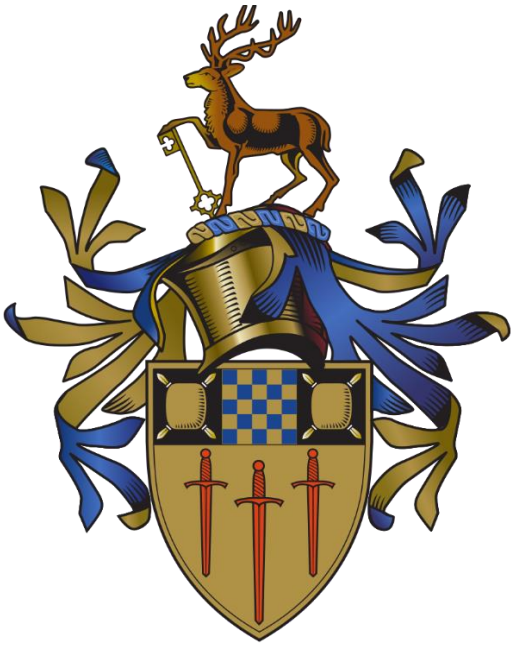
University of Surrey Master's Year Project

A Comparison of Tumour Growth Models



Author: Jonathan Evans
URN: 6250902

Supervisor: Dr Carina Dunlop

Submitted January 2017
Submitted as a report for the 6 month project for the MMath with Integrated Placement degree at the University of Surrey.

Executive Summary

The following report details the work completed during a six month industrial placement at the pharmaceutical company AstraZeneca, between July 2016 and December 2016. The objective of the work was to study a selection of tumour growth models, to better understand the applications of them within the field of Oncology.

The work can be split into 5 main sections; a review of Oncology and the problems currently within the field, a comparison of well-known tumour models by fitting them to a group of growth datasets, a comparison between two similar models to approximate an analytical solution to an unsolvable model, an analysis of a model of a protein phosphorylation system, and a study into the theory of modelling tumours as ellipsoids.

Table of Contents

Figures

Tables

1 Introduction The Growing Importance of Minimising the Rate of Failure of Late-Stage Trials, and of Accurately Translating Animal Model Data to Human Clinical Data

One of the major problems facing the pharmaceutical industry to this day is the high failure rate of clinical trials, with 85% failing at the early stages (Mak. I. W, 2014) and the majority then failing at the later phases. With regards to oncology specifically, 62.5% of trials have insignificant conclusions (Gan. H. K, 2012) where the results did not match the predictions for the trial; the predictions are often made overly optimistically with regards to the potency of the drug's effect which leads to disappointing conclusions in a large number of cases. Despite promising results during pre-clinical trials the vast majority of new products fall down later on in the testing process, which leads to the loss of hundreds of millions of pounds, several years of work, and ultimately the question of what went wrong.

1.1 Translating Animal Model Data to Clinical Data

One of the more frequently reoccurring themes when researching the failure rates of clinical trials is the poor focus on the differences between the pharmacokinetics and pharmacodynamics of humans and other animals, as well as the distinct biological differences too. Given recent advancements in scientific knowledge and techniques, predictions of the pharmacokinetics effects of a compound on humans are now far more accurate and transitioning from animal to human data is more successful. One method to translate the data is via an allometric scaling process (Mager. D. E, 2009) (West, Brown, & Enquist , 1997); it can shown that most biological processes can be described by a typical allometric scaling law $Y = Y_0 M^b$, where Y is the biological variable, Y_0 and b are constants, and M is the body mass.

Another approach, studied by Wong et al. (Wong. H, 2012), was to take the preclinical data and use it to estimate the pharmacokinetic-pharmacodynamic relationship for the compound, then combine this with the human pharmacokinetic data to create the simulated clinical trial data. The strength of the correlation between the simulated data and the observed data was then tested using the Spearman rank correlation coefficient; there was a significant correlation between the simulated and clinical data, which strongly suggests that this method can effectively help simulate the clinical trials. This suggests that pre-clinical data may provide better support to clinical trials if adequate precision and attention is put into considering the differences between animals and humans, and highlights the importance of mathematical modeling to the study of cancer. A major flaw in this process was that the human pharmacokinetics was gained from previously written literature, so is less applicable to new treatments, but the paradigm it sets is one which is easily transferable. By combining this with the pharmacokinetic data gained from other methods such as Phase 0 testing (reviewed later), this method could be used to more accurately predict the data from clinical trials.

Whilst mice and humans are similar in some aspects, there are others in which they differ which greatly affect the reliability of results gained from human tumours to grafted onto mice. Both mice and humans have similar organ systems in their body with a similar structure, and they also have similar genetics and physiology (Mestas. J, 2004) (Rangarajan. A, 2003); these biological similarities make mice theoretically a

good animal to model human tumours with, however small differences are the main reason for errors when modelling. Despite the similarities, mice are naturally susceptible to different types of tumours to humans (Anisimov. V. N, 2005), e.g. mice are more likely to develop sarcomas and humans are more likely to develop carcinomas; this difference could mean that a mice's body is not naturally adapted to cope with human tumours as well as a human would be, so would react differently to it. In addition to that, the carcinogens for mice and humans are different, meaning tumours that are triggered and cultivated by a carcinogen within the human body would not grow the same within a mouse, so the growth of a certain cell line in mice would not necessary accurately represent the growth of the same cell line in a human. As an alternative to standard mice, transgenic mice can be used to generate results which are more readily applicable (Jong. M, 2019); these are mice which have been genetically altered in order to display a different set of biological characteristics, and in the oncological setting the mice can be engineered to have a more similar carcinogenesis to humans than they would naturally. These mice share more biological traits with humans then before, and the results gained from using them are more translatable because of this, however due to the cost of this genetic modification as well as a series of other factors these type of mice are not frequently used in pre-clinical trials.

Other factors which vary between species make the extrapolation of data from murine models to human clinical trials difficult, such as toxicity. A study into the effects of 3 different drugs (Singh. B, 1985), Mithramycin, Chromomycin A_3, and Olivomycin, on a variety of species of animals and humans showed that the toxicity of a drug varied greatly across species, with some species showing high resistance where humans where sensitive. In the case of Olivomycin, the maximum tolerated dose was 8.4 mg/kg for mice and 0.3mg/kg for dogs, showing that the toxicity of a drug on a species is not only size dependent and is influenced by other biological factors. These differences make the animal study data more difficult to extrapolate, as the toxicity is species specific so the maximal dosage of a drug for mice may give little indication as to what the maximal dosage would be for humans.

Also, mice cells develop into tumours in less steps than in humans and the tumours then grow at a faster rate in mice as well (Hahn. W. C, 2002). Due to this, even after a drug has been introduced to the system, the tumour grows faster in the mice model and would recover quicker from the effect of the drug (subject to the differences in pharmacokinetics).

Although the use of subcutaneous murine models currently dominates the field when it comes to xenographs, it has also been shown that mice are not the most similar animals to humans in terms of pharmacokinetics (Sakai. I. S, 2014). However due to their relative availability for pre-clinical trials they're used over any others, even though they don't produce the most useful results. Mice models are the most easily accessible and widely used type of animal model, and while the data gained from pre-clinical trials is not perfect it still provides a valuable insight into how clinical data may turn out, however developing the ways in which the data is translated into usable clinical data is critical in minimizing failure rates. A stronger focus on understanding the differences and how to compensate for them, will help to narrow the gap between useful data and flawed data, saving both money and time as well as producing better results across the field.

Moving on from the issues with pre-clinical data and murine models, we now consider the reproducibility of trial data. Begley and Ellis (Begley. C. G, 2012) stated that the major reason for the high failure rate is the current standards of preclinical trials, and the unreproducible nature of many results. In a review of 53 studies with between 3 and 1909 citations each, only 11% of them were

reproducible, and any inaccuracies found in the data of the irreproducible ones were hence dispersed across the scientific community in the secondary publications. The reason behind this result is contributed partially to the high demand for results from the publishing authors, and the strong competition to gain funding for trials; this competition leads to data being specifically selected based on its ability to prove a result regardless of whether it's representative of the general effects on a large population, and when this skewed data is used to support clinical trials it often falls down at a later point. The pressure on authors to produce succinct, relevant, and moreover successful results overcomes the presumed scientific integrity of a paper that most would assume when reading it; the majority of the audience would hold a paper for what it says, and trust that anything published is trustworthy. Due to this the number of flawed papers in circulation is constantly increasing at an also increasing rate, with inaccurate papers being cited and creating further inaccurate papers.

A limitation on a lot of published work is the lack of rigor implemented within the pre-clinical trials, with details such as cells being taken from different cell lines and the results compared without this taken into consideration, or chosen cell lines being poorly characterized and not suitable to represent the patient population (Gan. H. K, 2012). Many publications also fail to properly account for the change in the pharmacokinetics between the mice and humans, the differences in which vastly effect a body's ability to process a drug and leading to errors within clinical trials as discussed earlier.

The obvious solution to this is to raise the standards of pre-clinical trials, and drive towards producing more accurate and transferable data to increase the success rate of later clinical trials. One method to achieve this would be to run co-clinical trials (Johnson, 2012), considering murine models alongside clinical trials; both use tumours from the same cell lines, and the data from the murine models is then transferred directly to provide the correct treatments to the human patients. As stated by Johnson (Johnson, 2012), this method of experimenting allows for better coordination and analysis of the results, both of which are critical in improving the standards of results. This kind of experimentation allows for better understanding of the biological systems involved in tumour growth, and the links between the effects on xenografts and human patients, which allows for better concordance between future trials; understanding the systems better means future pre-clinical results can have a greater effect on the outcome and direction of later clinical trials.

1.2 Minimising Late-Stage failures

As mentioned before, one of the major contributing factors in high failure rate is due to the fact that negative results in pre-clinical trials is not often published . Studies often overestimate the effectiveness of a drug (Sena. E. S, 2010), meaning the data around which clinical trials are based is bias as to the effectivity of the drug. In a study done on the relationship between in vivo and in vitro models and early clinical trials, the correlation between pre-clinical and early clinical trials in the majority of cases was statistically insignificant (Johnson. J, 2001), once again showing that the preclinical data doesn't accurately reflect what to expect from the clinical trials. This lack of correlation shows the real need for an increase in the standards of pre-clinical data published.

There are several methods gaining quick popularity in an attempt to minimize the number of late stage trial failures, including Window trials and Phase 0 testing (Cook. N, 2012). The former involves administering a new compound for a short period of time before then putting a patient on a standard therapy; this allows to test to see if a new compound will have any effect on patients before sending it

through pre-clinical and clinical trials, having large potential to avoid any losses whether monetary or otherwise. Conversely it can also push a compound onto further tests, and can save a compound from being discarded altogether (Glimelius. B, 2011). The latter involves administering a small group of patients with a sub-therapeutic size dose (Cook. N, 2012) to evaluate how a new compound reacts with the human body, and obtain basic pharmacokinetic and pharmacodynamic information (Angelo. L. S, 2009) which can be then used in tandem with animal model data to better understand the mechanisms behind the compound, as well as estimate the optimal dosing schedule.

As of recent years more and more studies are calling for various branches of academia and industry to work together in order to compile accurate data sets for pharmacokinetics and other useful data (Kleiman. R. J, 2016), which would in turn help the interpretation of the pre-clinical data and in theory raise the success rate of the later clinical trials.

1.3 Tumour Growth Models

One of the most useful techniques when predicting the results from clinical trials is the application of mathematical modelling; models can be used to describe systems such as human carcinogenesis (Moolgavkar, 1983), drug pharmacokinetics, and tumour growth, the latter of which allow simulations to predict how tumours will react to certain dosing schedules and stimulants. Due to the flexibility of parameters used in most models, modelling is arguably the most useful technique in drug prediction, however as there is an abundance in available growth models choosing the right one a difficult and important choice to make; different models often display an array of characteristics and can be more suited to certain types of cell lines (Marusic, 1994), drugs, and hosts for the tumours. To choose the most appropriate model to simulate clinical data, the models most be properly understood, from the relevance of the individual parameters to the limitations of the model and how it grows over both short and long time periods. In order to transfer the pre-clinical data to accurate and reliable clinical predictions it's important to choose the model which best replicates the data gathered, so studying the various available models would decrease the likeliness of late-stage failure.

Since the 1990's, there has been a large increase in the number of mathematical papers focusing on tumour growth models (Aranujo & McElwain, 2004), with both developments on existing models and new models being formulated constantly. In a lecture presented at the Mathematical Colloquium in Osijek, Miljenko Marušić (Marusic, Mathematical Communications, Vol. 1, No. 2, 2006) showed how a few models can be reduced to others by choice of parameter.

There are different categories of model, which each focus on different aspects of growth; these include empirical models which focus on standard growth laws, spatial/structural models which focus on the shape of tumours as a basis for the model, and models which are based on biological factors such as the doubling time or the fraction of the tumour which is growing (Marusic, 1994). Each of these categories prioritises different factors of growth and illustrates the diverse ability of models to simulate how tumours grow, furthermore studying the links between these models and how they react goes towards better understanding cancer and how it's affected by different stimuli.

Modelling gives standardized parameters which allow for easy comparison, and with a surge in the amount of experimental data being produced each year (Quaranta , 2005) the ability to interpret this data is more crucial. Whilst the current available models aren't perfect at describing systems, with assumptions taken in them often causing exponentially growing errors over time, mathematical

techniques are becoming a more regular part of experimental oncology there's much to suggest that they will play a large part in future of advances in Oncology (Gatenby & Maini, 2003).

1.4 SUMMARY

The call for more reliable pre-clinical data is becoming ever more dominant across both the academic and industrial side of the disease and medicinal field; growing frustrations at the number of late stage trial failures have led to a large portion of people know looking to increase the validity of preclinical data before clinical trials begin, with more emphasis being put on translating the differences in pharmacokinetics than before. The work done by Wong shows the benefits of having more accurate pharmacokinetic data, and although the fitting didn't work as well for very small or very large doses it was accurate for the more general case though a more in-depth study would be needed to comment with conviction the strength of this method.

The rise of methods such as Phase 0 testing and window testing provides a much earlier first glimpse into both the effectiveness of a new compound on tumours and its side effects on the body, providing the necessary backing to push forward with clinical trials or not when twinned with pre-clinical data. Having this better distinction in the strength of a trial from an earlier stage should reduce the risk of failure later on, and allow for either a revision of a molecule and its effects or a scrapping in the cases where it doesn't look likely to succeed.

In addition to this, fully understanding the mechanics of the many available tumor growth models in order to pick the most suitable one to fit pre-clinical data and make predictions for clinical trial results is crucial in successfully planning clinical trials and maximizing the chance for success. The available models range from simple ones with a small number of parameters to more complex ones, and the goodness of fit given by each model to a set of data can vary greatly; in many cases using a more complicated model doesn't result in a better fit for the data than a simpler model. Studying how each model works, and compliments different PK effects, means a larger variety of different cases can be effectively treated, and having this broader spectrum of treatable cases means a better understanding of cancer and how it grows within the body, which is one of the primary goals of any research into cancer.

Focusing on the accuracy of pre-clinical data and the methods used to translate the results of animal model data to human model data, as well as the interpretation of tumour growth models, should be the focus of the industry and academia in order to get the best results possible and increase the chances of success from all clinical trials. At the present time this means working to better interpret the differences between the biological systems of mice and humans, as well as how to accurately understand the results gained from pre-clinical trials and how to use them as predictors for clinical trials. The first steps towards sharpening overall results is to work from the beginning stage to the end, improving on the way, and in the case of clinical trials this means first understanding what animal models tell us, and how we can use the information from them.

2 A COMPARISON OF TUMOUR GROWTH MODELS

2.1 ABSTRACT

Given the large number of available models to simulate tumour growth, investigating which ones are best at fitting clinical data and looking and common factors in the best fitting ones could give insight into what mathematical techniques are best at representing tumour growth. Here we consider five models which are commonly used in the literature and compare their effectiveness at describing observed tumour growth data.

2.2 INTRODUCTION TO PARAMETER ESTIMATION

For each of the models available to simulate tumour growth, there are a number of parameters which control aspects of the rate of growth and decay in order to generate growth curves. Given a set of data $\{X_i\}$, for each model there exists a set of parameters which provide the best fit, and comparing the error between the data and the models for these optimal parameters provides a strong indicator as to which model was most suitable for simulating the data.

2.3 DATA

The data used came from five trials from 2014, each of which used a unique cell line and measured growth under two different conditions (with the exception of GCLS1433 which measured 4 different conditions). These sets of data were chosen to give a wide range of growth patterns and cell lines, as the aim is to test which model fits best for a generic set of data. Within the trials, data sets have been omitted when either the data was incomplete, the recorded data was unsuitable for parameter fitting (i.e the volume recorded stayed at a constant size for a long period of time, then dropped rapidly and reached a second constant state. This causes difficulties for the models, and is unrealistic at representing tumour growth), or could not be fit to all the models.

The datasets which couldn't be fit to all the models were omitted to allow consistency during the comparison.

2.4 THE MODELS

The models were chosen to represent a variety of different systems, and come from three papers (Marusic, 1994) (Jumbe & et al, 2010) (Evans, Dimelow, & Yates, 2014). The first of these papers (Marusic, 1994), is a comparison of the ability of 14 models to fit data, however there are a few differences between the methodologies used there and those used in this study. In (Marusic, 1994), an F-test was used to compare the logarithmic least sum of squared errors $(LLSSE)$ of each model for each data set, meaning the differences were considered for each individual dataset whereas here the mean of the error across the datasets is compared. Marušić fits the models to 9 total datasets, making the population size relatively small.

2.4.1 The Gompertz model

2.4.1.1 The Standard model

The first model looked at is one of the more common growth models, the Gompertz model,

which is given by;

$$\frac{dV}{dt} = aV - b\,\mathrm{Vlog}\,V \qquad (1)$$

The a coefficient is used to describe exponential growth rate, as in the case $b = 0$ the system reduces to $\frac{dV}{dt} = aV$ which has a solution $V = Ce^{at}$, where c is an arbitrary constant; The b coefficient is the growth limiting term, and is used to put an upper bound on the growth for it to tend towards over time which creates a new steady state.

The steady states of this model are $V = 0$, and $V = e^{\frac{a}{b}}$. To find the stability of these states, consider the function $f(V) = aV - b\,\mathrm{Vlog}\,V$, then $f'(V) = (a-b) - b\log V$. $f'\left(e^{\frac{a}{b}}\right) = -b < 0$, so this steady state is always stable; $f'(0)$ is undefinable, so its stability can't be determined via this method however it can be determined by logical reasoning; Take a small increment $\delta V < 1$, then $\log \delta V$<0 and $\left.\frac{dV}{dt}\right|_{V=\delta V} = a\delta V + b\delta V|\log \delta V| > 0$, so the model is constantly increasing at this point, therefore $V = 0$ cannot be a stable steady state.

If a and b are both positive, $e^{\frac{a}{b}} > 1$ so the model cannot have a steady state of size smaller than 1. To counteract this feature of the model, an additional death term must be added to make this model suitable for all sets of data.

A solution for (1) can be found by separation of variables as follows;

$$\frac{1}{V(a-b\log V)}\,dV = 1\,dt \qquad (2)$$

Integrating (2) with respect to *t* gives;

$$\frac{-1}{b}\log(a - b\log V) = t + C,$$

The constant C can be solved with the initial condition $V(0) = V_0$ to give $C = \frac{-1}{b}\log(a - b\log V_0)$, which is then substituted into the equation and rearranged for V;

$$a - b\log V = (a - b\log V_0)e^{-bt}$$

$$\Rightarrow V = e^{\frac{1}{b}\left(a-(a-b\log V_0)e^{-bt}\right)} \qquad (3)$$

So the model has an analytical solution, given by (3).

2.4.1.2 The Gompertz model with a death term

With an additional term to represent the death rate of tumour cells, the Gompertz model becomes $\frac{dV}{dt} = aV - b\,\mathrm{Vlog}\,V - DV = (a-D)V - bV\log V$. This can be rewritten as $\tilde{a}V - b\,\mathrm{Vlog}\,V$, where $\tilde{a}$ is no longer necessarily positive. The steady states of this equation remain the same as before, as does the solution to the model, with $\tilde{a}$ replacing a.

2.4.2 The Generic model

The Generic model is defined as $\frac{dV}{dt} = \frac{\beta V^{1-np}(k^n - V^n)^{1+p}}{k^n}$. The parameter β is a growth rate term, which describes how quickly the model reaches its steady state; the parameter k defines the non-zero steady state of the model, but has no effect on how quickly the function reaches that limit; the parameter n describes the shape of the curve, large n leads to large initial growth whereas small n gives exponential growth to begin with, which then levels out towards the steady state of the model; the parameter p does similar to n, however it has no effect on the non-zero steady state. When n is a whole number, there are $n + 1$ steady states defined by $V = 0$ and $V = ke^{i\frac{2\pi}{|n|}\alpha}$ where $\alpha \in \{1, \dots, n\}$, however with the exception of $\alpha = n$ (and $\frac{n}{2}$ if n is even) all of the exponential solutions are complex. The latter of the exceptions, where applicable, gives a negative value for V so the only real and useful steady states of the model are $V = 0$ and $V = k$, so as with the Gompertz model it will slowly tend towards this final value of $V = k$. Taking $f(V) = \frac{dV}{dt}$ gives $f'(V) = \frac{\beta V^{-np}}{k^n}\left(k^n - V^n\right)^p\left[(1 - np)\left(k^n - V^n\right) - n(1 - p)V^n\right]$. Using this gives $f'(0) = f'(k) = 0$ so this doesn't say much about the stability of either of these points, however this model is only decreasing when $V > k$, and increasing otherwise, making k a stable steady state and further making it that for decreasing data, the Generic model can still provide a good approximation for decreasing data, provided k is smaller than the smallest value of the data. When n is not a whole number, the model only has a zero steady state. p is the only parameter in any of the models which is allowed to be negative.

The Generic model has an analytical solution given in (Tsoularis & Wallace, 2002) by;

$$V(t) = \frac{k}{\left[1 + \left[pn\beta K^{1+(1-n)p}t + \left[\left(\frac{k}{V_0}\right)^n - 1\right]^{-p}\right]^{\frac{1}{-p}}\right]^{\frac{1}{n}}}$$

2.4.3 The Generalised Two-Parameter model

The Generalised Two-Parameter model is given by;

$$\frac{dV}{dt} = aV^\alpha - bV^\beta \qquad (4)$$

The parameters a and b in this are growth and death rate terms respectively whereas α and β effect the initial growth phase and how the model approaches its limit . The steady states of this equation are $V = 0$ and $V = \left(\frac{a}{b}\right)^{\frac{1}{\beta-\alpha}}$, the stability of which is determined by $f'(V) = V^{\alpha-1}\left(\alpha a - \beta b V^{\beta-\alpha}\right)$; $f'(0) = 0$, so this doesn't tell much about this steady state, however $f'\left(\left(\frac{a}{b}\right)^{\frac{1}{\beta-\alpha}}\right) = a\left(\frac{a}{b}\right)^{\frac{\alpha-1}{\beta-\alpha}}(\alpha - \beta)$, so this state is stable if $\beta > \alpha$. This model can again handle decreasing data sets as well as increasing ones.

The Generalised Two-Parameter model has a solution derived by Miljenko Marušić and Željko Bajzer (Marusic & Bajzer, 1993), and defined by 3 separate cases[1];

For $v_0 > 0, ab > 0$

$$V(t) = \left(\frac{a}{b}\right)^{\frac{1}{(\beta-\alpha)}} [1 - {\gamma_\delta}^{-1}(\gamma_\delta(v_0) + k(t - t_0))]^{\frac{1}{(\beta-\alpha)}} \quad (5)$$

For $v_0 = 0, ab > 0$

$$V(t) = V_0 \quad (6)$$

For $v_0 < 0, ab > 0$

$$V(t) = \left(\frac{a}{b}\right)^{\frac{1}{(\beta-\alpha)}} \left[1 - {\gamma_{-1-\delta}}^{-1}\left(\gamma_{-1-\delta}\left(\frac{v_0}{{v_0}^{-1}}\right) + k(t - t_0)\right)\right]^{\frac{1}{(\alpha-\beta)}} \quad (7)$$

Where $v_0 = 1 - \frac{b}{a}{V_0}^{\beta-\alpha}$, $k = -b(\beta - \alpha)\left|\frac{a}{b}\right|^{\frac{\beta-1}{\beta-\alpha}}$, $\delta = \frac{\beta-1}{\beta-\alpha}$, and $\gamma_\delta(x) = \int_{\frac{1}{2}}^{x}(v-1)^\delta v^{-1}\,dv$.

