Gompertz curve, which gives a tapered growth rate, at the expense of extra parameters.[3]

However, it has often been observed empirically that tumor diameters, as opposed to volumes, appear to grow in a roughly linear fashion. Indeed, this has been known since at least the 1930s. As Mayneord wrote in 1932: "The rather surprising fact emerges that the increase in long diameter of the implanted tumour follows a linear law."[4] Furthermore, he noted, there was "a simple explanation of the approximate linearity in terms of the structure of the sarcoma. On cutting open the tumour it is often apparent that not the whole of the mass is in a state of active growth, but only a thin capsule (sometimes not more than 1 cm thick)

[2] Collins, V. P., Loeffler, R. K. & Tivey, H. Observations on growth rates of human tumors. *The American journal of roentgenology, radium therapy, and nuclear medicine* **76**, 988-1000 (1956).
[3] Yorke, E. D., Fuks, Z., Norton, L., Whitmore, W. & Ling, C. C. Modeling the Development of Metastases from Primary and Locally Recurrent Tumors: Comparison with a Clinical Data Base for Prostatic Cancer. *Cancer Research* **53**, 2987-2993 (1993).
[4] 1932;16:841-846. Am J Cancer. W. V. Mayneord. On a Law of Growth of Jensen's Rat Sarcoma.

enclosing the necrotic centre of the tumour." As shown below, the existence of a growing layer leads to linear growth.

The linear growth in diameter translates to cubic growth of the tumour, which of course grows more slowly than an exponential curve at longer times. In other words, rather than use a modified exponential curve to fit volume growth, it may be better to use a linear equation to model diameter.

This idea that tumour growth is driven by an outer layer of proliferating cells, surrounding a quiescent or necrotic core, has been featured in a number of mathematical models, and it is the one we incorporate here.[5] If we assume that the proliferating layer has thickness *d*, and is growing at a rate *a*, then the volume of the layer is

$$V_p = 4\pi r^2 d$$

and it is growing at a rate

$$\frac{dV_p}{dt} = aV_p = a4\pi r^2 d$$

The growth equation for the radius is given by

$$\frac{dr}{dt} = \frac{dr}{dV}\frac{dV}{dt} = \left(\frac{1}{4\pi r^2}\right)a4\pi r^2 d = ad$$

which is solved to give the linear equation

$$r = R_0 + adt$$

To translate from cell population growth (with growth rate *a*) to tumour growth, we therefore need just two additional parameters, which are the thickness of the growing layer *d*, and the initial volume R_0.

The linear growth equation will of course not be a perfect fit for the growth of all tumours, but it is based on a consistent and empirically verified model of tumour growth, and can be easily parameterised and fit to data.

Using the CellCycler

[5] Bridging the gap between in vitro and in vivo: Dose and schedule predictions for the ATR inhibitor AZD6738. Stephen Checkley, Linda MacCallum, James Yates, Paul Jasper, Haobin Luo, John Tolsma, Claus Bendtsen. Scientific Reports.2015;5(3)13545.

The CellCycler model has been incorporated into a freely accessible Shiny web application. The starting point for the program is the Cells page, which is used to model the dynamics of a growing cell population. The key parameters are the average cell doubling time, and the fraction spent in each phase (G2 is set automatically since the proportions must add to 1). The doubling time is assumed to be variable, with a range that depends on the number of model compartments. This can be adjusted in the Advanced tab: 25 compartments gives a standard deviation for cell doubling times of about 20 percent, 50 compartments gives 14 percent, and 100 compartments gives 10 percent. Note that the number of compartments affects both the simulation time (more compartments is slower), and the discretisation of the cell cycle. For example with 50 compartments the proportional phase times will be rounded off to the nearest 1/50=0.02.

In addition the user selects the simulation time, and plotting choices such as growing or damaged cells. The simulation is not reactive, so the user must press the Run Simulation button to run the model. The plot will then show the volume of cells in each phase, as well as the total volume, normalised to an initial volume of 1. Model settings can be saved to or loaded from a csv file.