2.4.4 The Autostimulation Model

The Autostimulation model is composed of two separate differential equations, $\frac{dV}{dt} = \alpha V\frac{(1+S)}{(1+\beta V)} - \omega V$, and $\frac{dS}{dt} = aV - bS^2$, where $S(0) = 0$. The parameter a controls the growth rate of S, which in turn effects the growth of V; the parameter b controls the decay rate of S, and also effects the decay of B; the parameter α is the growth parameter for V; the parameter β is the decay parameter for V; the parameter ω adds a restriction on the growth given by the steady state below. The steady states of this are $V = 0$, and $V = \frac{1}{\omega\beta}(a(1+S) - \omega)$. The second of these changes with respect to changes in S, however using these steady states in $\frac{dS}{dt}$ gives the steady states of the system as;

$$(V, S) = (0,0)$$

$$(V, S) = \left(\frac{1}{\omega\beta}\left(a\left(1 + \frac{a\alpha}{2b\beta\omega} \pm \sqrt{\frac{a\alpha}{b\beta\omega}\left(1 + \frac{a\alpha}{b\beta\omega}\right) - \frac{a}{\beta}}\right) - \omega\right), \frac{a\alpha}{2b\beta\omega} \pm \sqrt{\frac{a\alpha}{b\beta\omega}\left(1 + \frac{a\alpha}{b\beta\omega}\right) - \frac{a}{\beta}}\right)^2$$

The stability of which is defined by $f'(V) = \frac{a(1+S)}{(1+\beta V)}\left(1 - \frac{\beta V}{(1+\beta V)^2}\right) - \omega$;

$f'(0) = a(1+S) - \omega$, and $f'\left(\frac{1}{\omega\beta}\left(a\left(1 + \frac{a\alpha}{2b\beta\omega} \pm \sqrt{\frac{a\alpha}{b\beta\omega}\left(1 + \frac{a\alpha}{b\beta\omega}\right) - \frac{a}{\beta}}\right) - \omega\right)\right) = \frac{\omega^2}{a(1+S)}\left[\frac{\omega}{a(1+S)} - 1\right]$

which is stable if $a\left(1 + \frac{1}{\omega\beta}\left(a\left(1 + \frac{a\alpha}{2b\beta\omega} \pm \sqrt{\frac{a\alpha}{b\beta\omega}\left(1 + \frac{a\alpha}{b\beta\omega}\right) - \frac{a}{\beta}}\right) - \omega\right)\right) > \omega$.

[1] There is a fourth solution for the case , $ab < 0$, however this paper is focusing on models with positive parameters so this case is omitted.

[2] For the negative case, this steady state only holds if $\frac{a\alpha}{b\beta\omega}\left(1 + \frac{a\alpha}{b\beta\omega}\right) - \frac{a}{\beta} \geq 0$

The model has no analytical solution.

2.4.5 The Generalised Bertalanffy-logistic Model

The Generalised Bertalanffy-logistic model is a 3 part model which changes based on the value of one its parameters. It can be summarized as the following;

$$\frac{dV}{dt} = \begin{cases} aV^{\alpha} - bV; & \alpha < 1 \quad (8) \\ aV - bV\log V\,; & \alpha = 1 \quad (9) \\ aV - bV^{\alpha}; & \alpha > 1 \quad (10) \end{cases}$$

with stable states given by $\begin{cases} V = 0, V = \left(\frac{b}{a}\right)^{\frac{1}{\alpha-1}}; \\ V = 0, V = e^{\frac{a}{b}}; \\ V = 0, V = \left(\frac{a}{b}\right)^{\frac{1}{\alpha-1}}; \end{cases}$. The stability function is given by $f'(V) =$

$\begin{cases} a\alpha V^{\alpha-1} - b; \\ a - b - blog(V); \\ a - b\alpha V^{\alpha-1}; \end{cases}$, which gives $\begin{cases} -b \\ a - b \\ a \end{cases}$ for the zero steady states and $\begin{cases} b(\alpha - 1); \\ -b; \\ a(1-\alpha); \end{cases}$ for the non-zero steady states. For the zero steady states, the stability is $\begin{cases} stable; \\ stable; \\ unstable; \end{cases}$ $if\ b > a$, and for the non-zero states all of them are stable.

The solution to the Generalised Bertalanffy model is split into 3 sections. (9) is identical to the Gompertz model so has the same solution, whereas the other two cases need to be solved separately.

(8) can be solved as follows;

$$\frac{1}{V(aV^{\alpha-1} - b)}\frac{dV}{dt} = 1$$

Separating the fractions gives;

$$\left(\frac{-1}{bV} + \frac{aV^{\alpha-2}}{b(aV^{\alpha-1} - b)}\right)\frac{dV}{dt} = 1$$

Integrating with respect to t gives;

$$\frac{-1}{b}\ln V + \frac{1}{b(\alpha-1)}\ln(aV^{\alpha-1} - b) = t - t_0 + c$$

$$\frac{1}{b(\alpha-1)}(-\ln V^{\alpha-1} + \ln(aV^{\alpha-1} - b)) = t - t_0 + c$$

$$\ln\left(\frac{aV^{\alpha-1} - b}{V^{\alpha-1}}\right) = b(\alpha-1)(t - t_0 + c)$$

$$a - bV^{1-\alpha} = Aexp\big(b(\alpha-1)(t - t_0)\big) \qquad (11)$$

Solving (11) for the condition $V(0) = V_0$ gives $A = \left(a - bV_0{}^{1-\alpha}\right)$, which can be substituted and rearranged to give;

$$a - bV^{1-\alpha} = \left(a - bV_0{}^{1-\alpha}\right)exp(b(\alpha - 1)(t - t_0))$$

$$V = \left(\frac{a}{b} - \left(\frac{a}{b} - V_0{}^{1-\alpha}\right)exp(b(\alpha - 1)(t - t_0))\right)^{\frac{1}{1-\alpha}}$$

For (10), the model can be rewritten as $\frac{dV}{dt} = aV - bV^{\alpha} = -(bV^{\alpha} - aV)$, which is the negative of (8) with the parameters interchanged, so the solution can be written as;

$$V = -\left(\frac{b}{a} - \left(\frac{b}{a} - V_0{}^{1-\alpha}\right)exp(a(\alpha - 1)t)\right)^{\frac{1}{1-\alpha}}$$

So the set of solutions for the Generalized Bertalanffy is given by;

$$V(t) = \begin{cases} \left(\frac{a}{b} - \left(\frac{a}{b} - V_0{}^{1-\alpha}\right)exp(b(\alpha - 1)(t - t_0))\right)^{\frac{1}{1-\alpha}}; & \alpha < 1 \\ e^{\frac{1}{b}\left(a - (a - b\log V_0)e^{-bt}\right)}; & \alpha = 1 \\ -\left(\frac{b}{a} - \left(\frac{b}{a} - V_0{}^{1-\alpha}\right)exp(a(\alpha - 1)(t - t_0))\right)^{\frac{1}{1-\alpha}}; & \alpha > 1 \end{cases}$$

This model is a combination of two specific cases of the Generalised Two-Parameter model and the Gompertz model, so is expected to perform similarly to both of them.

2.5 Fitting the Parameters

Given all the models can approximate increasing and decreasing sets of data, they can all be individually fitted to data and the results then compared to see if any of the specific models out-perform the others in general. The data used for this will be taken from the 2014 trials, GCATR1434, GCERK1420, GCGLS1404, GCERK1421, and GCGLS1433, which were all examining the growth of xenografted tumours in mice under laboratory conditions by measuring the length, width, volume and weight of the tumours. First each model will be fit to the data sets from each trial to get the optimal fit parameters, then the cumulative error calculated from each. The mean of the cumulative errors of each model will be compared to see if there is a significant difference in any of them.

Of the data contained within these trials, a selection have been omitted due to missing data points and incorrectly recorded data points, both of which would skew the parameter fit process so in order to properly analyze the performance of each model only the appropriate data sets were used.

The individual parameter values for each model are located within the appendix, Section 8.1.1.

2.6 MATHEMATICAL METHODS

2.6.1 Parameter optimisation

The method used to optimize the parameters will be to minimize the sum of the residuals squared between the observed data and the simulated data, i.e for the data set $\{X_i\}$, with $i \in \{1, \dots, n\}$ and the time at X_i being t_i, and the set of parameters $\{\tau_j\}$ such that $\tau_j > 0 \ \forall j$[3], the sum of the residual squares is given by $Residuals_{\frac{dV}{dt}|_{\{\tau_j\}}} = \sum_{i=1}^{n}(f(t_i) - X_i)^2$, where $f(t_i)$ is the solution for the differential system for $\{\tau_j\}$.[4] Taking the residuals squared makes the assumption that the measurement errors for the optimized fit are normally distributed.

There are several different algorithms available which can perform parameter optimization, and while some algorithms perform extremely well in specific cases, they aren't able to find viable solutions for all types of data so aren't suitable for this work. The method which was chosen was MATlab's built-in "*fminsearch*"; *"fminsearch"* is a function which outputs the set of parameters which minimizes a selected function subject to those parameters, so to fit the models to data a function was set up to calculate the sum of the residuals between the model simulation and the data. This function was then minimized w.r.t the models parameters. The major drawback of using this function is that it has no constraints on the output so it can output negative parameters, however this can often be averted by changing the initial values for each parameters. This happens as the algorithm for optimization finds a set of parameters which minimizes the error to below a certain tolerance, at which point the program ends, however if there is more than one set of parameters which give error below the tolerance then the program will output whichever it finds first.

2.6.2 Paired t-test

To compare the models, the optimal parameters for each model across each data set is recorded with the sum of the residual squared for those parameters, and the set of residuals from two models were compared in a paired t-test to determine if there is substantial evidence to suggest the mean error of the two models aren't equal. A paired t-test is used as the residuals from both models are determined from the same datasets.

2.6.3 Akaike Information Criterion (AIC)

The AIC (Burnham. K, 2011) is a way of comparing the effectiveness of models for a given set of data and the likelihood function for the model, which when comparing the residual sum of squares is $\frac{-n}{2}\ln\frac{RSS}{n}$ where n is the number of data sets considered and RSS is the residual sum of squares between the model and the data. The AIC value is then given by $AIC = 2k + n\ln\frac{RSS}{n}$, with k being the number of parameters. When the number of datasets is finite, a corrected version of the AIC can be used given by $AIC_C = 2k + n\ln\frac{RSS}{n} + \frac{(2K(K+1))}{(n-K-1)}$. From this, the difference between the AIC_C for each model and the smallest AIC_C value can be calculated, defined as $\Delta_i = AIC_{c,i} - AIC_{c,min}$, and finally the weighting of each in comparison to all of the models, given by $w_i = \frac{\exp(-\frac{1}{2}\Delta_i)}{\sum_{r=1}^{R}\exp(-\frac{1}{2}\Delta_r)}$. The weighting gives the probability

[3] The parameters are positive as defined in the literature.

[4] This method is also known as the least sum of squared errors (LSSE).

that a model will be the most suitable at fitting a given set of data. This test provides a comparison of the models which accounts for the number of parameters in the model, and gives the probability that each model is the best fit for a piece of data.

2.7 CALCULATING THE FIT

The model with the lowest difference for each dataset is highlighted in order to highlight any pattern on models which consistently outperformed the other models. In the event that one or more of the models performed the best, both are highlighted in a lighter shade of green.

2.7.1 GCATR1434

Mouse	Gompertz	Generic	Generalised Two-Parameter	Autostimulation	Bertalanffy
1	7.7237e-0.4	8.2503e-04	5.7170e-04	8.6376e-04	7.7435e-04
4	0.0012	0.0013	0.0011	0.0025	0.0012
5	0.0039	0.0049	0.0038	0.0030	0.0039
6	0.0019	0.0020	0.0179	0.0020	0.0019

Table 2-1: The difference between the optimised simulations of each model for the datasets from the GCATR1434 trial.

2.7.2 GCERK1420

Mouse	Gompertz	Generic	Generalised Two-Parameter	Autostimulation	Bertalanffy
1	0.0135	0.0162	0.0136	0.1406	0.0135
2	0.1074	0.0733	0.1015	0.0746	0.2149
3	0.0118	0.0227	0.0122	0.2149	0.0121
5	0.0074	0.0086	0.0092	0.0177	0.0080
6	0.0375	0.0160	0.0352	0.0265	0.0360
7	0.0522	0.0187	0.0498	0.1924	0.0511
8	0.0279	0.0117	0.0090	0.0141	0.0267
9	0.0198	0.0169	0.0159	0.0322	0.0188
10	0.0307	0.0309	0.0305	0.0432	0.0307

Table 2-2: The difference between the optimised simulations of each model for the datasets from the GCERK1420 trial.

2.7.3 GCGLS1404

Mouse	Gompertz	Generic	Generalised Two-Parameter	Autostimulation	Bertalanffy
1	0.0167	0.0165	0.0167	0.0160	0.0167
2	0.0994	0.0328	0.0281	0.0188	0.0980
3	0.0089	0.0085	0.0085	0.0043	0.0086
4	0.1426	0.0331	0.0276	0.0720	0.1402
6	0.0718	0.0542	0.0044	0.0053	0.0675
7	0.4011	0.1164	0.1162	0.2270	0.1178
8	0.0732	0.0724	0.0730	0.0522	0.0732
9	0.0345	0.1806	0.0341	0.0208	0.0344
10	0.0771	0.0810	0.0771	0.0615	0.0771
11	0.1102	0.1924	0.0355	0.0041	0.1054

Table 2-3: The difference between the optimised simulations of each model for the datasets from the GCGLS1404 trial.

2.7.4 GCERK1421

Mouse	Gompertz	Generic	Generalised Two-Parameter	Autostimulation	Bertalanffy
1	0.0054	0.0044	0.0014	0.0014	0.0054
2	0.0033	0.0028	0.0032	0.0249	0.0032
3	0.0010	0.0010	0.0010	0.0050	0.0010
5	0.0038	0.0031	0.0026	0.0031	0.0036
6	0.0026	0.0026	0.0021	0.0043	0.0021
8	0.0159	0.0155	0.0149	0.0113	0.0160
9	0.0034	0.0036	0.0011	0.0410	0.0036
10	0.0805	0.0943	0.0606	0.3072	0.0830
11	0.0101	0.0097	0.0053	0.0044	0.0096
12	0.0037	0.0032	0.0024	0.0320	0.0032
13	0.0317	0.0234	0.0143	0.0202	0.0207
14	0.0019	0.0015	0.0018	0.0016	0.0019
15	0.0144	0.0156	0.0049	0.0520	0.0139
16	0.0152	0.0149	0.0152	0.0454	0.0155
17	0.0023	0.5292	0.0016	0.0061	0.0023
19	0.0081	0.0082	0.0820	0.0168	0.0081

Table 2-4: The difference between the optimised simulations of each model for the datasets from the GCERK1421 trial.

2.7.5 GCGLS1433

Mouse	Gompertz	Generic	Generalised Two-Parameter	Autostimulation	Bertalanffy
1	0.0061	0.0059	0.0058	0.0124	0.0059
2	0.0038	0.0046	0.0035	0.0422	0.0035
3	0.0346	0.0388	0.0309	0.1110	0.0337
4	0.0418	0.0293	0.0213	0.0198	0.0418
6	0.0043	0.0046	0.0045	0.0050	0.0044
8	0.0033	0.0033	0.0033	0.0034	0.0033
9	8.8189e-04	6.5172e-04	6.7391e-04	0.0018	9.1969e-04
10	3.194e-04	4.6269e-04	5.3443e-04	8.3077e-05	5.0097e-04
17	0.0705	0.0996	0.0643	0.0272	0.0681
18	0.0616	0.0150	0.0116	0.0089	0.0544

Table 2-5: The difference between the optimised simulations of each model for the datasets from the GCGLS1433 trial.

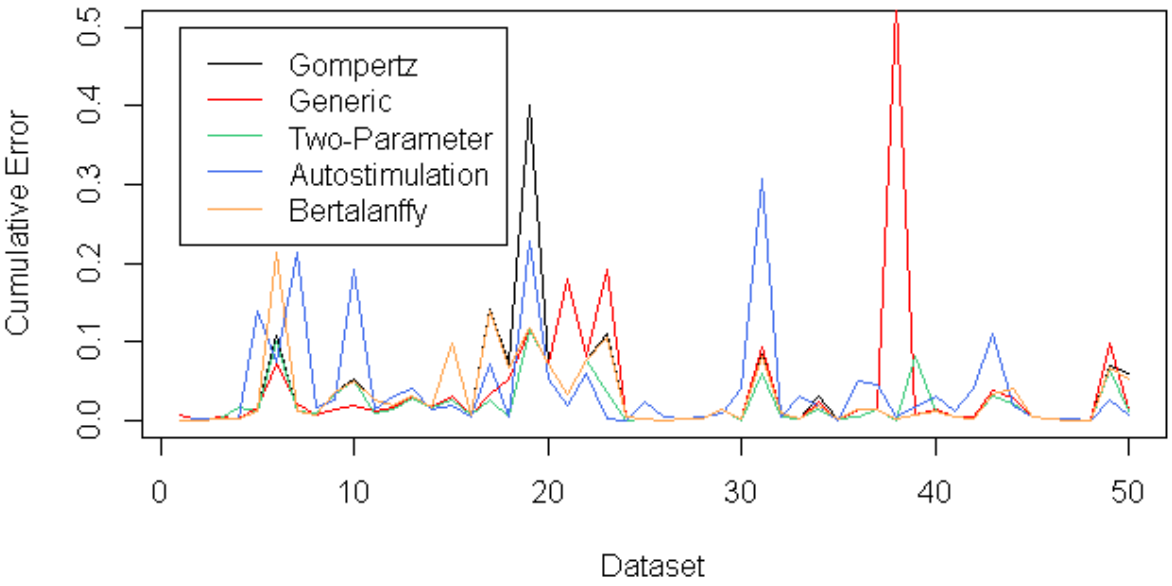


Figure 2-1: A comparison of the cumulative error of each of the models across each of the datasets.

As seen in Figure 2-1, across the majority of data the models performed very similarly with the exception of some peaks in the data, and from this it can be concluded that there is little difference between the models down to exceptional cases. The datasets are put in an order of 1-50, being datasets from GCATR1434, GCERK1420, GCGLS1404, GCERK1421, and GCGLS1433 in the order shown in Tables 2-1 to 2-5 inclusively.

2.8 COMPARING THE METHODS

The model which most frequently produced the smallest error was the Generalised Two-Parameter model, which produced either the best or joint best fit to the data 23 times out of the 50 data sets. In the majority of cases, the Gompertz and Generalised Bertalanffy Logistic models had very similar, if not equivalent values for the error, often varying by within $\pm 0.0003 mm^3$ in many cases; this variation can be seen as typical experimental errors. As seen in Section 2.4.5, for $\alpha = 1$ the Bertalanffy model is equivalent to the Gompertz model, and as seen across Appendix.a in the majority of data sets $\alpha = 1 \pm 0.025$ so seeing correlation between the suitability of fit for these models is expected in these cases, and overall the Bertalanffy model should be more accurate as it encompasses the Gompertz model with additional flexibility, however due to computational errors the Gompertz model does come out with smaller error than the Bertalanffy model in a few instances.

The mean of the cumulative differences generated by the models was;

	Gompertz	Generic	Generalised Two-Parameter	Autostimulation	Bertalanffy
Mean	0.03590547	0.03938129	0.0219096	0.04172934	0.0316119
Rank	3	4	1	5	2

Table 2-6: The mean of the cumulative differences for each of the model and their rank, from lowest to highest.

The mean of the error in each model is compared using a paired t-test, to determine whether any of the models were statistically better or worse at fitting to the data sets. Taking the Null Hypothesis as $\mu_1 = \mu_2$, and the Alternative Hypothesis as $\mu_1 \neq \mu_2$, the paired t-tests gave the following results (where non-significant means there is not sufficient evidence to assume the Null Hypothesis holds);

	Gompertz	Generic	Two-Param.	Autostimulation	Bertalanffy
Gompertz		Significant	Non-Significant	Non-Significant	Non-Significant
Generic			Non-Significant	Non-Significant	Non-Significant
Two-Param.				Non-Significant	Non-Significant
Autostimulation					Non-Significant
Bertalanffy					

Table 2-7: The results of the paired t-test, comparing whether there is significant evidence to suggest that two models have equivalent mean cumulative difference.

As Table 2-7 shows, with the exception of the Gompertz and Generic models, there wasn't sufficient evidence to suggest that the mean of the errors for any of the models was equal. Given this, the model with the lowest mean error is the one which best fit the datasets used, which was the Generalised Two-Parameter model.

Below is the AIC comparison for the models for the datasets used. The weightings show that the Generalised Two-Parameter model is statistically the most likely model to fit to one of the datasets.

	Number of parameters	RSS	AIC_c	Δ_i	w_i	Rank
Gompertz	3	0.03590547	-159.8215	24.3311	5.2061e-06	3
Generic	4	0.03938129	-154.8343	29.3183	4.3009e-07	4
Two-Parameter	4	0.0219096	-184.1526	0	0.9999	1
Autostimulation	5	0.04172934	-147.4693	36.6887	1.0792e-08	5
Bertalanffy	3	0.0316119	-166.1893	17.9633	1.2568e-04	2

Table 2-8: The results of the AIC test for each of the models.

2.9 ADDITIONAL MODELS

In addition to the more popular growth models, we consider two models which were of significant interest to AstraZeneca; the Warwick model and the Jumbe model.

2.9.1 The Warwick Model

The Warwick model is given by $\frac{dV}{dt} = kV\left(1 - \left(1 - \frac{r_d}{\left(\frac{3V}{4\pi}\right)^{\frac{1}{3}}}\right)^3\right)$ (Evans, Dimelow, & Yates, 2014), and is a model which assumes that cell replication only happens in an outer proliferating shell section of the tumour. The parameter k represents the growth rate, and r_d represents the depth of the proliferating shell. This model can be expressed as $\frac{dV}{dt} = 3k\phi V^{\frac{2}{3}} - 3k\phi^2 V^{\frac{1}{3}} + k\phi^3$ by using a Taylor expansion on $\left(1 - \frac{r_d}{\left(\frac{3V}{4\pi}\right)^{\frac{1}{3}}}\right)^3$, where $\phi = \left(\frac{3}{4\pi}\right)^{\frac{-1}{3}} r_d$, which can then be written as $\frac{dV}{dt} = 3k\phi\left[\left(V^{\frac{1}{3}} - \frac{\phi}{2}\right)^2 - \frac{\phi^2}{4}\right] + k\phi^3$. Taking $\frac{dV}{dt} = 0$ in this form, it can be seen that these are no non-zero real steady states to this model, so the growth is continuous over time.

2.9.2 The Jumbe Model

The Jumbe model is given by $\frac{dV}{dt} = KV^{\frac{2}{3}}$ (Jumbe & et al, 2010), which assumes the growth of a tumour is proportional to the surface area of the tumour, with K being the growth rate parameter. This model only has one steady state, $V = 0$, the stability of which is defined by $f'(V) = \frac{2}{3}KV^{-\frac{1}{3}}$, however at $V = 0$ this is undefinable. Due to this, the model is unbounded above so any simulations by it will continue to grow exponentially over time.

2.9.3 Parameter Fitting and Error

Below are the parameter and error values for the Jumbe and Warwick models across all the applicable datasets. None of the data sets from the GCATR1434 trials where viable here as they all contained data which decreased over time, however the data from the rest of the trials was suitable.

2.9.3.1 GCERK1420

	Jumbe		Warwick		
	K	Error	k	r_d	Error
1	0.1031	0.0648	2.2608	0.0096	0.0673
2	0.1182	0.1406	2.2648	0.011	0.1429
3	0.1113	0.0282	1.9400	0.0121	0.0367
5	0.0952	0.0286	2.1896	0.0092	0.0300
6	0.1088	0.0903	1.9325	0.0119	0.0935
7	0.1148	0.0606	0.9635	0.0259	0.0603
8	0.1130	0.0505	1.7537	0.0136	0.0527
9	0.1048	0.1487	3.0143	0.0073	0.1516
10	0.0843	0.0321	1.1987	0.0150	0.0327

Table 2-9: The difference between the optimised simulations of each model for the datasets from the GCERK1420 trial.

2.9.3.2 GCGLS1404

	Jumbe		Warwick		
	K	Error	k	r_d	Error
1	0.0405	0.0182	0.0457	0.4783	0.0166
2	0.0427	0.1357	0.0521	0.5084	0.0997
3	0.0363	0.0086	0.0839	0.1124	0.0086
4	0.0496	0.2390	0.0571	0.5463	0.1437
6	0.0502	0.0996	0.0949	0.0615	0.0005949
7	0.0742	0.1207	0.2259	0.0795	0.1188
8	0.0530	0.0817	0.0711	0.2963	0.6724
9	0.0679	0.0423	0.1240	0.1643	0.0348
10	0.0576	0.1110	0.0682	0.4556	0.0760
11	0.0572	0.2399	0.0663	0.5243	0.1107

Table 2-10: The difference between the optimised simulations of each model for the datasets from the GCGLS1404 trial.

2.9.3.3 GCERK1421

	Jumbe		Warwick		
	K	error	k	r_d	error
1	0.0214	0.0075	0.1570	0.135	0.0055
2	0.0968	0.0069	0.2617	0.0956	0.0037
3	0.0693	0.0015	1.0375	0.0143	0.0017
5	0.0916	0.0059	1.356	0.0144	0.0063
6	0.1071	0.3737	3.4969	0.0064	0.5774
8	0.0936	0.0166	0.3013	0.0748	0.0152
9	0.0936	0.0465	2.7766	0.0071	0.0477
10	0.1100	0.0921	2.1515	0.0108	0.0941
11	0.0906	0.0178	1.9294	0.0099	0.0184
12	0.0820	0.0051	1.1664	0.0150	0.0054
13	0.0967	0.0356	1.2032	0.0173	0.0357
14	0.0900	0.0120	0.1775	0.1729	0.0017
15	0.0933	0.0179	1.4590	0.0137	0.0181
16	0.0930	0.1470	2.2093	0.0088	0.1493
17	0.0846	0.0072	2.0026	0.0089	0.0076
19	0.0423	0.0045	0.1177	0.2149	0.0020

Table 2-11: The difference between the optimised simulations of each model for the datasets from the GCERK1421 trial.

2.9.3.4 GCGLS1433

	Jumbe		Warwick		
	K	error	k	r_d	error
1	0.0924	0.0148	0.4779	0.0520	0.0068
2	0.1179	0.0310	0.5855	0.0523	0.0063
3	0.1476	1.3209	0.2807	0.2128	0.2197
4	0.1117	0.5840	0.1896	0.2871	0.0761
6	0.2159	0.0162	0.6886	0.0825	0.0044
8	0.1653	0.0369	0.3055	0.2933	0.0030
9	0.1569	0.0053	0.4505	0.1029	0.00086714
10	0.1692	0.00040134	1.4626	0.0260	0.000081387
17	0.1076	0.1021	0.3418	0.0837	0.0795
18	0.1068	0.1026	0.3062	0.0968	0.0702

Table 2-12: The difference between the optimised simulations of each model for the datasets from the GCGLS1433 trial.

2.9.4 Comparing the models

	Gompertz	Generic	Generalised Two-Parameter	Autostimulation	Bertalanffy	Jumbe	Warwick
Mean	0.03895655	0.04244814	0.02330888	0.04516311	0.03419175	0.1035465	0.07413442
Rank	3	4	1	5	2	7	6

Table 2-13: The mean of the cumulative differences for each of the model and their rank, from lowest to highest.

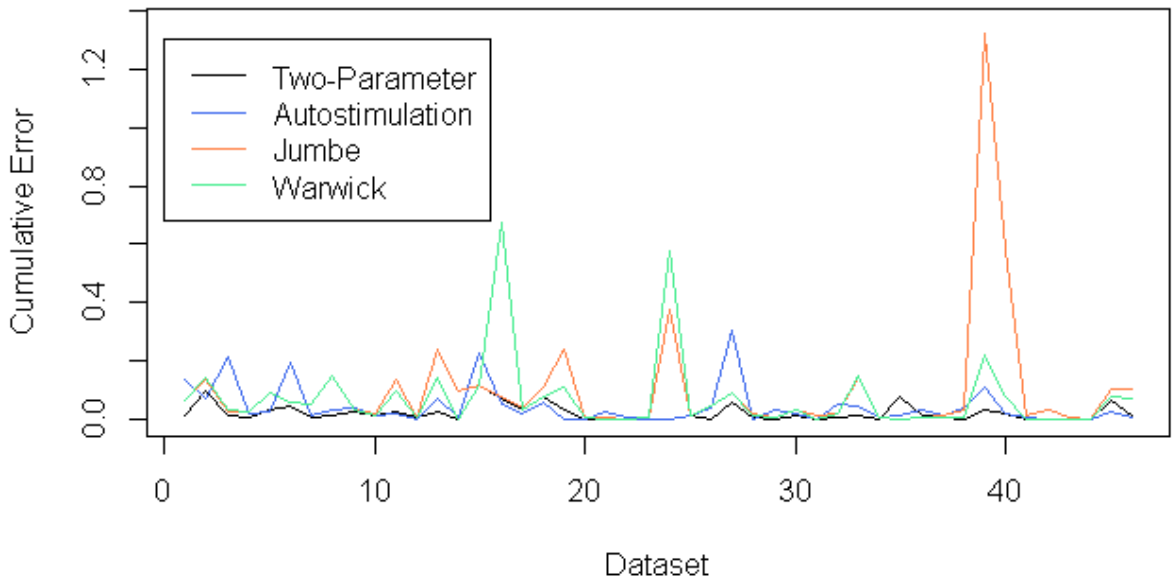


Figure 2-2: A comparison of the error of the Jumbe and Warwick models, along with the best and worst fitting of the previous models, across all the viable datasets.

In Figure 2-2, the datasets are put in an order of 1-48, being datasets from GCERK1420, GCGLS1404, GCERK1421, and GCGLS1433 in the order shown in Tables 2-9 to 2-13 inclusively.

Comparing the mean error of all the models, the Jumbe and Warwick model both performed worse than the previous models with the Jumbe model performing the worse overall. Figure 17 shows the variation in the error between the Jumbe model, the Warwick model, and the best and worse fitting models from the previous five (The Generalised Two-Parameter model and Autosimulation model respectively). As the graph shows, for the majority of data sets the Jumbe and Warwick model perform similarly to the Generalised Two-Parameter and Autosimulation models, however the exceptions stand out greatly.

These exceptions highlight the limitations of the models, as seen by the large spike in the Jumbe model found at dataset 3 from GCGLS1433. Figure 2-3 shows the fit of Jumbe model for this data set; the final two data points show a significant growth in comparison with the rest of the data and as the Jumbe model is only subject to one parameter it has little versatility and can't simulate this sudden growth.

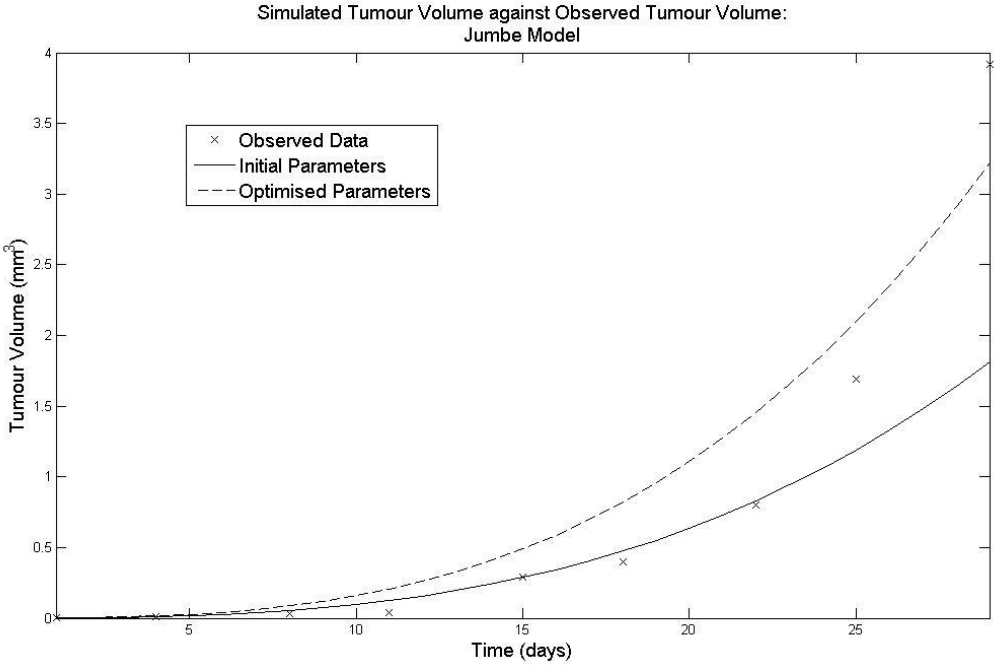


Figure 2-3: The simulation of the Jumbe model to fit to dataset 3 from the GCGLS1433 trials.

This also provides a suitable explanation as to why the Jumbe model performed worse than the other models, in that it can only accurately simulate data which follows certain growth patterns.

Performing the t-test for these new values gives the same result as before for the previous models. For the Warwick and Jumbe models, there's also non-significant evidence to suggest that the mean error of the two models are equal to each other, or the rest of the models.

Below is the table of values for the AIC comparison for the models. The weightings suggest that the Two-Parameter model has a 99.95% chance of fitting the data best, whilst the Jumbe model has the least chance of providing the best fit. The statistical ranking correlates with the ranking of the average mean of each model, further strengthening the evidence that to support the ordering of the models.

	Number of parameters	RSS	AIC_c	Δ_i	w_i	Rank
Gompertz	3	0.03895655	-142.7128	21.222	2.463e-05	3
Generic	4	0.04244814	-136.3601	27.5747	1.028e-06	4
Two-Parameter	4	0.02330888	-163.9348	0	0.9995	1
Autostimulation	5	0.04516311	-130.9838	32.9510	6.991e-08	5
Bertalanffy	3	0.03419178	-148.7140	15.2208	4.950e-04	2
Jumbe	1	0.1035465	-102.2249	61.7099	3.978e-14	7
Warwick	2	0.07413442	-115.4072	48.5276	2.898e-11	6

Table 2-14: The results of the AIC test for each of the original models, with the inclusion of the Jumbe and Warwick models.

2.10 CONCLUSION

Table 2-15 below compares the ranking of each the models with the number of parameters and the relative complexity of each of the models. The relative complexity is subjective, based on the mathematical operations used in the model and the number of parameters in the model, however it shows that there is a relation between the complexity and accuracy of a model; the least accurate models were the ones that were either most or least complex. This indicates that for a model to provide accurate fittings to data and also be a reliable predictor for simulating the effects of changes to a system it cannot be too simple as it won't have the versatility to deal with a range of different growth patterns, and conversely it can't be too complex, or the model would be too sensitive to changes in each of the parameters which makes finding an accurate fit for multiple parameters becomes difficult. Comparing the mean of the errors gave the ranking of each model with respect to its ability to fit the data. Whilst the t-test concluded that there was significant evidence to suggest the Gompertz and the Generic model perform as well as each other, the AIC comparison showed that the Generic model was ten times less likely than the Gompertz to provide the best fit for a model, justifying why they aren't listed in the same rank bracket.

Model	Equation	Number of Parameters	Data Fitting Rank	Complexity
Gompertz	$aV - b\,\mathrm{V}\log V$	2	3	2
Generic	$\frac{\beta V^{1-np}(k^n - V^n)^{1+p}}{k^n}$	4	4	6
Generalised Two-Parameter	$aV^{\alpha} - bV^{\beta}$	4	1	3
Autosimulation	$\alpha V \frac{(1+S)}{(1+\beta V)} - \omega V$ $\left(\frac{dS}{dt} = aV - bS^2, S(0) = 0\right)$	5	5	7
Generalised Bertalanffy	$\begin{cases} aV^{\alpha} - bV; & \alpha < 1 \\ aV - bV\log V\,; & \alpha = 1 \\ aV - bV^{\alpha}; & \alpha > 1 \end{cases}$	3	2	4
Jumbe	$KV^{\frac{2}{3}}$	1	7	1
Warwick	$\frac{dV}{dt} = kV\left(1 - \left(1 - \frac{r_d}{\left(\frac{3V}{4\pi}\right)^{\frac{1}{3}}}\right)^3\right)$	2	6	5

Table 2-15: An overview of the number of parameters in each model, their rank in the data fit, and subjective relative complexity.

The two best fitting models for the datasets, the Generalised Two-Parameter model and the Generalised Bertalanffy model, were derived from the principle that biological processes can be simulated using an allometric scaling process (West, Brown, & Enquist , 1997). Allometric scaling assumes that a process can be modelled as $M(t) = kW(t)^{\lambda}$, where $M(t)$ is the biological variable and $W(t)$ is the mass of the organism, and as $Mass = Density\ x\ Volume$ then $M(t) = k\rho^{\lambda}V(t)^{\lambda} = \tilde{k}V(t)^{\lambda}$ if the density, ρ, of the tumour is assumed constant. The Generalised Two-Parameter model assumes that the growth and decay of the tumour are both modelled using an allometric scaling process giving $\frac{dV}{dt} = aV^{\alpha} - bV^{\beta}$, where aV^{α} models the growth and bV^{β} models decay, and for the cases of the Bertalanffy model where $\alpha \neq 1$, the model is also represented by two separate allometric scaling processes.

The Jumbe model is another model represented by an allometric scaling process, however it only contains a growth term with a fixed power. As discussed in Section 15.2, this means the model is unbounded and the growth increases exponentially with the size of the tumour. This affects the overall versatility of the model and could explain why it was the least suitable of all the models.

The data used suggests that the best way to model the growth of tumours is with allometric scaling processes, with one to define tumour growth and the other for decay, and models based around these processes are best suited to model tumour growth.

3 Approximating an Analytical Solution to the Warwick model

3.1 Abstract

The Warwick model is derived from the principle that the growth of a tumour occurs strictly in an outer proliferating shell of the tumour. It has no solvable analytical solution, however it can be approximated to the Jumbe model by assuming that the proliferating shells depth is sufficiently small, and the value of the proliferating shell depth which minimizes the error between the two models can be found. Through this an approximate solution to the Warwick model can be gained, which can then be shown to hold for a generic proliferating shell size and extended to include the effects of a pharmacodynamic drug system.

3.2 Initial Approximation

3.2.1 Systems

The Warwick and Jumbe models are two separate models that can be used to model the growth of a tumor, both of which assume the tumor is spherical; the Warwick assumes there is a proliferating outer shell of constant depth which grows, while the Jumbe model assumes only the surface of the tumour grows. These systems are defined by the following equations;

Warwick:

$$\frac{dV}{dt} = k_g V\left(1 - \left(1 - \frac{r_d}{r}\right)^3\right) = k_g V_{pro} \tag{12}$$

Jumbe:

$$\frac{dV}{dt} = KV^{\frac{2}{3}} \tag{13}$$

Taking the volume to be $V = \frac{4\pi}{3}r^3$, and the surface area of the tumor to be $A = 4\pi r^2$, the Jumbe model becomes $\frac{dV}{dt} = K\left(\frac{4\pi}{3}r^3\right)^{\frac{2}{3}} = K\left(\frac{1}{36\pi}\right)^{\frac{1}{3}} 4\pi r^2 \propto A$, so (18) is proportional to the surface area of the tumor. Using that $\frac{dV}{dt} = 4\pi r^2 \frac{dr}{dt}$ gives $\frac{dr}{dt} = K\left(\frac{1}{36\pi}\right)^{\frac{1}{3}}$, which has a solution defined by $r = K\left(\frac{1}{36\pi}\right)^{\frac{1}{3}} t + r_0$, so the (18) has an analytical solution given by $V = \frac{4\pi}{3}\left(K\left(\frac{1}{36\pi}\right)^{\frac{1}{3}} t + r_0\right)^3$. As shown in Section (…), (12) can't be solved implicitly for $V(t)$, but if an approximation can be found from (12) to (13) then there might be an approximate solution for (12).

As the width of the proliferating shell becomes sufficiently small, to a point $\tilde{r}$, the volume of the proliferating shell will converge to the surface area of the tumor, hence as $r_d \to \tilde{r}$, $\frac{dV}{dt} \to k_g A \propto A$, so with the right parameters values the two models should be identical.

For sufficiently small values of r_d, terms of $O({r_d}^2)$ or higher are negligible, so;

$$k_g V\left(1-\left(1-\frac{r_d}{r}\right)^3\right) = k_g V\left(1-\left(1-3\frac{r_d}{r}+3\frac{{r_d}^2}{r^2}-\frac{{r_d}^3}{r^3}\right)\right)$$

$$\cong k_g V\left(3\frac{r_d}{r}\right)$$

$$= k_g V\left(3r_d\left(\frac{3V}{4\pi}\right)^{\frac{-1}{3}}\right)$$

$$= 36\pi r_d k_g V^{\frac{2}{3}} \propto A$$

Under this assumption, the Warwick model then has a solution given by $V = \frac{4\pi}{3}\left((36\pi)^{\frac{2}{3}} r_d k_g t + r_0\right)^3$.

As shown, for small enough r_d, the Warwick model should approximate the Jumbe model given the right parameters. As $\lim_{r_d=0} \frac{dV}{dt} = 0$, the value of r_{diff} which makes this approximation suitable must be non-zero, so there follows that there is a region $(0, \delta r)$ in which r_d is sufficiently small enough to make the approximation hold. It then follows to find the value of r_d which minimizes the error between the two models.

3.2.2 Fixing parameters

To find a value of r_d for which the assumption above holds, ${r_d}^*$, set the rate of change of each system equivalent for the initial condition to minimize the difference over a short time frame;

$$k_g V_0\left(1-\left(1-\frac{{r_d}^*}{r_0}\right)^3\right) = K{V_0}^{\frac{2}{3}} \qquad (14)$$

(14) can be solved to give ${r_d}^* = r_0\left(1-\left(1-\frac{K}{r_0 k_g}\left(\frac{3}{4\pi}\right)^{\frac{1}{3}}\right)^{\frac{1}{3}}\right)$[5], which is used as an initial estimate of r_d for the approximation to hold.

[5] Note: For the Warwick model to be valid, ${r_d}^*$ must have real values, so $1-\frac{K}{r_0 k_g}\left(\frac{3}{4\pi}\right)^{\frac{1}{3}} > 0$ must hold.

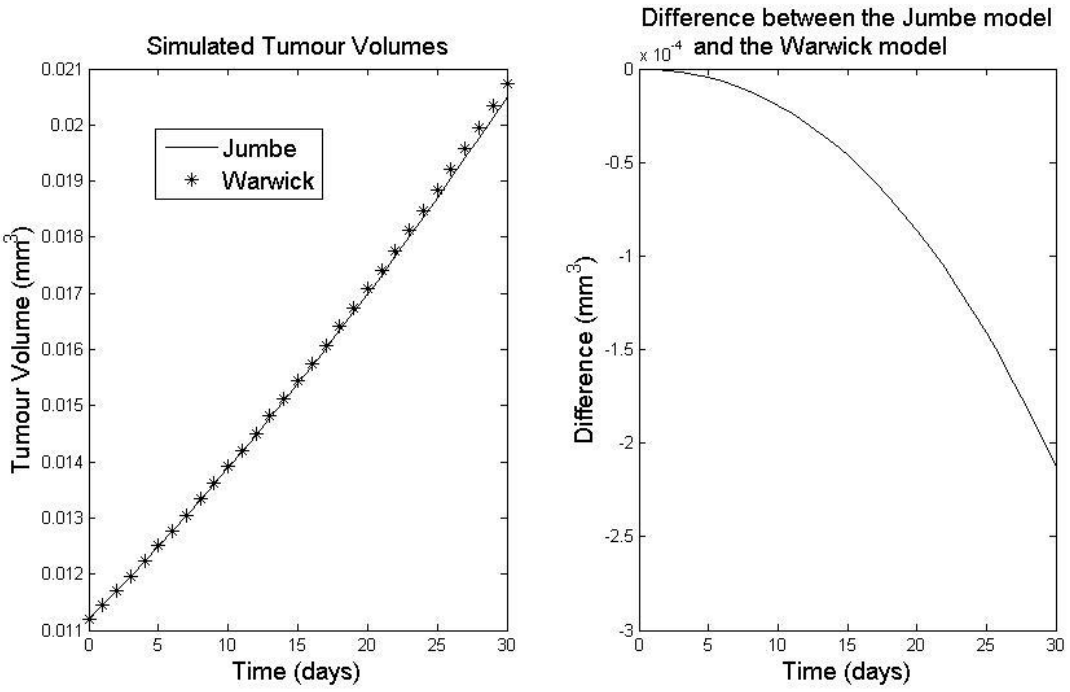


Figure 3-1: A plot comparing the simulation of the Jumbe model and the Warwick model for the initial estimate of the proliferating shell depth, and the difference between the two models for this estimate.

3.2.3 Simulation and Difference

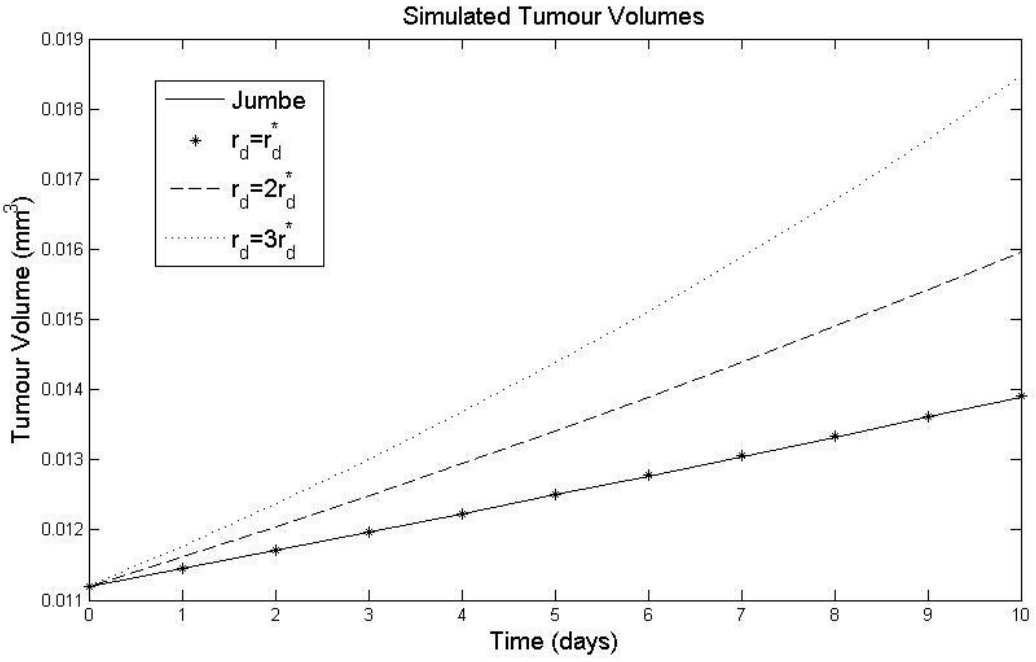


Figure 3-2: A plot comparing the Warwick model for values of the shell depth above the initial estimate.