The next pages, PK1 and PK2, are used to parameterise the PK models and drug effects. The program has a choice of simple decay (K-PD), or one or two-compartment models, with adjustable parameters such as Dose/Volume. A schematic of the compartment PK models is given below. In addition the phase of action (choices are G1, S, G2, M, or all), and rates for death, damage, and repair can be adjusted.

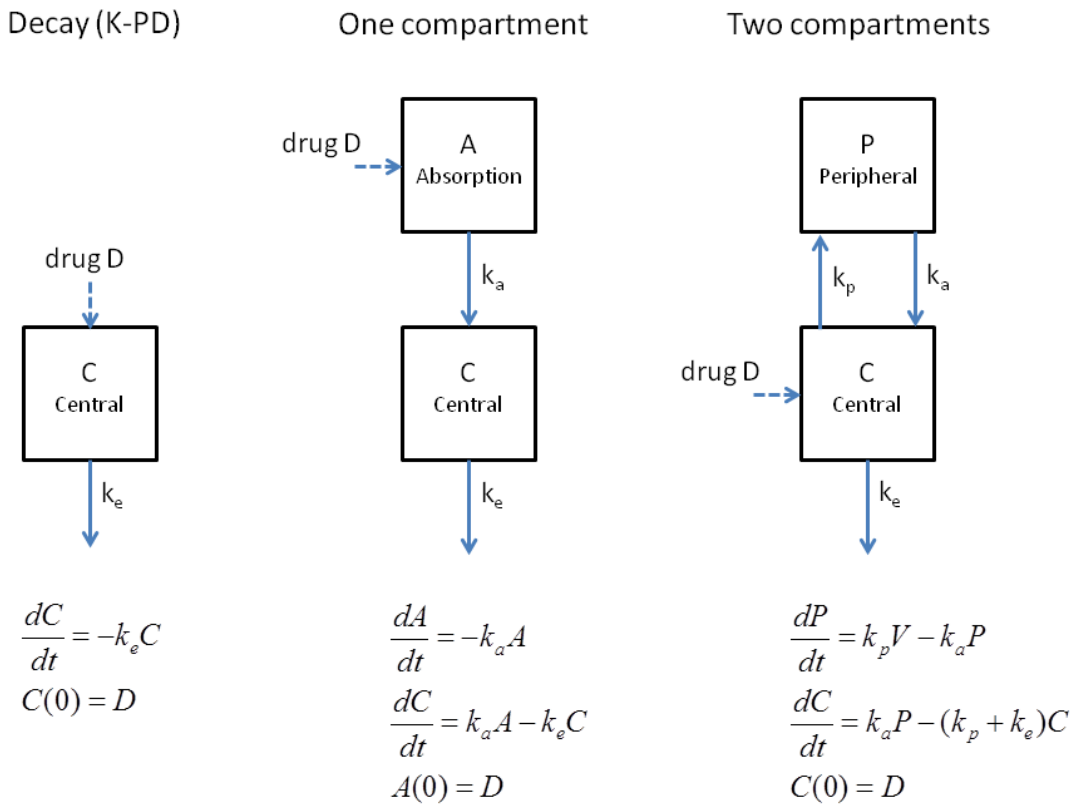

Figure 2. Schematic diagram of the compartment PK models available in the CellCycler program, with rate constants and equations.

Finally, the Tumor page uses the model simulation to generate a plot of tumor radius, given an initial radius and growing layer. A table is shown giving total radius gain; the maximum gain that would be obtained in the absence of drug; the radius loss due to drug; and the proportions of this loss that are due to death or cell damage.

The results can be compared with experimental results by using the "Read data from file" option. This should be a text file containing a matrix, where the top row is times in hours, and the other rows are individual experiments with the radius at each time. If the user selects "show linear fit" then the plot will include a linear interpolation, with estimates for initial volume and growing layer thickness given in the title. If the data corresponds to control curves with no drug, then these values can be used to parameterise the model.

For cases with drug, a two-piece linear fit to the data should be done separately. The first line segment corresponds to growth before drug, with a slope corresponding to that of the control,

and the second segment corresponds to when growth has resumed after treatment, which again should have the same slope but a lower y-intercept. The vertical distance between the two line segments can then be used to parameterise the radius lost due to drug.