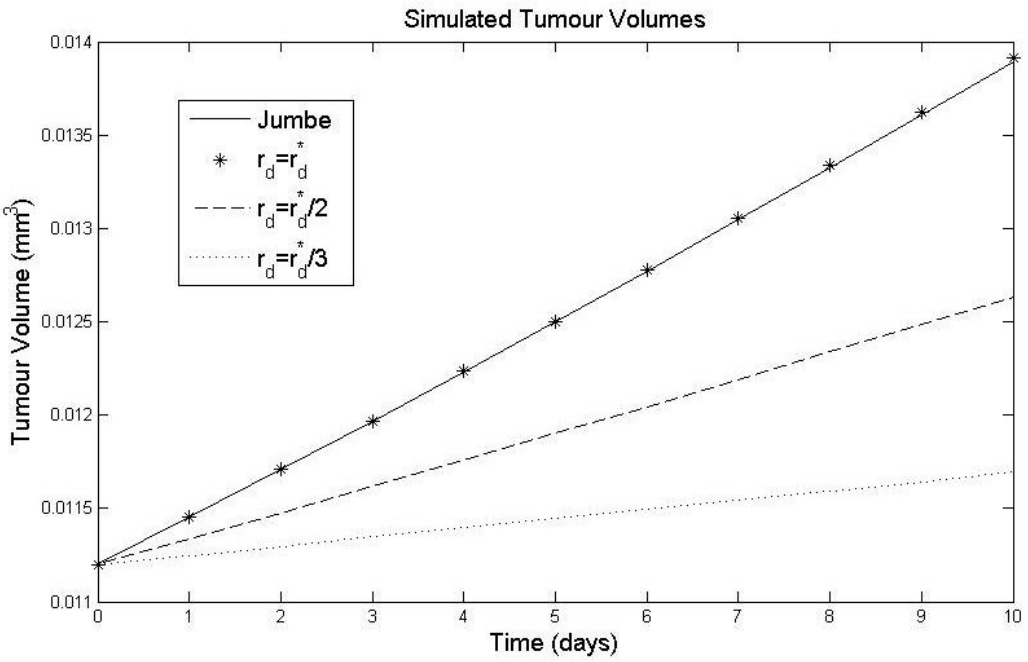


Figure 3-3: A plot comparing the Warwick model for values of the shell depth below the initial estimate.

Figures 3-2 and 3-3 show the simulation of the Jumbe model with K = 0.005 relative to 5 different cases of the Warwick model with k_{grow} = 0.05; these cases were $r_d = \{{r_d}^*, 2{r_d}^*, 3{r_d}^*, \frac{{r_d}^*}{2}, \frac{{r_d}^*}{3}\}$. As shown, large diversions away for the value of ${r_d}^*$ either above or below cause rapid changes to the strength of correlation between the two models.

The reason for this large variance in correlation in the case of ${r_d}^* + \delta r$ can be explained by the error as follows;

$$Error(E) = V_J - V_W$$

$$\Rightarrow \frac{dE}{dt} = \frac{dV_J}{dt} - \frac{dV_W}{dt}$$

$$\Rightarrow \frac{dE}{dt} = KV_J^{\frac{2}{3}} - k_g V_W \left(1 - \left(1 - \frac{r_d}{r}\right)^3\right) \qquad (15)$$

For $\frac{dE}{dt} = 0$ for all time, r_d would need to be a function of time, however as ${r_d}^*$ is constant and r is increasing, we get that ${r_d}^* \geq r\left(1 - \left(1 - \frac{K}{rk_g}\left(\frac{3}{4\pi}\right)^{\frac{1}{3}}\right)^{\frac{1}{3}}\right)$, which further gives $\frac{dD}{dt}\Big|_{r_d = {r_d}^*} < 0$ at time t >0 (proof in Appendix).

Considering equation (15) for $r_d = {r_d}^*$ and $r_d = {r_d}^* + \delta r$, where $r - {r_d}^* > \delta r > 0$, it can be shown that

$$\frac{dE}{dt}\Big|_{r_d = {r_d}^* + \delta r} < \frac{dE}{dt}\Big|_{r_d = {r_d}^*} < 0,$$

which holds for all t, showing that for small increments above $r_d{}^*$ the approximation is less accurate. Considering $\frac{d^2E}{dt^2}$;

$$\frac{d^2E}{dt^2} = \frac{d^2V_J}{dt^2} - \frac{d^2V_W}{dt^2}$$

$$\Rightarrow \frac{d^2E}{dt^2} = \frac{2}{3}KV_J{}^{\frac{-1}{3}}\frac{dV_J}{dt} - k_g\frac{dV_W}{dt}\left(1 - \left(1 - \frac{r_d}{r}\right)^2\left(1 - 2\frac{r_d}{r}\right)\right)$$

Using that $\left(1 - \frac{r_d{}^* + \delta r}{r}\right)^2\left(1 - 2\frac{r_d{}^* + \delta r}{r}\right) < \left(1 - \frac{r_d{}^*}{r}\right)^2\left(1 - 2\frac{r_d{}^*}{r}\right)$, it can be shown that

$$\left.\frac{d^2E}{dt^2}\right|_{r_d = r_d{}^* + \delta r} < \left.\frac{d^2E}{dt^2}\right|_{r_d = r_d{}^*}$$

As the error produced from $r_d = r_d{}^* + \delta r$ decreasing quicker than $r_d = r_d{}^*$, and the error for $r_d{}^*$ is always non-positive in the absence of drug effects, the difference between the Jumbe model and the Warwick model for $r_d = r_d{}^* + \delta r$ will always less than 0, and of greater magnitude than the difference from simulating the Warwick model for $r_d{}^*$.

3.2.4 Optimising $r_d{}^*$

Conversely to the case of $r_d{}^* + \delta r$, $\left.\frac{dD}{dt}\right|_{r_d = r_d{}^* - \delta r}$ can't be bound . For $\delta r > 0$, $r_d{}^* - \delta r$ is a solution for $\frac{d^2E}{dt^2} = 0$ at point in time t>0, leading to the graphs shown in Figure 3-4;

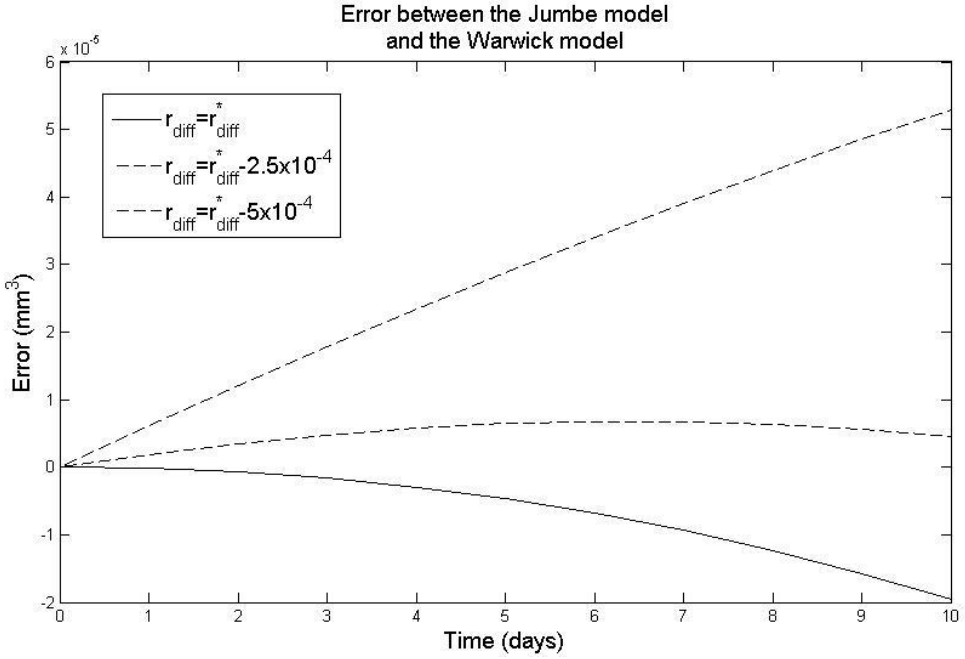


Figure 3-4: A plot illustrating the effects of taking small increments from the initial estimate for the shell depth on the error between the Jumbe model and Warwick model.

For $\delta r = 2.5 \times 10^{-4}$, the difference travels up to a peak before tending downwards again, and as seen provides a better fit in terms of total difference than $r_d = {r_d}^*$. This improvement has a limit shown by $\delta r = 5 \times 10^{-4}$ where the difference is far greater across all the time points.

Figure 3-5 shows a similar comparison for the same growth parameters over a longer time period, which illustrates how the variations progress over time; comparing the difference curves for $\delta r = 5 \times 10^{-4}$ shows that for some $t > 10$, $\left.\frac{dD}{dt}\right|_{r_d = {r_d}^* - \delta r} = 0$, and the difference between the two models becomes exponentially more negative after this point.

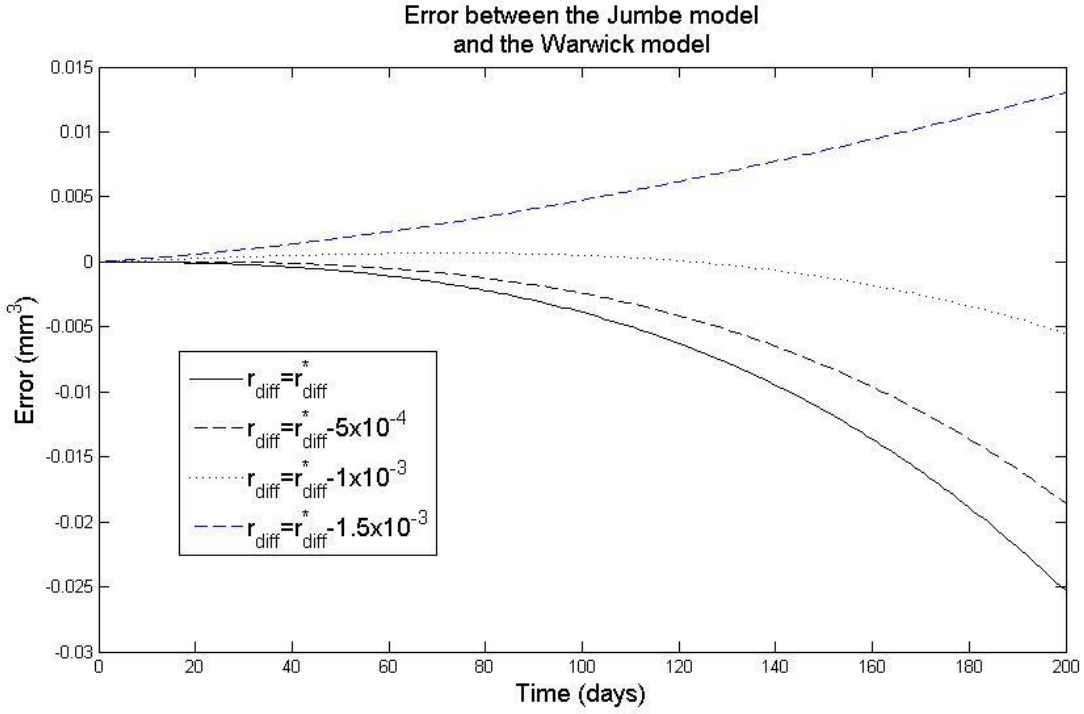


Figure 3-5: A plot illustrating the effects of taking small increments from the initial estimate for the shell depth on the error between the Jumbe model and Warwick model.

For a given number of time points, to find the optimal value of r_d it's sufficient to consider the absolute value of the sum of errors across the time points for a given r_d; if the absolute value of the sum of errors for $r_d - \delta r$ is greater than it was for r_d, then r_d was the optimal value to within δr degree of accuracy. As seen in Figure 3-6, the optimal value for the data used above with 200 time points is ${r_d}^* - 1.75 \times 10^{-3}$, however Figure 3-7 shows that for 500 time points, the optimal value becomes ${r_d}^* - 2.5 \times 10^{-3}$.

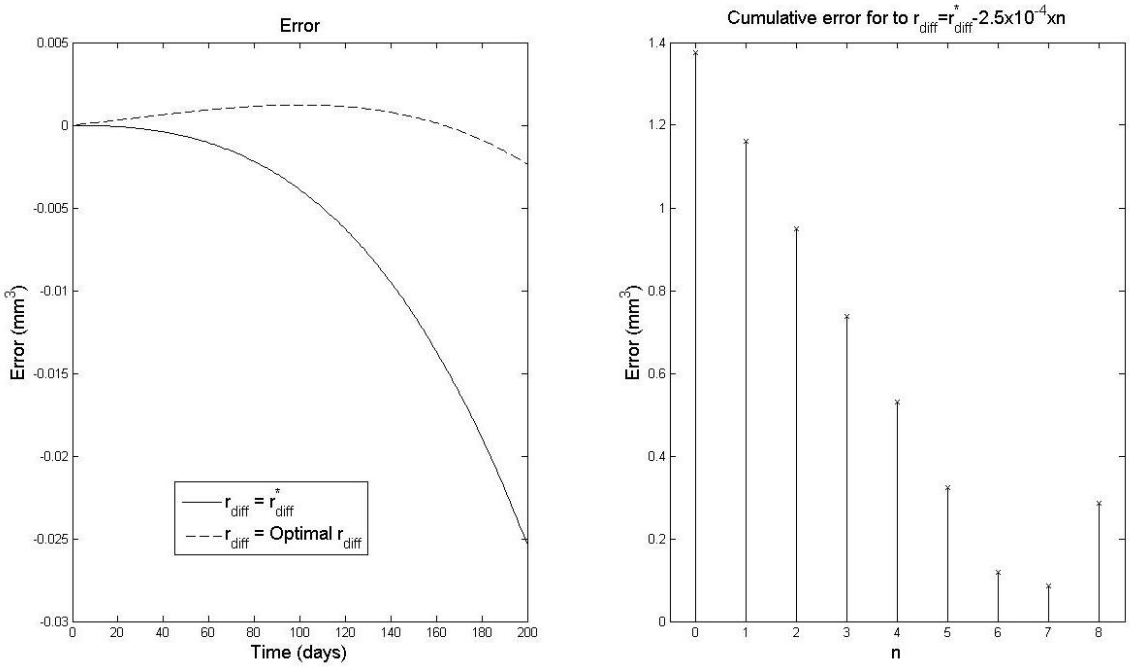


Figure 3-6: A comparison between the initial estimate for the proliferating shell depth and the optimized value, along with the modulus of the cumulative error for each iteration to show how the error changes with the changes in the pro

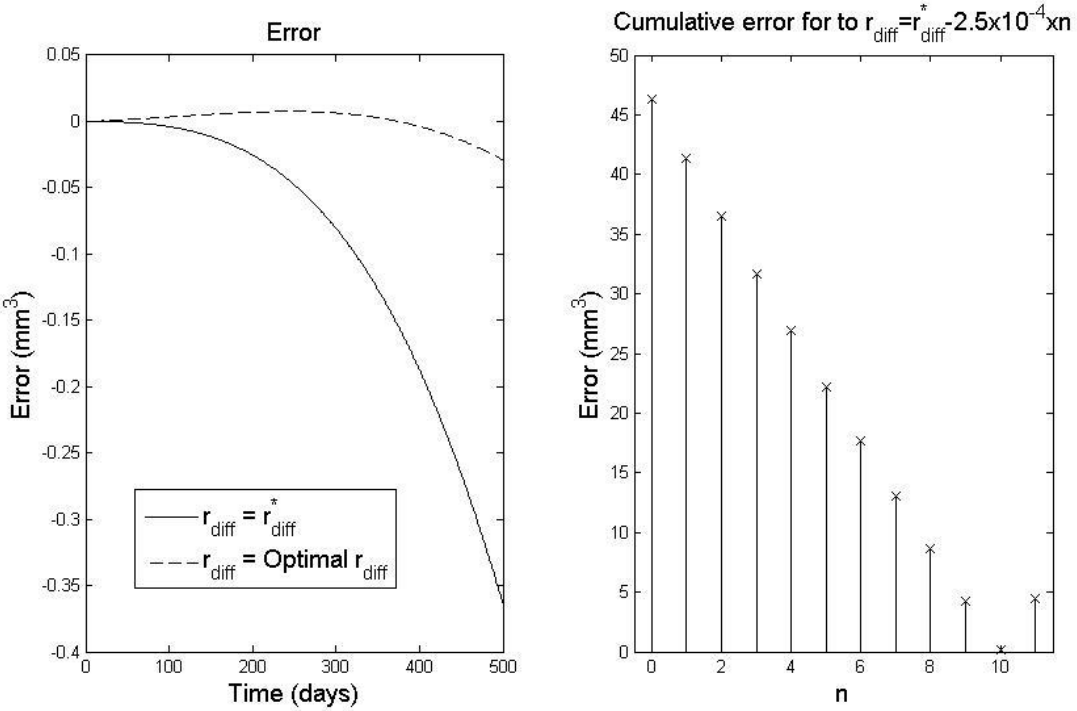


Figure 3-7: A comparison between the initial estimate for the proliferating shell depth and the optimized value over 500 time steps, along with the modulus of the cumulative error for each iteration.

3.3 INTRODUCING DRUG EFFECTS

3.3.1 Reconfiguring the system

When the presence of a drug is added to each model, the governing equations change to take into account this change to growth. They then become;

Warwick:

$$\frac{dV_W}{dt} = (k_g - d_{eW})V_W\left(1 - \left(1 - \frac{r_d}{r}\right)^3\right) = (k_g - d_{eW})V_{pro} \qquad (16)$$

Jumbe:

$$\frac{dV_J}{dt} = (K - d_{eJ})V_J^{\frac{2}{3}} \qquad (17)$$

In this new system, the previous initial value of $r_d{}^*$ no longer provides a good approximation from (16) to (17) and leads to divergence between the two, as shown below in Figure 3-8;

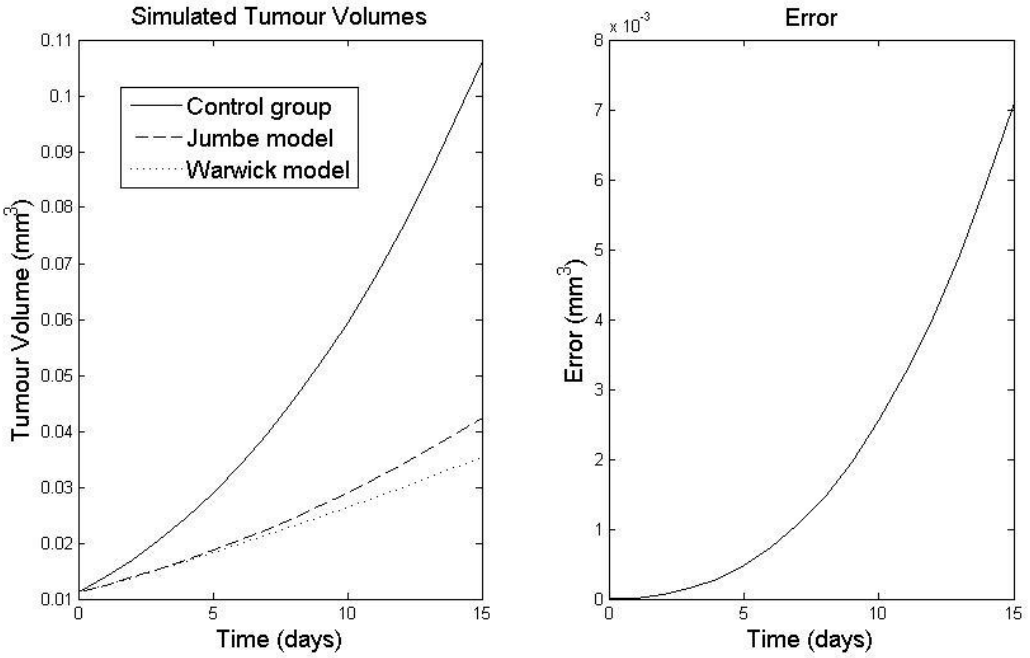


Figure 3-8: The effect of adding a drug effect to the system on the accuracy of the initial estimate.

3.3.2 Re-estimating $r_d{}^*$

To make the systems equivalent a new optimal value of r_d is required. For this new system, the initial value of $r_d{}^*$ to make the two systems equivalent is found in the same manner by solving the system $\left.\frac{dV_J}{dt}\right|_{t=0} = \left.\frac{dV_W}{dt}\right|_{t=0}$ for r_d;

$$(k_g - d_{eW})V_0\left(1 - \left(1 - \frac{r_d{}^*}{r_0}\right)^3\right) = (K - d_{eJ})V_0{}^{\frac{2}{3}}$$

$$\Rightarrow r_d{}^* = r_0\left(1-\left(1-\frac{(K-d_{eJ})}{r_0(k_g-d_{eW})}\left(\frac{3}{4\pi}\right)^{\frac{1}{3}}\right)^{\frac{1}{3}}\right)$$

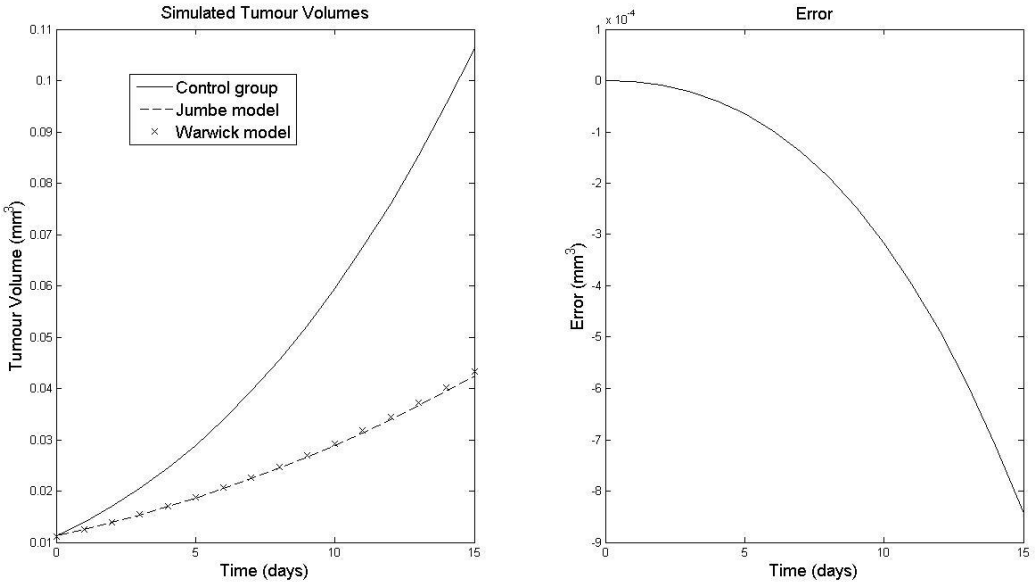


Figure 3-9: The comparison between the Jumbe model and Warwick model with the revised initial estimation of the proliferating shell depth which approximates the Warwick model as the Jumbe model in the presence of drug effects.

Figure 3-9 shows how (16) acts under this new value of $r_d{}^*$ in relation to (17), compared to (17) for $d_{eJ} = 0$ (Control group). This re-estimated value of $r_d{}^*$ converges the Warwick models to the Jumbe model and the approximation becomes less accurate over time, as before. As $d_{eJ} \longrightarrow 0$ and $d_{eW} \longrightarrow 0$, this new value collapses into the same one as the value from Section 3.2.2.

Note: For the revised formula for $r_d{}^*$ to be valid (real and non-negative), the following condition must now hold, $0 < 1-\frac{K-d_{eJ}}{r_0(k_g-d_{eW})}\left(\frac{3}{4\pi}\right)^{\frac{1}{3}} < 1$. This is similar to the condition from Section 3.2, but with an additional non-negativity constraint on $\frac{K-d_{eJ}}{r_0(k_g-d_{eW})}\left(\frac{3}{4\pi}\right)^{\frac{1}{3}}$ to ensure that $r_d{}^*$ is non-negative.

3.3.3 Optimizing $r_d{}^*$

To optimize r_d, once again look to minimize the sum of errors over the time points. As the drug effect has altered the constant terms in the governing equations of the system, the results are expected to be similar to Section 3.2.3 when considering $r_d{}^* - \delta r$. Figures 3-10 and 3-11 show the results of this for 100 and 500 time points respectively, showing again that the optimal value of r_d is dependent on the number of time points available.

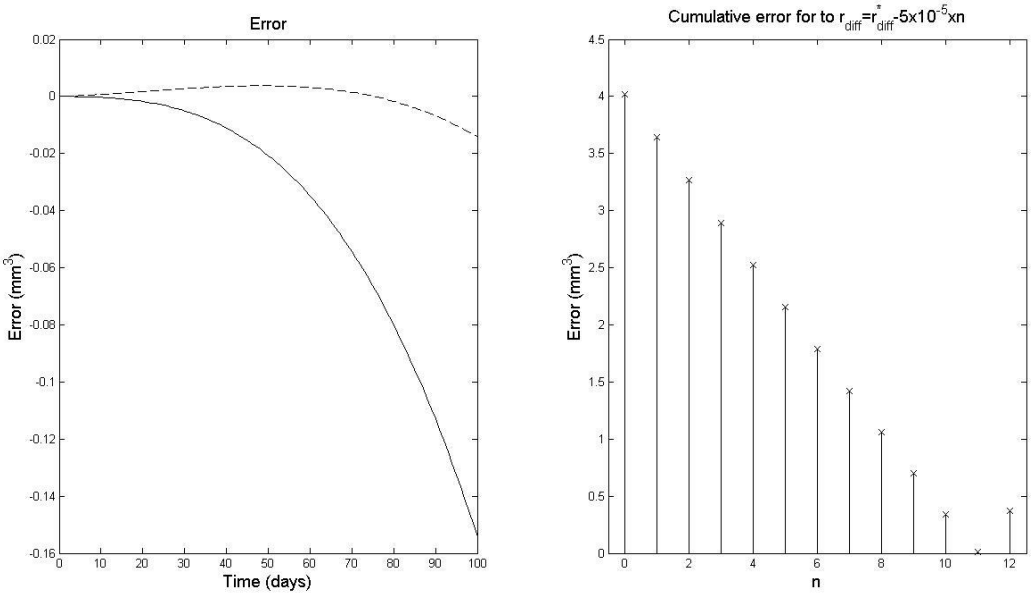


Figure 3-10: Optimising the initial estimate for the shell depth over 100 time points, along with the modulus of the cumulative error for each iteration.

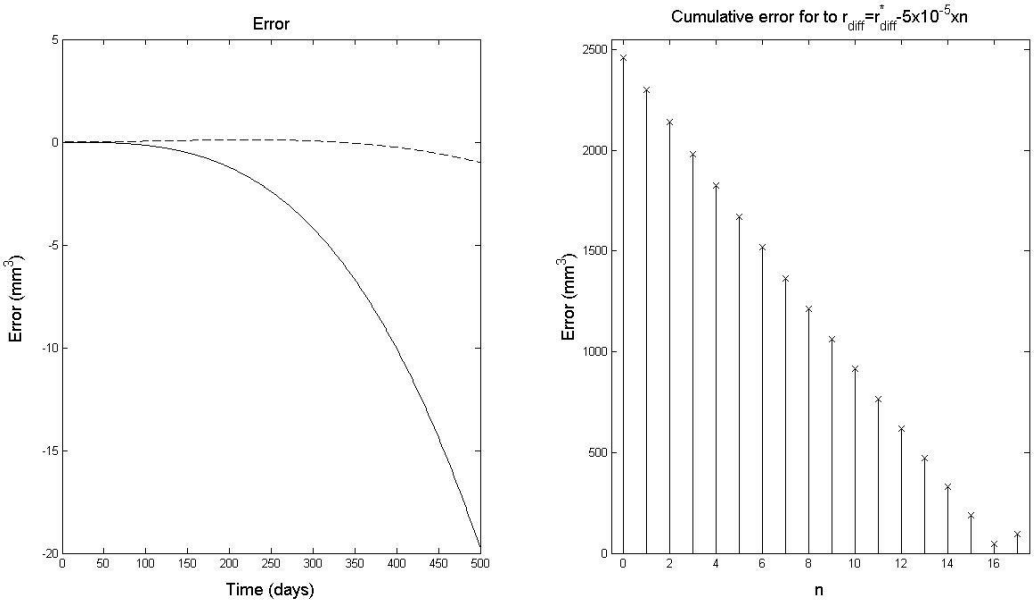


Figure 3-11: Optimising the initial estimate for the shell depth over 500 time points, along with the modulus of the cumulative error for each iteration.

3.4 PK EFFECTS

3.4.1 Reconfiguring the system

At the moment, the models both assume that the quantities of drug present is constant, so the drug effect is constant. This means that both models neglect the transfer of drug from body tissue to plasma, or the removal of drug from the body, so the next step is to introduce a varying drug quantity to the

system and see if that further effects the results. The system is now adapted to take into account two new quantities, the volume of drug in the body tissue, q_t, and the volume of drug in the body plasma, q_p. This new system can be defined by the following;

$$\frac{dq_p}{dt} = -(k_{pe} + k_{pt})q_p + k_{tp}q_{t,}$$

$$\frac{dq_t}{dt} = k_{pt}q_p - k_{tp}q_t,$$

$$\frac{dV_W}{dt} = \left(k_g - d_{eW}\left(\frac{q_p}{v_p}\right)\right)V_W\left(1 - \left(1 - \frac{r_d}{r}\right)^3\right),$$

$$\frac{dV_J}{dt} = \left(K - d_{eJ}\left(\frac{q_p}{v_p}\right)\right)V_J^{\frac{2}{3}}.$$

This gives a new initial estimate for the shell's width;

$$r_d{}^* = r_0\left(1 - \left(1 - \frac{\left(K - d_{eJ}\left(\frac{q_{p0}}{v_p}\right)\right)}{r_0\left(k_g - d_{eW}\left(\frac{q_{p0}}{v_p}\right)\right)}\left(\frac{3}{4\pi}\right)^{\frac{1}{3}}\right)^{\frac{1}{3}}\right).$$

However, the presence of the varying q_p changes the outcome of previous simulations drastically. Taking the same parameters as before ($\{K, k_g, d_e, V_0\}$ = {0.05,0.5,0.0025,0.0112}), along with $\{k_{pe}, k_{pt}, k_{tp}, v_p\}$ = {0.382,0.523,0.196,1.3}, the simulation now produces Figure 3-12.

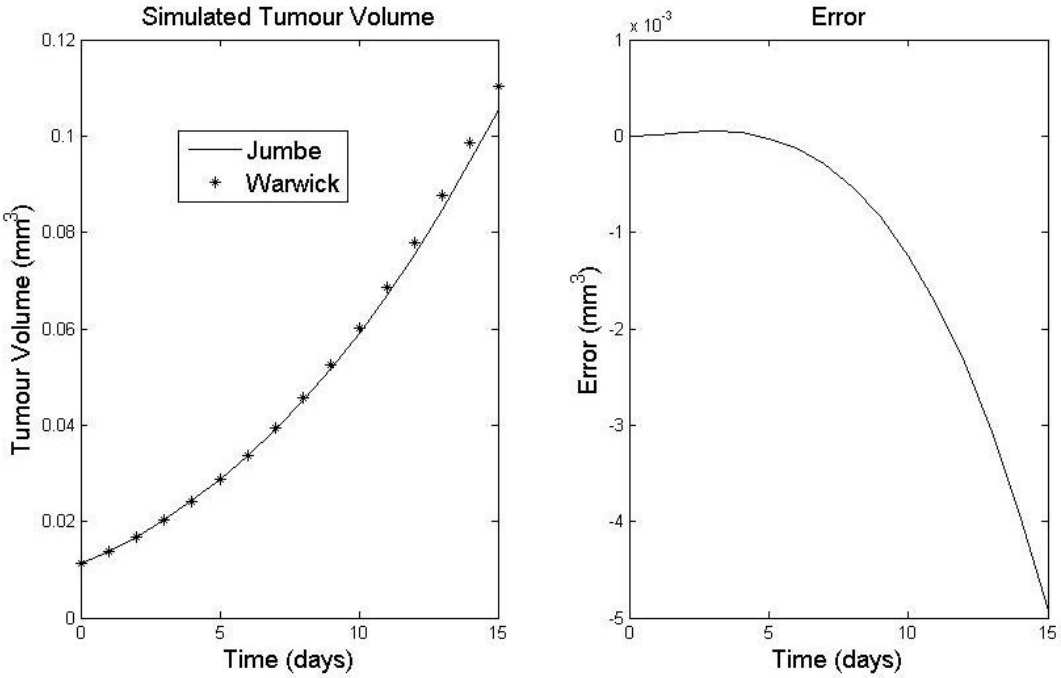


Figure 3-12:A plot comparing the simulation of the Jumbe model and the Warwick model under the effect of a pharmacokinetic drug system for the redefined value of the initial shell depth estimate.

Considering $r_d{}^* \pm \delta r$, we get the following results;

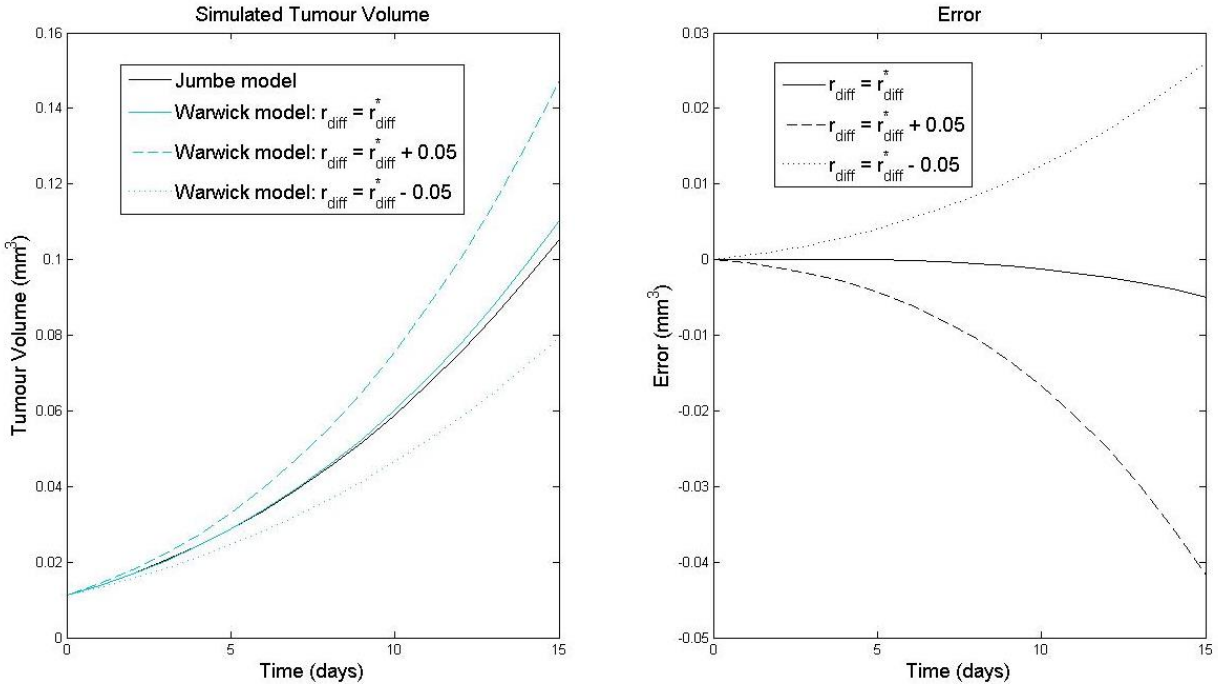


Figure 3-13: Varying the initial estimate for the shell depth by small increments above and below.

3.4.2 Optimising r_d^*

As with the constant drug effect, an optimized value of r_d is dependent on the number of time points the model is simulated over. Figure 3-14 and Figure 3-15 below show the optimization for the parameters above for 100 and 500 time points, and as before the optimal value differs greatly.

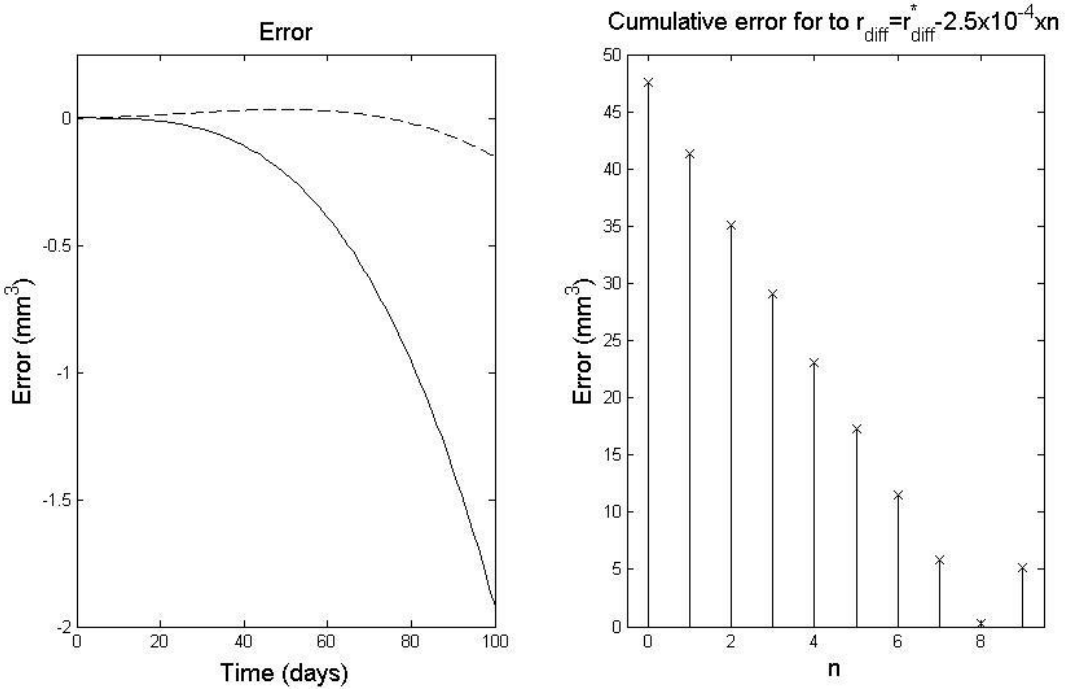


Figure 3-14: Optimising the initial estimate for the proliferating shell depth over 100 time points, along with the modulus of the cumulative error for each iteration.

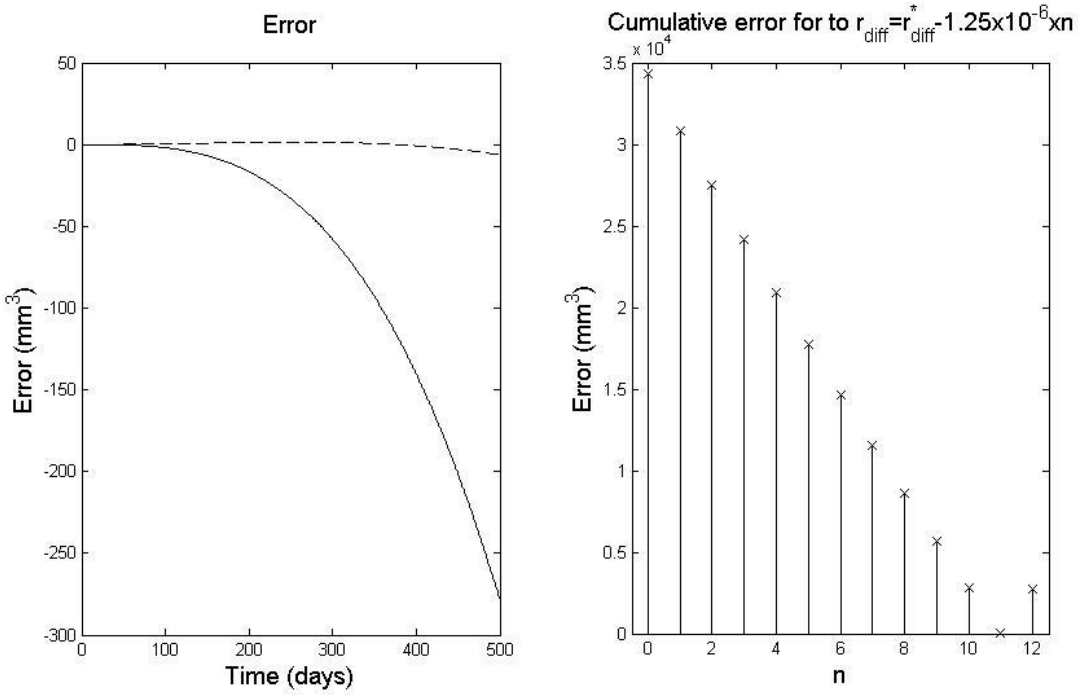


Figure 3-15:Optimising the initial estimate for the proliferating shell depth over 500 time points, along with the modulus of the cumulative error for each iteration.

As shown above, for a small enough value of r_d the Warwick model can provide an accurate approximation for the Jumbe model for any growth parameters of both systems. There isn't a set formula for the optimal value as it varies subject to the number of time points the simulations are generated over, however this strengthens the idea that there may be an approximate analytical for the Warwick model attainable through the approximation to the Jumbe model.

3.5 CONCENTRIC SPHERE

3.5.1 Premise and systems

Consider splitting the proliferating shell up into N separate shells, each of depth r_d which is sufficiently small enough to make a realistic approximation from the Warwick model to the Jumbe model for each shell. In this case, the goal is to see if each of these shells can be approximated at the same time by a Jumbe model with only one coefficient, $\widetilde{K}$, however first it will be considered as the composition of multiple Jumbe models. A visual representation of this idea is shown in Figure 3-16.

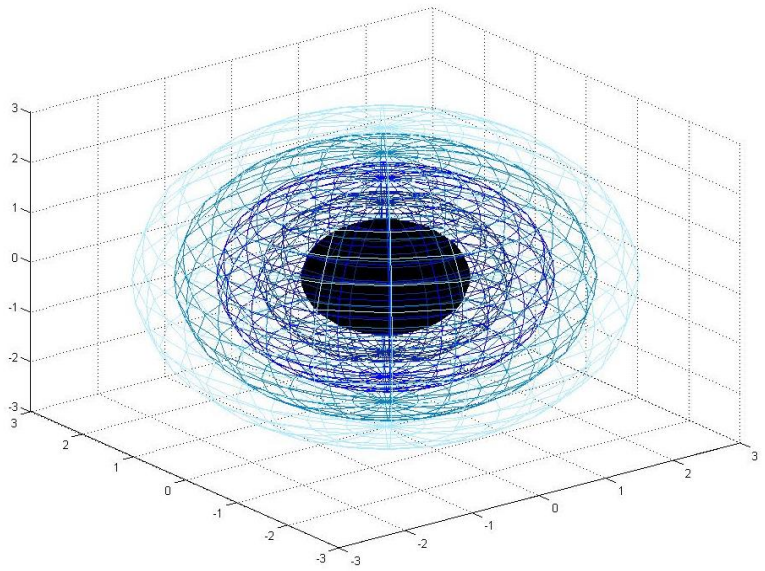


Figure 3-16: Simple diagram to represent a tumour with 4 proliferating shells, each of equal depth.

As seen before, the growth given by the Warwick model for a single proliferating shell is modelled as $\frac{dV}{dt} = k_g V_{pro}$. Assuming the growth constant to be the same in each proliferating shell[6], the combined rate of change in volume is then gained by the superposition principle;

$$\frac{dV_W}{dt} = k_g \sum_{i=1}^{N} {V_{pro}}^i$$

$$\Rightarrow \frac{dV_W}{dt} = k_g \sum_{i=1}^{N} V^i \left(1 - \left(1 - \frac{r_d}{r^i}\right)^3\right),$$

[6] It may however be more realistic to make the growth constant in the inner shells smaller, to account for the decrease in resources needed for growth.

where ${V_{pro}}^i$ is the volume of the i^{th} proliferating shell, V^i is the volume of the tumour up to the i^{th} shell, and r^i is the corresponding radius.

3.5.2 Splitting the Warwick model into two shells

Consider two tumours, one with volume V_1 made up of two proliferating shells, each of depth r_d, and a core, and one with volume V_2 made up of a single proliferating shell, of depth $2r_d$ and a core.

For the tumour with two shells, $V_{1;}$ is the volume of the core, inner, and outer shell together, V_{11} is the volume of the core and the inner shell.

The growth of each of the proliferating shells can be described by $k_g V\left(1-\left(1-\frac{r_d}{r}\right)^3\right)$, where V is the volume of the tumour up to the outside of the proliferating shell, and r is the corresponding radius.

The total change in V_1 can then be given by;

$$\frac{dV_1}{dt} = k_g V_1\left(1-\left(1-\frac{r_d}{r_1}\right)^3\right) + k_g V_{11}\left(1-\left(1-\frac{r_d}{r_{11}}\right)^3\right)$$

V_{11} can be written as the total volume minus the volume of the outer shell, so $V_{11} = \left(V_1 - V_1\left(1-\left(1-\frac{r_d}{r_1}\right)^3\right)\right)$, substituting this gives;

$$\frac{dV_1}{dt} = k_g V_1\left(1-\left(1-\frac{r_d}{r_1}\right)^3\right) + k_g\left(V_1 - V_1\left(1-\left(1-\frac{r_d}{r_1}\right)^3\right)\right)\left(1-\left(1-\frac{r_d}{r_1-r_d}\right)^3\right)$$

This can then be rearraged to give a simplified equation for $\frac{dV_1}{dt}$;

$$\frac{dV_1}{dt} = k_g V_1\left(1-\left(1-\frac{r_d}{r_1}\right)^3\right) + k_g V_1\left(1-\frac{r_d}{r_1}\right)^3\left(1-\left(1-\frac{r_d}{r_1-r_d}\right)^3\right)$$

$$\Rightarrow \frac{dV_1}{dt} = k_g V_1\left(1-\left(1-\frac{r_d}{r_1}\right)^3\left(1-\frac{r_d}{r_1-r_d}\right)^3\right) \quad (18)$$

The total change in V_2 is given by;

$$\frac{dV_2}{dt} = k_g V_2\left(1-\left(1-\frac{2r_d}{r_2}\right)^3\right) \quad (19)$$

Given the same starting values, for these two models to grow equivalently to each other, (18) and (19) would need to be equal for all t, so take t = 0;

$$k_g V_0\left(1-\left(1-\frac{r_d}{r_0}\right)^3\left(1-\frac{r_d}{r_0-r_d}\right)^3\right) = k_g V_0\left(1-\left(1-\frac{2r_d}{r_0}\right)^3\right)$$

$$\Rightarrow \left(1-\frac{r_d}{r_0}\right)^3\left(1-\frac{r_d}{r_0-r_d}\right)^3 = \left(1-\frac{2r_d}{r_0}\right)^3$$

$$\Rightarrow \frac{1}{{r_0}^3}(r_0 - r_d)^3 \left(\frac{r_0 - 2r_d}{r_0 - r_d}\right)^3 = \left(1 - \frac{2r_d}{r_0}\right)^3$$

$$\Rightarrow \left(1 - \frac{2r_d}{r_0}\right) = \left(1 - \frac{2r_d}{r_0}\right)^3 \blacksquare$$

This means that the single shell model can be accurately approximated by a model with two shells, and by repeating the splitting process the single shell model can be approximated by 2^n shells, of depth ${r^*}_d = \frac{r_d}{2^n}$.

3.5.3 Splitting the Warwick model into N shells

The next step is to prove the Warwick model can be split into a generic number of N shells, which can be done via induction;

Let the inductive hypothesis be that $\sum_{i=1}^{n} k_g V_i \left(1 - \left(1 - \frac{r_d}{r_i}\right)^3\right) = k_g V_1 \left(1 - \left(1 - \frac{nr_d}{r_1}\right)^3\right)$. This is clearly true when $n = 1$, and as proven above true when $n = 2$.

Assume it holds for $n = k$, so

$$\sum_{i=1}^{k} k_g V_i \left(1 - \left(1 - \frac{r_d}{r_i}\right)^3\right) = k_g V_1 \left(1 - \left(1 - \frac{kr_d}{r_1}\right)^3\right) \quad (20)$$

Considering the case when $n = k + 1$;

Let $\sum_{i=1}^{k+1} k_g V_i \left(1 - \left(1 - \frac{r_d}{r_i}\right)^3\right) = I(k+1)$;

$$I(k+1) = k_g V_1 \left(1 - \left(1 - \frac{r_d}{r_1}\right)^3\right) + \sum_{i=2}^{k+1} k_g V_i \left(1 - \left(1 - \frac{r_d}{r_i}\right)^3\right)$$

From (20) we have $\sum_{i=2}^{k+1} k_g V_i \left(1 - \left(1 - \frac{r_d}{r_i}\right)^3\right) = k_g V_2 \left(1 - \left(1 - \frac{kr_d}{r_2}\right)^3\right)$, substituting this gives;

$$I(k+1) = k_g V_1 \left(1 - \left(1 - \frac{r_d}{r_1}\right)^3\right) + k_g V_2 \left(1 - \left(1 - \frac{kr_d}{r_2}\right)^3\right)$$

Using that $V_2 = \left(V_1 - V_1 \left(1 - \left(1 - \frac{r_d}{r_1}\right)^3\right)\right)$, this can be written as;

$$I(k+1) = k_g V_1 \left(1 - \left(1 - \frac{r_d}{r_1}\right)^3\right) + k_g \left(V_1 - V_1 \left(1 - \left(1 - \frac{r_d}{r_1}\right)^3\right)\right)\left(1 - \left(1 - \frac{kr_d}{r_1 - r_d}\right)^3\right)$$

$$\Rightarrow I(k+1) = k_g V_1 \left(1 - \left(1 - \frac{r_d}{r_1}\right)^3\right) + k_g V_1 \left(1 - \frac{r_d}{r_1}\right)^3 \left(1 - \left(1 - \frac{kr_d}{r_1 - r_d}\right)^3\right)$$

$$\Rightarrow I(k+1) = k_g V_1 \left(1 - \left(1 - \frac{r_d}{r_1}\right)^3 \left(1 - \frac{kr_d}{r_1 - r_d}\right)^3\right)$$

$$\Rightarrow I(k+1) = k_g V_1 \left(1 - \left(\frac{r_1 - r_d}{r_1}\right)^3 \left(\frac{r_1 - (k+1)r_d}{r_1 - r_d}\right)^3\right)$$

$$\Rightarrow I(k+1) = k_g V_1 \left(1 - \left(1 - \frac{(k+1)r_d}{r_1}\right)^3\right) \blacksquare$$

As it holds for $n = k + 1$, it holds for all n, so the Warwick model can be estimated by any number of concentric shells. So considering the single shell model;

$$\frac{dr}{dt} = \frac{k_g r}{3}\left(1 - \left(1 - \frac{r_d}{r}\right)^3\right) = \sum_{i=1}^{n} \frac{k_g r_i}{3}\left(1 - \left(1 - \frac{\widetilde{r_d}}{r_i}\right)^3\right)$$

Integrating both sides with respect to time, and using the approximation that for small enough $\widetilde{r_d}$, $\frac{k_g r_i}{3}\left(1 - \left(1 - \frac{\widetilde{r_d}}{r_i}\right)^3\right) = k_g \widetilde{r_d}$, this gives;

$$\int \frac{dr}{dt} dt = \sum_{i=1}^{n} \int \frac{k_g r_i}{3}\left(1 - \left(1 - \frac{\widetilde{r_d}}{r_i}\right)^3\right) dt$$

$$\Rightarrow \int \frac{dr}{dt} dt = \sum_{i=1}^{n} \int k_g \widetilde{r_d} dt$$

$$\Rightarrow r = \textstyle\sum_{i=1}^{n} \left(k_g \widetilde{r_d} t + c_i\right)$$

$$\Rightarrow r = n k_g \widetilde{r_d} t + C$$

$$\Rightarrow r = k_g r_d t + r_0 \qquad (21)$$

From this, the Warwick model should have a linear solution over time, which is dependent on k_g and r_d.

Figure 3-17 below shows the simulation of this model for 4 different sets of parameters (from left to right, top to bottom);

a) $\{k_g, r_d, r_0\} = \{0.5, 0.0005, 0.05\}$, b) $\{k_g, r_d, r_0\} = \{0.5, 0.0025, 0.05\}$,
c) $\{k_g, r_d, r_0\} = \{0.1, 0.005, 0.05\}$, d) $\{k_g, r_d, r_0\} = \{1, 0.005, 0.05\}$.

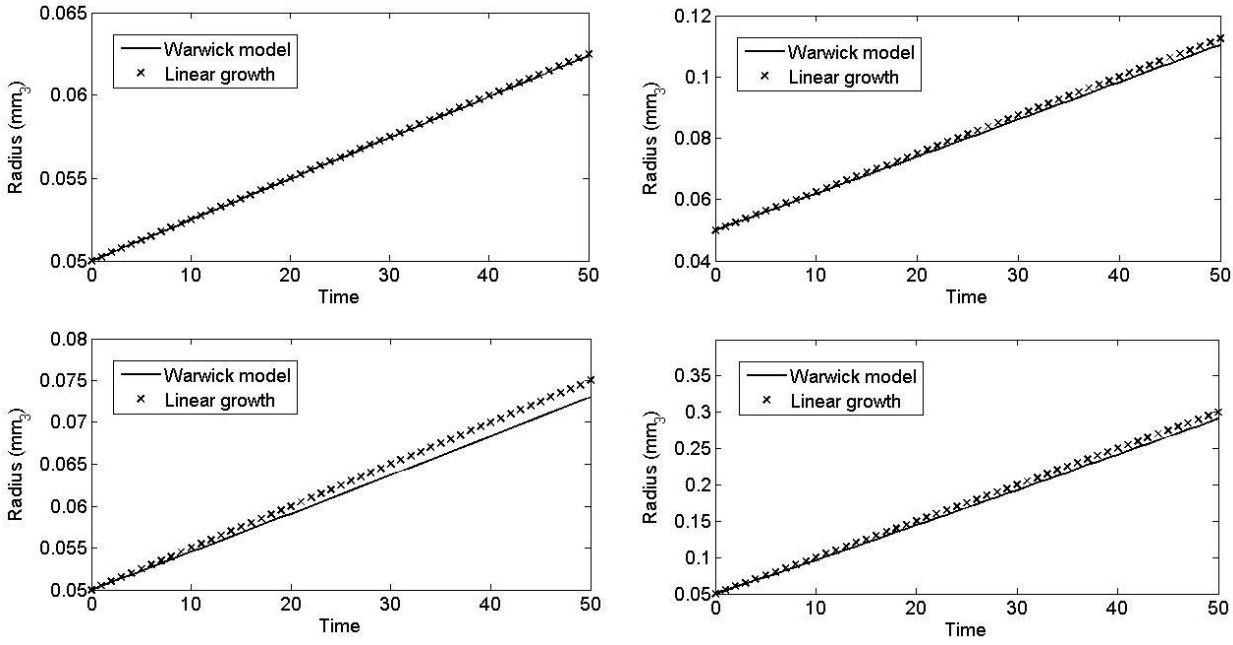


Figure 3-17: A comparison between the Warwick model and the Linear approximation (26).

As shown across the examples in Figure 3-17, the linear model provides a good approximation for the Warwick model for non-extreme parameter values. It becomes less accurate for large values of r_d or small values of k_g.

3.5.4 Concentric Spheres with Constant Drug Effect

Under constant drug effects, the Warwick model is $\frac{dV}{dt} = (k_g - d_e)V\left(1 - \left(1 - \frac{r_d}{r}\right)^3\right)$. As this is the same as the original model down to the constant term, so the solution is the same as before with a revised linear coefficient, giving $r = (k_g - d_e)r_d t + r_0$.

3.5.5 Concentric Sphere with Pharmacokinetic Effects

When the system is under the effect of a drug and pharmacokinetics, the solution changes due to the fact that the concentration of the drug within the body plasma, q_p, is a function of time. Under these new settings, the model is $\frac{dr}{dt} = \frac{1}{3}\left(k_g - d_e\left(\frac{q_p(t)}{v_p}\right)\right)r\left(1 - \left(1 - \frac{r_d}{r}\right)^3\right)$. This can be solved as follows;

$$\frac{dr}{dt} = \frac{k_g}{3}r\left(1 - \left(1 - \frac{r_d}{r}\right)^3\right) - \frac{d_e}{3}\left(\frac{q_p(t)}{v_p}\right)r\left(1 - \left(1 - \frac{r_d}{r}\right)^3\right)$$

$$\int_0^t \frac{dr}{dt}dt = \int_0^t \frac{k_g}{3}r\left(1 - \left(1 - \frac{r_d}{r}\right)^3\right)dt - \int_0^t \left(\frac{d_e q_p(t)}{3v_p}\right)r\left(1 - \left(1 - \frac{r_d}{r}\right)^3\right)dt$$

$$r(t) - r_0 = k_g r_d t - \int_0^t \left(\frac{d_e q_p(t)}{3v_p}\right)r\left(1 - \left(1 - \frac{r_d}{r}\right)^3\right)dt$$

Let $I = \int_0^t \left(\frac{d_e q_p(t)}{3v_p}\right) r\left(1 - \left(1 - \frac{r_d}{r}\right)^3\right) dt$, as q_p is a function of time this integral isn't a standard one, however it can be solved using partial integration;

$$u' = \left(\frac{d_e}{3v_p}\right) r\left(1 - \left(1 - \frac{r_d}{r}\right)^3\right) \qquad v = q_p(t)$$

$$u = \frac{d_e r_d}{v_p} t + r_0 \qquad v' = \frac{dq_p}{dt}$$

$$I = \left[q_p(t)\left(\frac{d_e r_d}{v_p} t + r_0\right)\right]_{t=0}^{t=t} - \int_0^t \left(\frac{d_e r_d}{v_p}\frac{dq_p}{dt} t + r_0 \frac{dq_p}{dt}\right) dt$$

$$\Rightarrow I = q_p(t)\left(\frac{d_e r_d}{v_p} t + r_0\right) - r_0 q_p(0) - \left(r_0 q_p(t) - r_0 q_p(0)\right) - \int_0^t \frac{d_e r_d}{v_p}\frac{dq_p}{dt} t \, dt$$

$$\Rightarrow I = q_p(t)\frac{d_e r_d}{v_p} t - \int_0^t \frac{d_e r_d}{v_p}\frac{dq_p}{dt} t \, dt$$

Taking the last integral and solving again by partial integration;

$$u = \frac{d_e r_d}{v_p} t \qquad v' = \frac{dq_p}{dt}$$

$$u' = \frac{d_e r_d}{v_p} \qquad v = q_p(t)$$

$$I = q_p(t)\frac{d_e r_d}{v_p} t - \left(q_p(t)\frac{d_e r_d}{v_p} t\right) + \int_0^t \frac{d_e r_d}{v_p} q_p(t) \, dt$$

$$I = \int_0^t \frac{d_e r_d}{v_p} q_p(t) \, dt$$

$$I = \frac{d_e r_d}{v_p}\left(Q_p(t) - Q_p(0)\right)$$

So;

$$r(t) - r_0 = k_g r_d t - \frac{d_e r_d}{v_p}\left(Q_p(t) - Q_p(0)\right)$$

$$r(t) = k_g r_d t + r_0 - \frac{d_e r_d}{v_p}\left(Q_p(t) - Q_p(0)\right)$$

Where $Q_p(t) - Q_p(0) = \int_0^t q_p(t) \, dt$, the integral of the drug concentration.

3.5.6 Analysisng Pharmacokinetic Systems

As given in Section 3.4, the pharmacokinetic system is defined by;

$$\frac{dq_p}{dt} = -(k_{pe} + k_{pt})q_p + k_{tp}q_{t,}$$

$$\frac{dq_t}{dt} = k_{pt}q_p - k_{tp}q_t$$

This can be rewritten as;

$$\begin{pmatrix} \dot{q_p} \\ \dot{q_t} \end{pmatrix} = \begin{pmatrix} -(k_{pe} + k_{pt}) & k_{tp} \\ k_{pt} & -k_{tp} \end{pmatrix} \begin{pmatrix} q_p \\ q_t \end{pmatrix}$$

The solution to this is given by;

$$\begin{pmatrix} q_p \\ q_t \end{pmatrix} = C_1 e^{\lambda_+ t} \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix} + C_2 e^{\lambda_- t} \begin{pmatrix} v_{21} \\ v_{22} \end{pmatrix}$$

where λ_+ and λ_- are the eigenvalues of $\begin{pmatrix} -(k_{pe} + k_{pt}) & k_{tp} \\ k_{pt} & -k_{tp} \end{pmatrix}$, and $\begin{pmatrix} v_{+1} \\ v_{+2} \end{pmatrix}$ and $\begin{pmatrix} v_{-1} \\ v_{-2} \end{pmatrix}$ are the corresponding eigenvectors. The values of λ_+ and λ_- are given by;

$$\lambda_{\pm} = \frac{-(k_{pe} + k_{pt} + k_{tp}) \pm \sqrt{{k_{pe}}^2 + {k_{pt}}^2 + {k_{tp}}^2 + 2(k_{pe}k_{pt} - k_{pe}k_{tp} + k_{pt}k_{tp})}}{2}$$

$$\Rightarrow \lambda_{\pm} = \frac{-(k_{pe} + k_{pt} + k_{tp}) \pm \sqrt{(k_{pe} - k_{tp})^2 + {k_{pt}}^2 + 2(k_{pe}k_{pt} + k_{pt}k_{tp})}}{2}$$

which are always unique as $(k_{pe} - k_{tp})^2 + {k_{pt}}^2 + 2(k_{pe}k_{pt} + k_{pt}k_{tp}) > 0$ as long as $k_{pe} > 0$, $k_{pt} > 0$, $k_{tp} > 0$.

They're both negative as ${k_{pe}}^2 + {k_{pt}}^2 + {k_{tp}}^2 + 2(k_{pe}k_{pt} - k_{pe}k_{tp} + k_{pt}k_{tp}) < (k_{pe} + k_{pt} + k_{tp})^2$, so $\begin{pmatrix} q_p \\ q_t \end{pmatrix} \to \begin{pmatrix} 0_+ \\ 0_+ \end{pmatrix}$ as $t \to \infty$.

Considering the eigenvectors of this gives;

$$\begin{pmatrix} -(k_{pe} + k_{pt}) & k_{tp} \\ k_{pt} & -k_{tp} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

Which gives;

$$\frac{k_{pt}}{(k_{tp} + \lambda)} v_1 = v_2$$

$$\begin{pmatrix} v_{+1} \\ v_{+2} \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{k_{pt}}{(k_{tp} + \lambda_+)} \end{pmatrix}; \begin{pmatrix} v_{-1} \\ v_{-2} \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{k_{pt}}{(k_{tp} + \lambda_-)} \end{pmatrix}$$

Considering the sign of $k_{tp}+\lambda_{\pm}=\frac{-(k_{pe}+k_{pt}-k_{tp})\pm\sqrt{{k_{pe}}^2+{k_{pt}}^2+{k_{tp}}^2+2(k_{pe}k_{pt}-k_{pe}k_{tp}+k_{pt}k_{tp})}}{2}$;

$$(k_{pe}+k_{pt}-k_{tp})^2=(k_{pe}+k_{pt})^2+{k_{tp}}^2-2k_{tp}(k_{pe}+k_{pt})$$

$$(k_{pe}+k_{pt}-k_{tp})^2={k_{pe}}^2+{k_{pt}}^2+{k_{tp}}^2+2(k_{pe}k_{pt}-k_{pe}k_{tp}-k_{pt}k_{tp})$$

$$(k_{pe}+k_{pt}-k_{tp})^2<{k_{pe}}^2+{k_{pt}}^2+{k_{tp}}^2+2(k_{pe}k_{pt}-k_{pe}k_{tp}+k_{pt}k_{tp})$$

So $k_{tp}+\lambda_+>0$, and $k_{tp}+\lambda_-<0$.

This gives $\binom{q_p}{q_t}=C_1e^{\lambda_+t}\begin{pmatrix}1\\ \frac{k_{pt}}{(k_{tp}+\lambda_+)}\end{pmatrix}+C_2e^{\lambda_-t}\begin{pmatrix}1\\ \frac{k_{pt}}{(k_{tp}+\lambda_-)}\end{pmatrix}$. Assuming the initial conditions are $\binom{q_{p0}}{0}$, the coefficients can be solved to give;

$$\binom{q_{p0}}{0}=C_1\begin{pmatrix}1\\ \frac{k_{pt}}{(k_{tp}+\lambda_+)}\end{pmatrix}+C_2\begin{pmatrix}1\\ \frac{k_{pt}}{(k_{tp}+\lambda_-)}\end{pmatrix}$$

This gives a simultaneous system defined by;

$$q_{p0}=C_1+C_2 \tag{22}$$

$$and\ 0=\frac{C_1k_{pt}}{(k_{tp}+\lambda_+)}+\frac{C_2k_{pt}}{(k_{tp}+\lambda_-)}$$

$$\Rightarrow C_1=-C_2\frac{(k_{tp}+\lambda_+)}{(k_{tp}+\lambda_-)} \tag{23}$$

Substituting (23) into (22) gives;

$$C_2\left(1-\frac{(k_{tp}+\lambda_+)}{(k_{tp}+\lambda_-)}\right)=q_{p0}$$

$$\Rightarrow C_2=\frac{q_{p0}}{\left(1-\frac{(k_{tp}+\lambda_+)}{(k_{tp}+\lambda_-)}\right)}>0$$

$$and\ C_1=-C_2\frac{(k_{tp}+\lambda_+)}{(k_{tp}+\lambda_-)}>0$$

So the solution to the pharmacokinetic system is;

$$\binom{q_p}{q_t}=\frac{q_{p0}}{\left(1-\frac{(k_{tp}+\lambda_+)}{(k_{tp}+\lambda_-)}\right)}\left[-\frac{(k_{tp}+\lambda_+)}{(k_{tp}+\lambda_-)}e^{\lambda_+t}\begin{pmatrix}1\\ \frac{k_{pt}}{(k_{tp}+\lambda_+)}\end{pmatrix}+e^{\lambda_-t}\begin{pmatrix}1\\ \frac{k_{pt}}{(k_{tp}+\lambda_-)}\end{pmatrix}\right]$$

$$\Rightarrow q_p(t) = \frac{q_{p0}}{\left(1 - \frac{(k_{tp}+\lambda_+)}{(k_{tp}+\lambda_-)}\right)}\left[-\frac{(k_{tp}+\lambda_+)}{(k_{tp}+\lambda_-)}e^{\lambda_+ t} + e^{\lambda_- t}\right]$$

Integrating w.r.t time gives;

$$Q_p(t) = \frac{q_{p0}}{\left(1 - \frac{(k_{tp}+\lambda_+)}{(k_{tp}+\lambda_-)}\right)}\left[-\frac{(k_{tp}+\lambda_+)}{\lambda_+(k_{tp}+\lambda_-)}e^{\lambda_+ t} + \frac{e^{\lambda_- t}}{\lambda_-}\right] + C$$

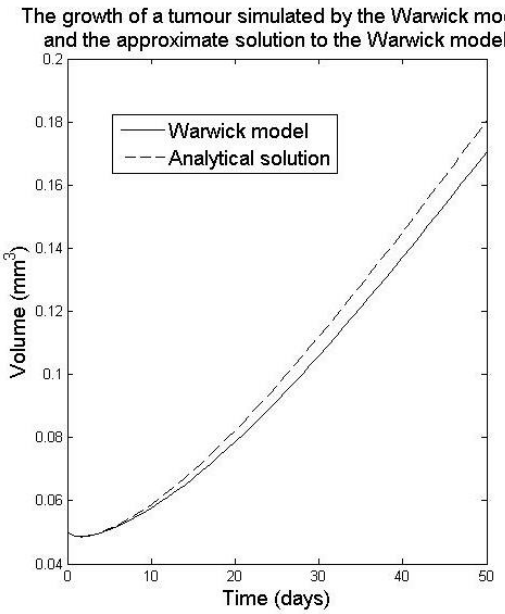
So the final solution to the radius is;

$$r(t) = k_g r_d t + r_0 - \frac{d_e r_d q_{p0}}{v_p\left(1 - \frac{(k_{tp}+\lambda_+)}{(k_{tp}+\lambda_-)}\right)}\left[-\frac{(k_{tp}+\lambda_+)}{\lambda_+(k_{tp}+\lambda_-)}(e^{\lambda_+ t} - 1) + \frac{(e^{\lambda_- t} - 1)}{\lambda_-}\right]$$

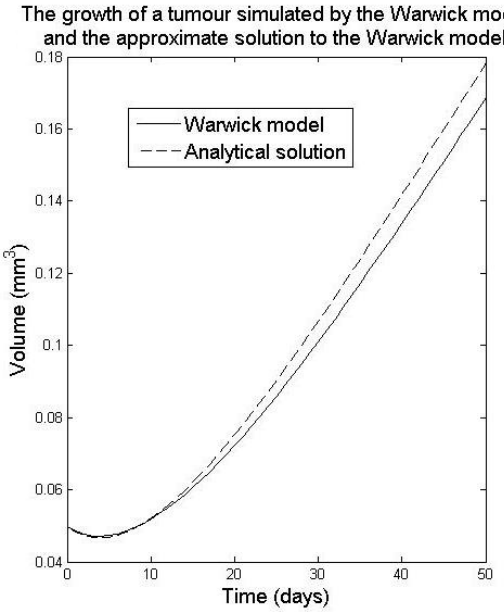
Where $\lambda_\pm = \frac{-(k_{pe}+k_{pt}+k_{tp}) \pm \sqrt{{k_{pe}}^2 + {k_{pt}}^2 + {k_{tp}}^2 + 2(k_{pe}k_{pt} - k_{pe}k_{tp} + k_{pt}k_{tp})}}{2}$.

Below are the simulations of this solution next to the Warwick model for a variety of different parameter values, with a graph simulating the PK effects on the system. The blue line on the PK graph indicates the concentration of drug at which the tumour starts growing instead of shrinking, found by solving $k_g - d_e\left(\frac{q_p(t)}{v_p}\right) = 0$.

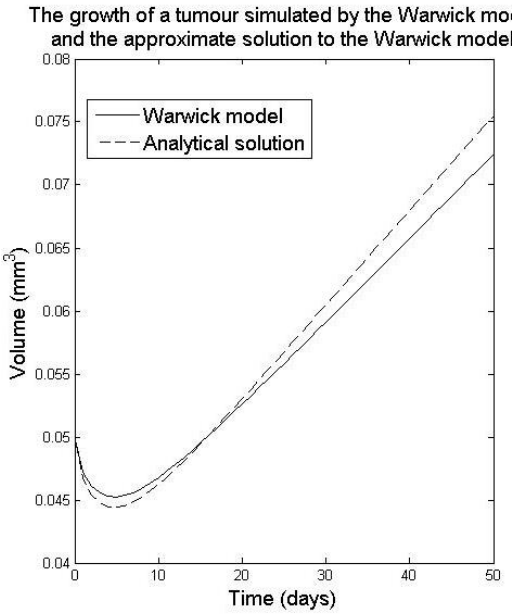
1) $[q_{p0}, q_{t0}, r_0, k_{pe}, k_{pt}, k_{tp}, v_p, k_g, d_e, r_d] = [1,0,0.05,0.1,0.3,0.5,1,0.5,0.8,0.0075]$



2) $[q_{p0}, q_{t0}, r_0, k_{pe}, k_{pt}, k_{tp}, \nu_p, k_g, d_e, r_d] = [1,0,0.05,0.1,0.1,0.8,1,0.5,0.8,0.0075]$



3) $[q_{p0}, q_{t0}, r_0, k_{pe}, k_{pt}, k_{tp}, \nu_p, k_g, d_e, r_d] = [10,0,0.05,0.5,0.5,0.8,10,0.1,0.8,0.0075]$



4) $[q_{p0}, q_{t0}, r_0, k_{pe}, k_{pt}, k_{tp}, \nu_p, k_g, d_e, r_d] = [10,0,0.05,0.5,0.5,0.8,10,0.1,0.8,0.0075]$

(over a time scale of 250 days)

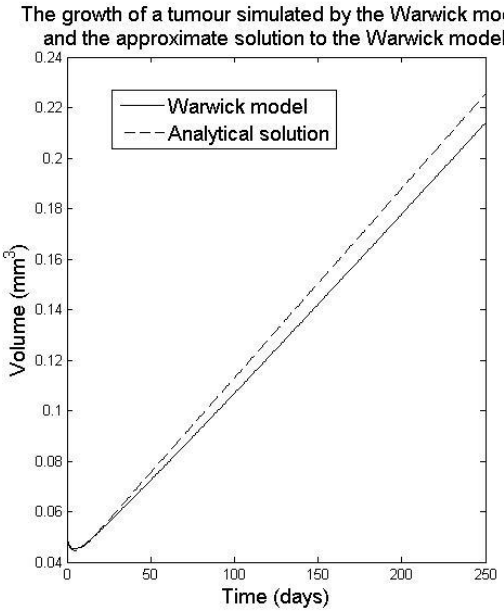


Figure 3-18: A comparison of the Warwick model and the analytical solution.

3.6 CONCLUSION

With the addition of pharmacokinetic drug effects, the analytical solution for the Warwick model based on the initial approximation to the Jumbe model still holds a reasonable degree of accuracy, however it grows less accurate over time. The solution to the PK system is accurate, as in all cases the solution was similar to the simulation from the model, as well as the minimum point of each of the growth simulations. In each case, the proliferating shell radius is 15% of the whole tumour radius which is large enough to be deemed significant. During the decaying sections of the model, the approximate solution under-estimated the actual model, and during the growth the approximate solution over-estimated the model, and the error between the two grew over time in each case.

This shows that an approximate analytical solution to the Warwick model does exist, which gives that the radius of the tumour grows linearly over time. The approximation is less accurate the larger the depth of the proliferating shell, which can be attributed to the splitting of the shell in to a series of small shells in Section 3.5.3. As there is a separate approximation for each of the smaller shells, they each generate additional errors, so having a larger initial shell means there is a larger number of smaller shells and hence a larger number of error terms.

4 KINASE MODEL

4.1 ABSTRACT

The phosphorylation of protein is a process which has many applications in biology, such as in enzyme regulation. The rate of the phosphorylation reaction is currently modelled using a simple linear ODE, and this model is used to predict the IC50 of a drug, however this model fails to accurately describe the reaction of the system so there's adequate reason to question the validity of the prediction made by it. A new non-linear model was proposed which fully defines the biology of the system, and this model is compared with the linear model to see how different factors in the system effects the prediction of IC50 from each model.

4.2 PREMISE

Protein phosphorylation is a biological process which alters the structure of a protein through the addition of a phosphate group, in the presence of a protein kinase; a protein kinase is an enzyme which acts as a catalyst for the reaction. The inverse of this process is de-phosphorylation, which is catalyzed by a phosphatase.

Under the effect of a drug which inhibits the rate of phosphorylation, the amount of phosphorylated protein (pP) in a system is often simulated by the model;

$$\frac{dpP}{dt} = \frac{1}{\tau}\left(I(q_p)pP_0 - pP\right) \tag{24}$$

where $I\left(q_p\right) = \frac{IC50}{IC50+q_p(t)}$, q_p is the amount of drug in the body's plasma, and the $IC50$ is defined as the concentration of drug at which halves the rate of phosphorylation. This model can be used in conjunction with clinical data to predict the $IC50$ of a drug, however it makes a number of assumptions about the system such as the rate of phosphorylation being constant and the rate of de-phosphorylation being first order, so this prediction might not be accurate in comparison to the true $IC50$ of the drug. Realistically the rate of phosphorylation is subject to a number of factors, such as the saturation levels of the kinase and phosphatase, and the capacity of both phosphorylated and unphosphorylated protein in the system.

An alternative model to simulate this system which takes into account the aforementioned factors, derived by James Yates (see Appendix), is given by;

$$\frac{d\widehat{pP}}{dt} = \frac{1}{\tau}\left[I(q_p)\left(\frac{1}{K_{eq}pP_0} - \frac{\widehat{pP}}{K_{eq}}\right)\frac{(K\prime_{m1}+1)}{\left(K\prime_{m1}+\frac{1}{K_{eq}pP_0}-\frac{\widehat{pP}}{K_{eq}}\right)} - \widehat{pP}\,\frac{K\prime_{m2}+1}{K\prime_{m2}+\widehat{pP}}\right], \tag{25}$$

where $\widehat{pP}$ is the fraction of the protein pool which is phosphorylated, with the initial condition $\widehat{pP}(0) = 1$. This normalizes the system, as to make comparison of the effects of changing parameters such as pP_0 easier. K'_{m1} and K'_{m2} are defined as the saturation levels of kinase and phosphatase in the system, and $K_{eq} = \frac{1-pP_0}{pP_0}$ is the fractional partitioning between the phosphorylated and dephosphorylated protein.

This model is derived by assuming the rate of phosphorylation and dephosphorylation are functions of the amount of unphosphorylated and phosphorylated protein respectively which is a more physically

realistic and justifiable assumption than the assumption of rates in the linear model. Due to this, comparing how the parameters $IC50$ and τ of each model vary through a number of simulations can be used to indicate how well the linear model can estimate $IC50$.

4.3 The Adapted Model

To check that the new model represents the system correctly a general simulation can be taken, then each of the parameters varied to note the effect of each on the system. If each affects the simulation in the predicted manner, and the simulation fits the profile of clinical data, then the adapted model is an accurate fit for the system.

The pharmacodynamic system for both the full and the linear model is given by;

$$\frac{dq_p}{dt} = -(k_{pe} + k_{pt})q_p + k_{tp}q_{t,}$$

$$\frac{dq_t}{dt} = k_{pt}q_p - k_{tp}q_t$$

Where k_{pe} is the extermination rate of drug out of the system, k_{pt} and k_{tp} are the transfer rates of the drug from the body plasma to tissue and visa versa, and q_p and q_t are the volumes of drug in the body plasma and tissue respectively. For the remainder of this section, the pharmacodynamics will be given by the parameters $\{k_{pe}, k_{pt}, k_{tp}\}$={0.25,0.3,0.2}, which is illustrated in Figure 4-1;

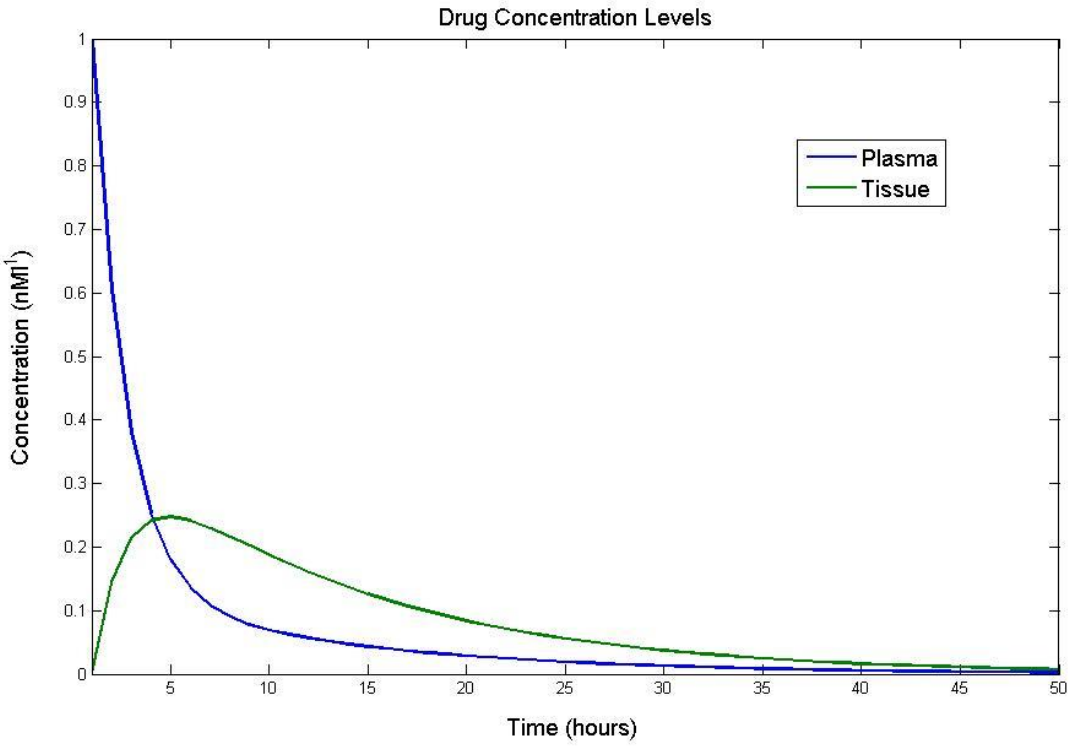


Figure 4-1: The pharmacokinetic system used in all of the simulations for this section.

The parameters used in the simulation are $\{K_{m1}, K_{m2}, pP_0, V_{max2}, IC50\}$[7], which fit into the model as $\tau = \frac{K_{m2}+pP_0}{V_{max2}}$, $K'_{m1} = \frac{K_{m1}}{pP_0}$, $K'_{m2} = \frac{K_{m2}}{pP_0}$, and $K_{eq} = \frac{1-pP_0}{pP_0}$. The reason these parameters were used as the parameters in the final non-Linear model are all functions of these, and it's easy to describe the effects of changing these parameters on the system. V_{max2} is the maximal speed the reaction of dephosphorylation is able to achieve, relative to the concentration of the dephosphorylated protein. K_{m1} and K_{m2} are the rate constants for the phosphorylation and dephosphorylation reactions, which are also representative of the saturation of the kinase and phosphatase. The former of these is inhibited by the reaction of the drug. pP_0 is the size of the protein pool in the system.

4.3.1 τ

τ is the time constant of the system, defined by $\tau = \frac{K_{m2}+pP_0}{V_{max2}}$, and to change this within the model the parameter V_{max2} will be edited. This can be done without effecting the rest of the equation, so for the later simulations V_{max2} will be used to keep τ constant. Below are the simulations for 3 different initial doses of a drug for two different values of τ. As $\frac{1}{\tau}$ is the time constant of the system, increasing τ slows down growth/decay of the system. This is shown in Figures 4-2 and 4-3, a lower value of τ means the minimum point of the simulation occurs earlier. The parameters used are $\{K_{m1}, K_{m2}, pP_0, V_{max2}$,IC50}={0.1,0.1,0.1,0.2,10}.

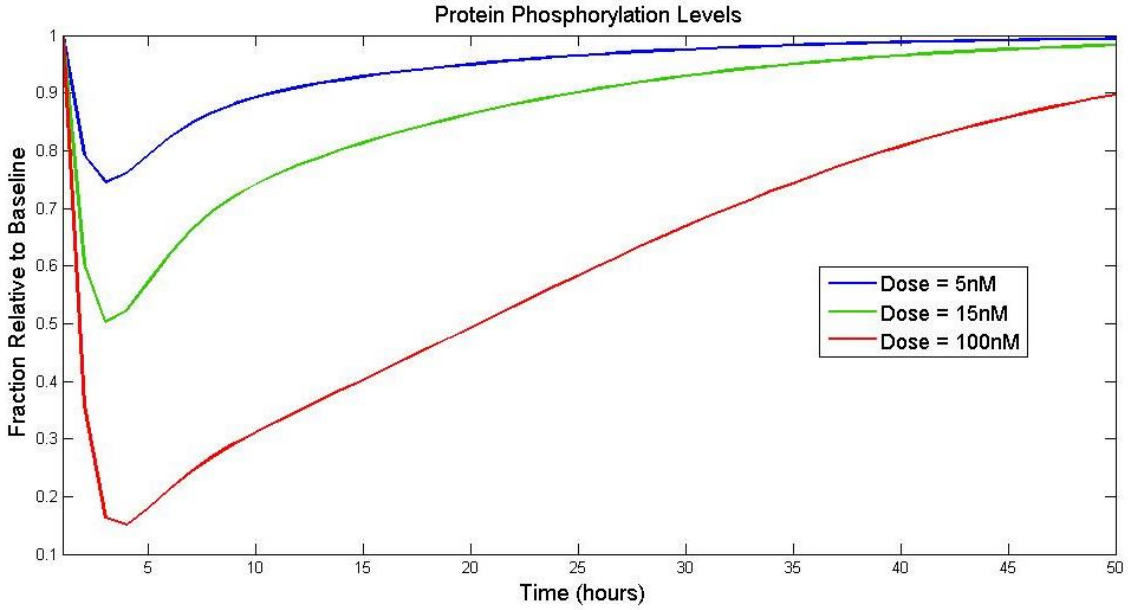


Figure 4-2: A simulation of the non-linear system under parameters which make τ=1.

[7] See derivation for relevance of these parameters.

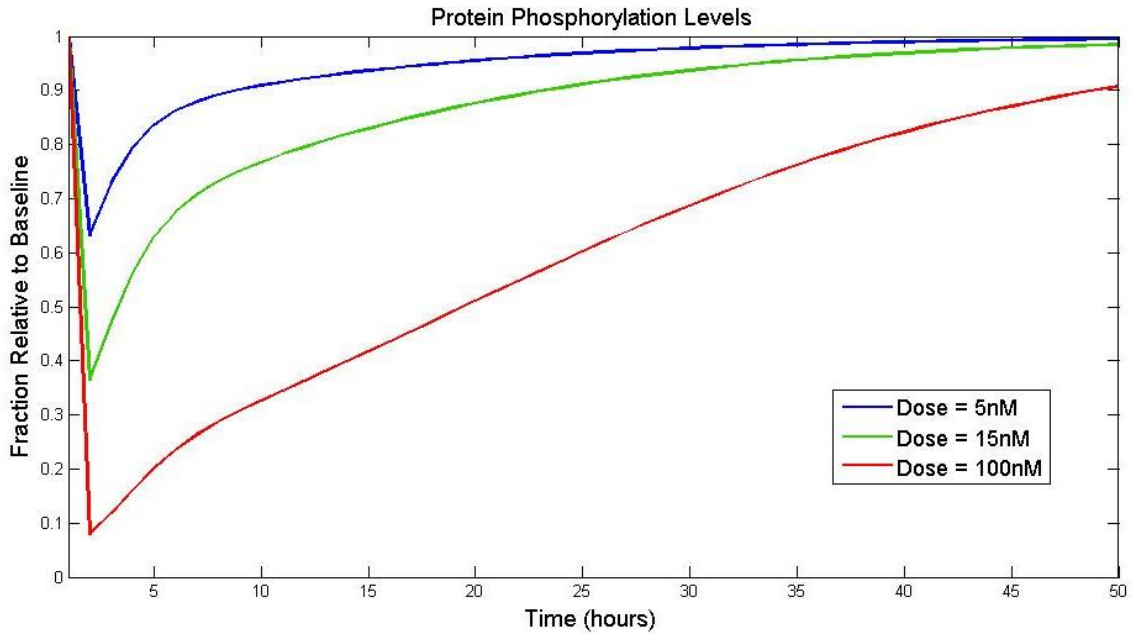


Figure 4-3: A simulation of the non-linear system under parameters which make τ=0.1.

4.3.2 pP_0

pP_0 is the size of the protein pool. Having a larger protein pool means that the minimum fraction of protein which is phosphorylated is lowered.

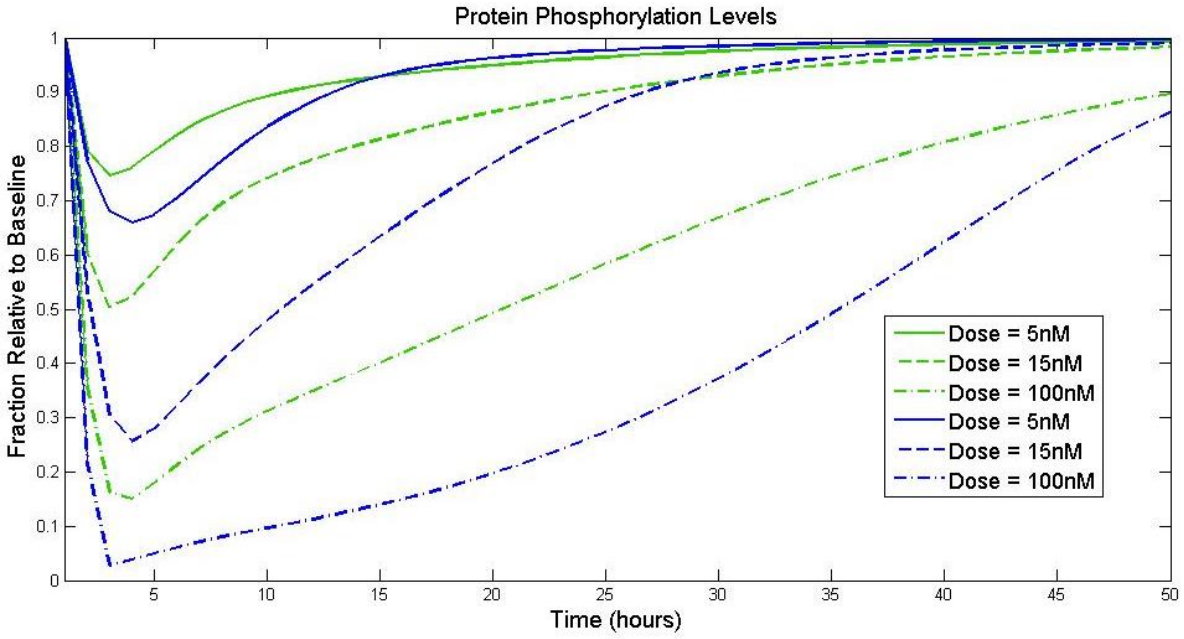


Figure 4-4: A simulation comparing the effect of increasing the size of the protein pool in the non-linear model for three different initial doses of drug.

Figure 4-4 shows that for a larger initial protein pool the fraction of phosphorylated protein was lower than for the smaller protein pool, so (25) reacts to variations in the protein pool as would be expected. The parameters used are {K_{m1}, K_{m2},IC50}={0.1,0.1,10}, with V_{max2} varying to keep τ constant. The green lines represent $pP_0 = 0.1$, and the blue lines respresent $pP_0 = 0.9$.

4.3.3 K'_{m1}

K'_{m1} is the normalised saturation of the kinase given by $K'_{m1} = \frac{K_{m1}}{pP_0}$, so to see the effect of varying K'_{m1} the parameter K_{m1} is varied. Figures 4-5 and 4-6 below show simulations for the model for variations in K'_{m1} across 3 different dosing schedules. Raising the value of K'_{m1} increases the rate of phosphorylation, and doesn't affect the rate of dephosphorylation, so the fraction of the protein which is phosphorylated is always slightly higher for larger values of K'_{m1}. The difference is minimal in the case of a smaller protein pool, and much greater for a larger protein pool. The more saturated the Kinase is, the more dramatic the effect of the drug is on the system. The parameters used below are $\{K_{m2}, V_{max2}, \text{IC50}\}=\{0.1,0.1,2,10\}$, with $pP_0 = 0.1$ in Figure 4-5 and $pP_0 = 0.9$ in Figure 4-6.

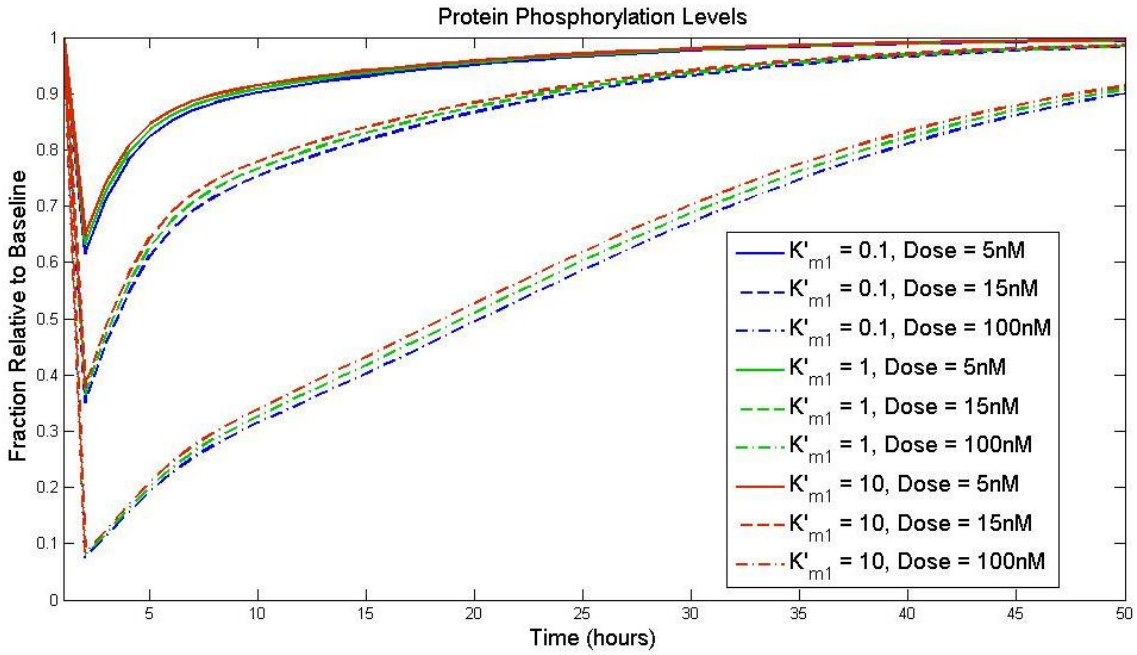


Figure 4-5: A comparison between the simulation of the non-linear model for three different saturation levels of the kinase, and three different dosing schedules.

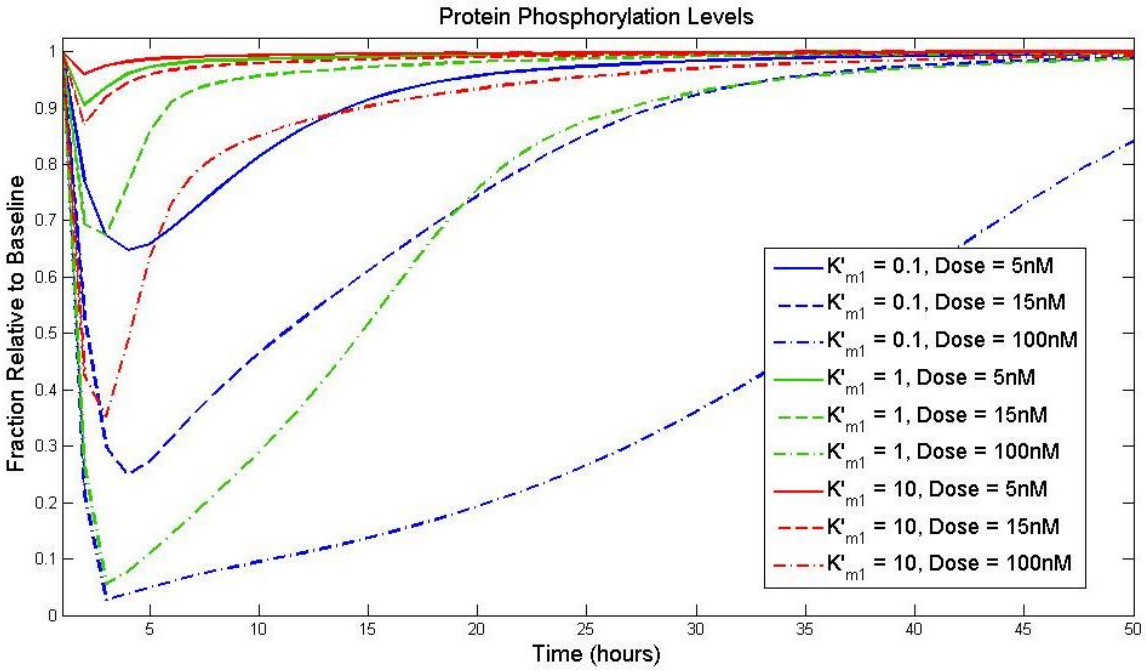


Figure 4-6:A comparison between the simulation of the non-linear model for three different saturation levels of the kinase, and three different dosing schedules.

4.3.4 K'_{m2}

K'_{m2} is the normalised saturation of the Phosphatase given by $K'_{m2} = \frac{K_{m2}}{pP_0}$, so to see the effect of varying K'_{m2} the parameter K_{m2} is varied. Similar to before, the more saturated the phosphatase, the larger the effect of the drug on the system. The parameters used below are $\{K_{m1}, V_{max2}, \text{IC50}\}=\{0.1,0.1,2,10\}$, with $pP_0 = 0.1$ in Figure 4-7 and $pP_0 = 0.9$ in Figure 4-8.

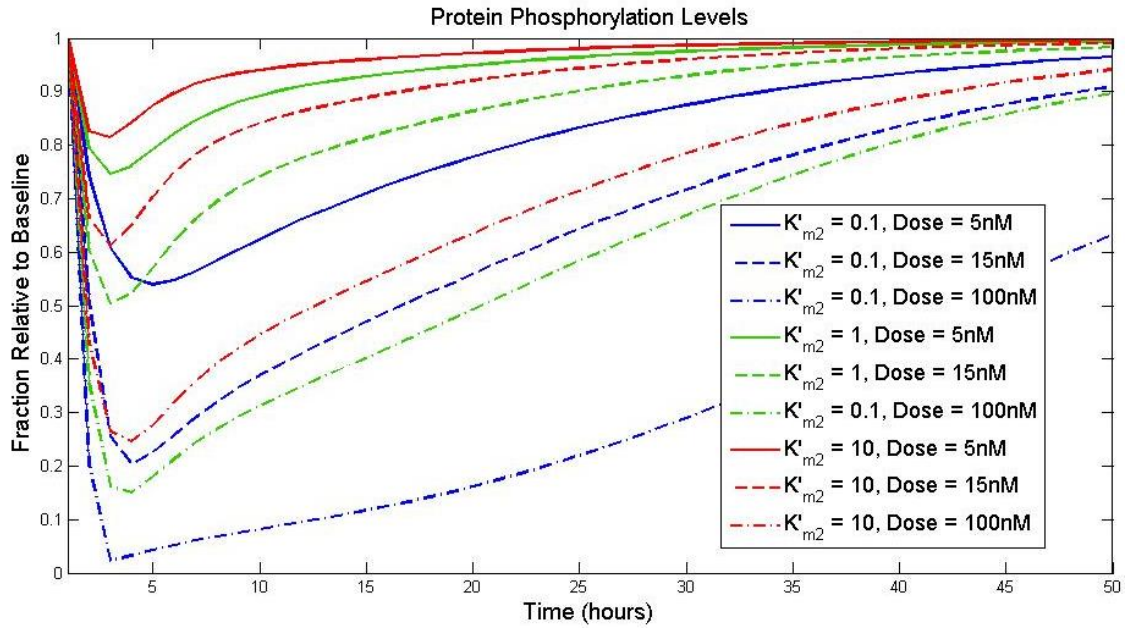


Figure 4-7: A comparison between the simulation of the non-linear model for three saturation levels of the phosphatase, and three different dosing schedules.

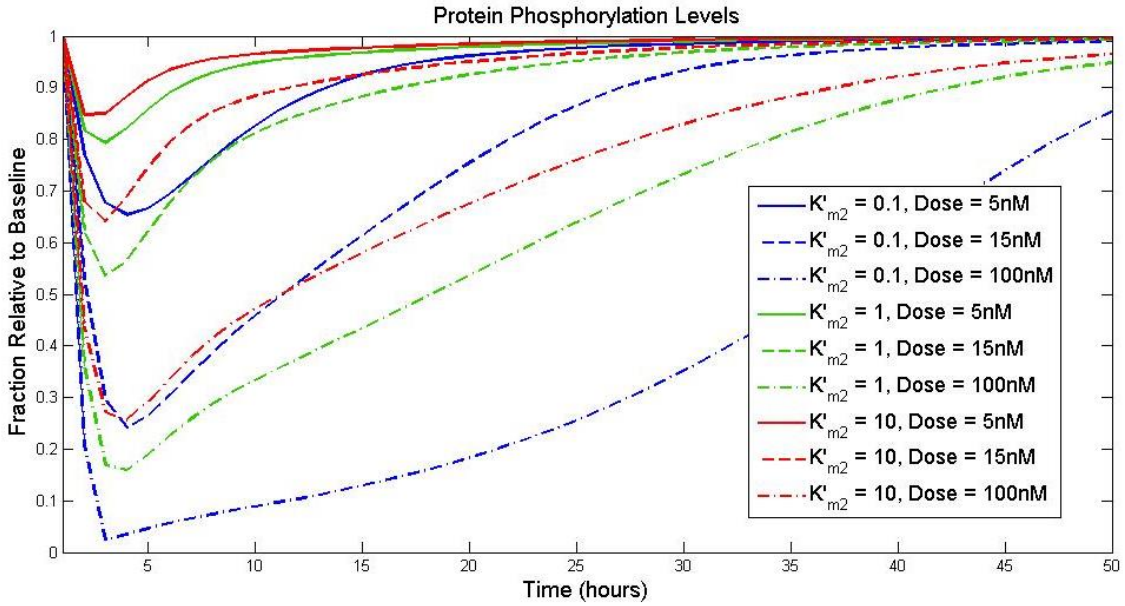


Figure 4-8: A comparison between the simulation of the non-linear model for three saturation levels of the phosphatase, and three different dosing schedules.

4.4 Comparisons Between the Two Models

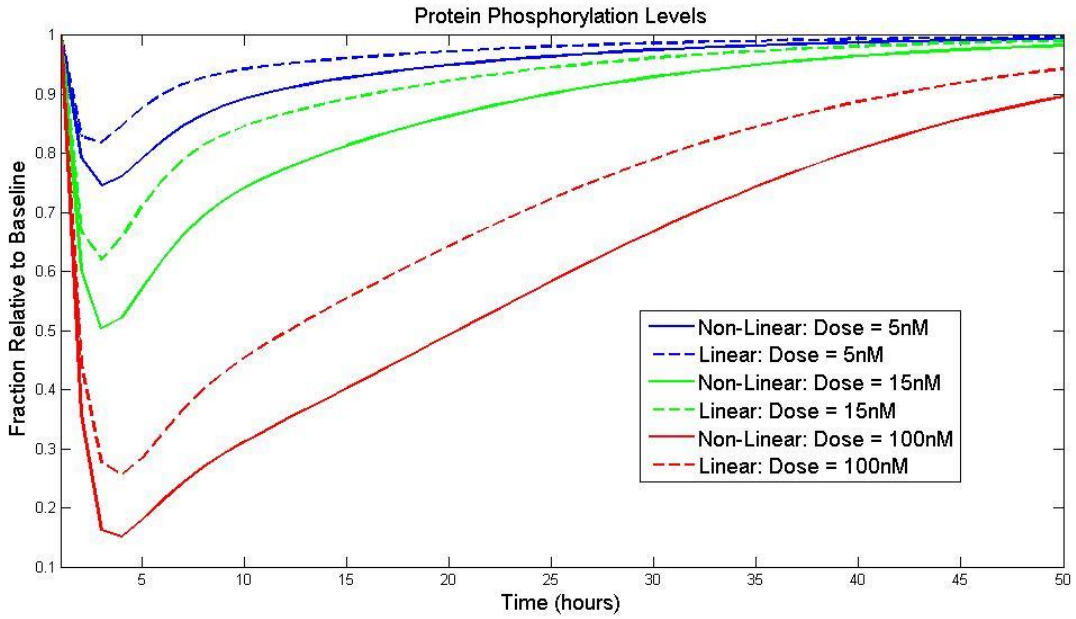


Figure 4-9: A comparison between the Linear and non-Linear models for 3 dose schedules, with the time constant and the IC50 of each system are equal.

Figure 4-9 shows the simulations from the Linear and non-Linear models which have the same pharmacokinetics, and the parameters τ and $IC50$ in the Linear model are the same as those in the non-Linear model. As shown, the Linear model overestimates the relative phosphorylation levels when compared with the non-Linear model, so in order to make the two models similar the parameter values of the Linear model must vary.

As the Linear model is widely used within clinical trials and research in order to estimate the $IC50$ of compounds, determining whether the difference between the value from the Linear model and the non-Linear model is important in determining how useful the Linear model is at accurately simulating the system.

To analyze the differences between the two, the Linear model can be fit to simulations of the non-Linear model, and then the difference between the IC50 of each model compared to see how changing the different parameters effects the validity of the Linear model. To differentiate between the two models, the $IC50$ for the non-linear and linear models will be denoted $IC50_{full}$ and $IC50_{lin}$ respectively.

4.4.1 Increased $IC50$

	Fig. 10: $IC50_{full} = 10nM$			Fig. 11: $IC50_{full} = 25nM$		
Dose	5	15	100	5	15	100
τ	1.4256	1.1366	0.8327	3.2951	2.8967	2.0166
$IC50_{lin}$	5.3951	5.3183	5.914	13.6625	13.3451	13.0146
$\frac{IC50_{lin}}{IC50_{full}}$	0.5395	0.5318	0.5914	0.5465	0.5338	0.5206

Table 4-1: An analysis of the difference between the simulated IC50 of the linear model, with the IC50 of the non-linear model, for two different values of IC50.

Table 4-1 shows the comparison between the fitted IC50 of the Linear model against the IC50 of the non-linear model. Changing the value of the IC50 in the full model did not effect the relative accuracy of the fitted IC50 along, as did changing the dose level of the drug.

4.4.2 Varying V_{max2}

	Fig. 12: $V_{max2} = 0.01$			Fig. 13: $V_{max2} = 0.1$			Fig. 14: $V_{max2} = 1$		
Dose	5	15	100	5	15	100	5	15	100
τ	20	20	20	2	2	2	0.2	0.2	0.2
τ_{lin}	32.4114	29.1330	22.5301	2.9820	2.4441	1.7686	0.1302	0.1416	0.1210
$IC50_{lin}$	5.4009	5.2988	5.1261	5.3598	5.2520	5.1990	5.3091	5.3692	5.3736
$\frac{IC50_{lin}}{IC50_{full}}$	0.5401	0.5299	0.5126	0.5360	0.5252	0.5199	0.5309	0.5369	0.5374

Table 4-2: An analysis of the difference between the simulated IC50 of the linear model, with the IC50 of the non-linear model, for three values of Vmax2.

V_{max2} is the maximum rate at which the phosphatase can work at. As Table 4-2 shows, varying the value of V_{max2} once again has little effect on the linear models estimate of $IC50_{lin}$, however it does change the estimate of the time constant of the system.

4.4.3 Varying K_{m1}

	Fig. 12: $K_{m1} = 0.01$			Fig. 13: $K_{m1} = 0.1$			Fig. 14: $K_{m1} = 1$		
Dose	5	15	100	5	15	100	5	15	100
τ	2	2	2	2	2	2	2	2	2
τ_{lin}	3.1939	2.5484	1.7858	3.1564	2.5308	1.7830	2.9820	2.4441	1.7686
$IC50_{lin}$	4.8361	4.7614	4.7776	4.9219	4.8402	4.8450	5.3598	5.2520	5.1990
$\frac{IC50_{lin}}{IC50_{full}}$	0.4836	0.4761	0.4778	0.4922	0.4840	0.4845	0.5360	0.5252	0.5199

Table 4-3: An analysis of the difference between the simulated IC50 of the linear model, with the IC50 of the non-linear model, for three values of Km1.

Table 4-3 shows that varying the value of K_{m1} does not affect the relative accuracy of the fitted IC50 of the Linear model to the IC50 of the non-Linear model. This shows that the saturation of the Kinase does not affect the accuracy of the Linear model.

4.4.4 Varying K_{m2}

	$K'_{m2} = 0.1$			$K'_{m2} = 1$			$K'_{m2} = 10$		
Dose	5	15	100	5	15	100	5	15	100
τ	1.1	1.1	1.1	2	2	2	11	11	11
τ_{lin}	2.9934	1.3397	0.6787	2.9820	2.4441	1.7686	11.331	11.241	11.961
$IC50_{lin}$	1.1164	1.0302	1.0277	5.3598	5.2520	5.1990	9.6234	9.5990	9.5760
$\frac{IC50_{lin}}{IC50_{full}}$	0.1116	0.1030	0.1028	0.5360	0.5252	0.5199	0.9623	0.9599	0.9576

Table 4-4: An analysis of the difference between the simulated IC50 of the linear model, with the IC50 of the non-linear model, for three values of K'm2.

Table 4-4 shows that decreasing the value of K_{m2} dramatically effects the correlation between the optimized value of $IC50_{lin}$ and the value of $IC50_{full}$. This shows the saturation of the Phosphatase affects the accuracy of the Linear model, indicating that the Linear model cannot accurately represent this aspect of the system.

4.4.5 Varying pP_0

	Fig. 12: $pP_0 = 0.01$			Fig. 13: $pP_0 = 0.1$			Fig. 14: $pP_0 = 0.5$		
Dose	5	15	100	5	15	100	5	15	100
τ	1.1	1.1	1.1	2	2	2	11	11	11
τ_{lin}	1.1676	1.1294	1.0750	2.9820	2.4441	1.7686	14.713	11.543	5.4225
$IC50_{lin}$	9.1683	9.1426	9.1500	5.3598	5.2520	5.1990	2.7872	2.3445	1.8241
$\frac{IC50_{lin}}{IC50_{full}}$	0.9168	0.91426	0.9150	0.5360	0.5252	0.5199	0.2787	0.2345	0.1824

Table 4-5: An analysis of the difference between the simulated IC50 of the linear model, with the IC50 of the non-linear model, for three values of pP0.

Table 4-5 shows that increasing the value pP_0 decreases the correlation between the optimized value of $IC50_{lin}$ and the value of $IC50_{full}$. However, as the derivation of the non-Linear model shows, $K'_{m2} = \frac{K_{m2}}{pP_0}$. This indicates that the only factor which affects the accuracy of the non-Linear model is the saturation of the phosphatase.

4.5 Conclusion

As shown above, the two parameters which affected the correlation between the $IC50_{lin}$ and $IC50_{full}$ were pP_0, the size of the protein pool, and K_{m2}, the amount of phosphatase in the system. By considering how these effect the different attributes of the system it can be concluded that the only factor which effects the correlation is the saturation of the phosphatase, given in the non-linear system by K'_{m2}.

Whilst this is based on the assumption that the non-linear model provides a better fit, there's strong evidence in the structure of each model to suggest this holds; as the linear model only has 2 parameters its likely to be unsuitable at representing such a diverse system.

Whilst the results shown from the fitting of the Linear model to the non-Linear model are arguably skewed by the fact that both the IC50 and τ are being fitted, it still holds significant reasoning that the Linear model is not an accurate representation of the protein phosphorylation system.

5 MODELLING TUMOURS AS ELLIPSOIDS

5.1 ABSTRACT

The majority of models for tumour growth assume that the tumour is spherical and maintains this shape through growth, this is a look into how the model can be adapted if the shape is assumed to be an ellipsoid instead; how the equations which define the system change, and if there's a general formula for the tumour in terms of each axis of a 3D Cartesian plane. Through polar coordinates and manipulation of the equation of an ellipsoid, an equation for each axis of the plane can be derived for a generic model.

5.2 METHODOLOGY

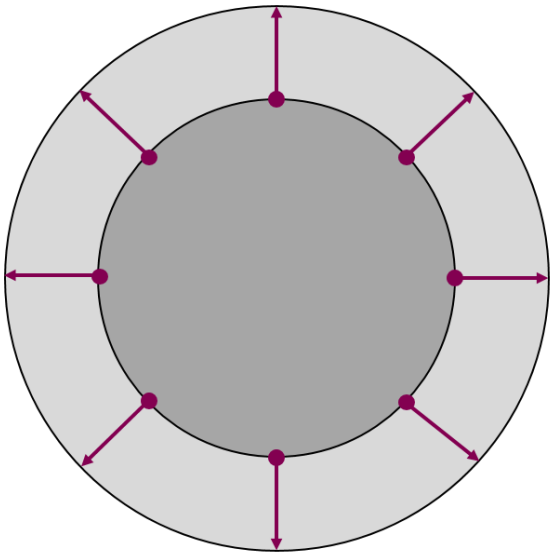
Using the initial condition that the x and the y axes are symmetric to each other, and the z axis is initially longer by a set ratio (r_a), the initial ellipsoid is generated at a set number of points (x, y, z) which were used to create a mesh of the shape. These points satisfy $x \in [-x_0, x_0], y \in [-x_0, x_0], z \in [-r_a x_0, r_a x_0]$, where x_0 is the x axis value when $y = z = 0$. Each of the initial set of (x, y, z) can then converted into a set of spherical polar coordinates (r, θ, ϕ), using the following system of equations;

$$r = \sqrt{x^2 + y^2 + z^2}$$

$$x = r \sin\phi \cos\theta$$

$$y = r \sin\phi \sin\theta$$

$$z = r\cos(\phi)$$



The set of radii $\{r\}$ was then extracted from these points, and each of the radius points was then used as the initial condition in the growth model to get a set of new radius points $\{\tilde{r}\}$. These were then paired up to the sets of (θ, ϕ) they initially corresponded, giving a new set of points $(\tilde{r}, \theta, \phi)$ which define the new set of (x, y, z) for the ellipsoid after the growth when converted back into Cartesian coordinates. This will be used later.

A solution for the axes will be looked for by rearranging the general equation for an ellipsoid, and using substitutions to isolate a single axial component.

5.3 Examining the General Equation of an Ellipsoid

The general form of an equation for an ellipsoid is;

$$\left(\frac{x(t)}{x_m(t)}\right)^2 + \left(\frac{y(t)}{y_m(t)}\right)^2 + \left(\frac{z(t)}{z_m(t)}\right)^2 = 1 \qquad (26)$$

where $x_m(t)$, $y_m(t)$, $z_m(t)$ are the functions describing the maximal values achieved by the x, y, and z axes respectively at time t.

By using the substitution $y_m(t) = x_m(t)$ and $z_m(t) = r_a x_m(t)$, the volume can be expressed as $V(x_m) = \frac{4\pi}{3} r_a {x_m}^3$. Considering $\frac{dV(x_m)}{dt}$ gives;

$$\frac{dV(x_m)}{dt} = \frac{dV(x_m)}{dx_m}\frac{dx_m}{dt},$$

$$\frac{dV(x_m)}{dt} = 4\pi r_a {x_m}^2 \frac{dx_m}{dt}.$$

By rearranging this and integrating with respect to t, a formula for x_m can be found, given by;

$$x_m = \int_0^t \frac{1}{4\pi r_a {x_m}^2}\frac{dV(x_m)}{dt}dt + x_0$$

Similar equations can be found for y_m and z_m given by;

$$y_m = \int_0^t \frac{1}{4\pi r_a {y_m}^2}\frac{dV(y_m)}{dt}dt + y_0$$

$$z_m = \int_0^t \frac{{r_a}^2}{4\pi {z_m}^2}\frac{dV(z_m)}{dt}dt + z_0$$

Taking the polar coordinates defined by $x = r\sin\phi\cos\theta, y = r\sin\phi\sin\theta, z = rcos\phi$ the following can be derived;

$$y = x\tan\theta \qquad (27)$$

$$z = x\cot\phi\sec\theta \qquad (28)$$

As seen in above, $x_m(t)$, $y_m(t)$, $z_m(t)$ can each be described by integrals, and when substituting these solutions into (26), it becomes;

$$\left(\frac{x(t)}{\int_0^t \frac{1}{4\pi r_a {x_m}^2}\frac{dV(x_m)}{dt}dt + x_0}\right)^2 + \left(\frac{y(t)}{\int_0^t \frac{1}{4\pi r_a {y_m}^2}\frac{dV(y_m)}{dt}dt + y_0}\right)^2 + \left(\frac{z(t)}{\int_0^t \frac{{r_a}^2}{4\pi {z_m}^2}\frac{dV(z_m)}{dt}dt + z_0}\right)^2 = 1$$

Substituting in (33) and (34) gives;

$$x^2(t)\left(\left(\frac{x(t)}{\int_0^t \frac{1}{4\pi r_a {x_m}^2}\frac{dV(x_m)}{dt}dt + x_0}\right)^2 + \left(\frac{\tan\theta}{\int_0^t \frac{1}{4\pi r_a {y_m}^2}\frac{dV(y_m)}{dt}dt + y_0}\right)^2 + \left(\frac{\cot\phi\sec\theta}{\int_0^t \frac{{r_a}^2}{4\pi {z_m}^2}\frac{dV(z_m)}{dt}dt + z_0}\right)^2\right) = 1$$

Rearranging this gives the final solution:

$$x_{\theta,\phi}(t) = \left(\left(\frac{1}{\int_0^t \frac{1}{4\pi r_a {x_m}^2}\frac{dV(x_m)}{dt}dt + x_0}\right)^2 + \left(\frac{\tan\theta}{\int_0^t \frac{1}{4\pi r_a {y_m}^2}\frac{dV(y_m)}{dt}dt + y_0}\right)^2 + \left(\frac{\cot\phi\sec\theta}{\int_0^t \frac{{r_a}^2}{4\pi {z_m}^2}\frac{dV(z_m)}{dt}dt + z_0}\right)^2\right)^{\frac{-1}{2}} \quad (29)$$

This is the solution for x at the position (θ, ϕ), so can be denoted $x_{\theta,\phi}(t)$. The corresponding solutions for y and z are;

$$y_{\theta,\phi}(t) = \left(\left(\frac{\cot\theta}{\int_0^t \frac{1}{4\pi r_a {x_m}^2}\frac{dV(x_m)}{dt}dt + x_0}\right)^2 + \left(\frac{1}{\int_0^t \frac{1}{4\pi r_a {y_m}^2}\frac{dV(y_m)}{dt}dt + y_0}\right)^2 + \left(\frac{\cot\phi\,\mathrm{cosec}\,\theta}{\int_0^t \frac{{r_a}^2}{4\pi {z_m}^2}\frac{dV(z_m)}{dt}dt + z_0}\right)^2\right)^{\frac{-1}{2}} \quad (30)$$

$$z_{\theta,\phi}(t) = \left(\left(\frac{\tan\phi\cos\theta}{\int_0^t \frac{1}{4\pi r_a {x_m}^2}\frac{dV(x_m)}{dt}dt + x_0}\right)^2 + \left(\frac{\tan\phi\sin\theta}{\int_0^t \frac{1}{4\pi r_a {y_m}^2}\frac{dV(y_m)}{dt}dt + y_0}\right)^2 + \left(\frac{1}{\int_0^t \frac{{r_a}^2}{4\pi {z_m}^2}\frac{dV(z_m)}{dt}dt + z_0}\right)^2\right)^{\frac{-1}{2}} \quad (31)$$

Using these axial solutions, the simulation generates an a set of points satisfying the equation $\left(\frac{x(t)}{x_m(t)}\right)^2 + \left(\frac{y(t)}{y_m(t)}\right)^2 + \left(\frac{z(t)}{z_m(t)}\right)^2 = 1$ so retains the shape of an ellipsoid, and due to the derivation of $x_m(t)$, $y_m(t)$, and $z_m(t)$ the ratio between these lengths remains the same.

The solutions are all defined based on the point of the tumour being considered, defined by the angles θ and ϕ, which also puts a limit on the region the solutions are valid over. When $\theta \in \left\{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\right\}$ or $\phi \in \left\{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\right\}$, the solutions equal to 0, or x_m, y_m, z_m dependent on which axis is being considered.

5.4 Conclusion

As seen in Section 5.3, a general solution for each of the axial components of a tumour can be found, defined by their position on the tumour. The solutions take the assumption that the tumour will keep an ellipsoid of constant ratio, however if the ratio was made a function of time then the solutions could be further adapted to include a variety of problems. As mentioned at the beginning of the chapter, tumours are often observed to growth into ellipsoids over time; this suggests that adapting these solutions to incorporate a ratio as a function of time could be useful at describing these problems, however this would cause $\frac{dV(x_m)}{dt} = \frac{dV(x_m)}{dx_m}\frac{dx_m}{dt}$ to no longer hold as partial derivatives would need to be considered instead.

The solutions given above, whilst limited, create a potential basis for further study however.

6 CONCLUDING STATEMENTS

Mathematical models are widely used across the study of oncology, and with this use comes a responsibility to understand how each model works, and the assumptions that are made in each of them. In the literature there are an increasingly large number of models, here we compared five which are widely popular along with two that were of interest to AstraZeneca to compare features which make models successful at representing tumour growth. Through fitting the models to 48 sets of growth data from a selection of 4 clinical trials, the models provided the best fit were derived from the principle of allometric scaling. The analysis showed that the more complex models tended to have a harder time providing an accurate fit for the data, which shows that although a model needs to successfully describe the biological systems of tumour growth there is a level of complexity at which the usefulness of the model begins to decline. This correlates well with the industrial applications of a model, as for a model to be reliable for use it need to be easy to use on a large scale; whilst an incredibly intricate model may look promising it may be difficult to use and lead to a variety of computational issues.

Examining links between models can give insight into the mechanisms of them, and help better understand them. This was illustrated by the approximation of an analytical solution to the Warwick model, which was found by examining a link from the Warwick model to the simpler Jumbe model. This was done by taking the limit of the proliferating shell depth in the Warwick model, which reduced the Warwick model to the Jumbe. By taking a value for the parameter that made the models equivalent for the initial condition and considering small increments to this value, a parameter value could be found which minimizes the difference between the two models and allowed for an approximate solution to the Warwick model. By then considering splitting a general proliferating shell into a finite number of shells with a small enough depth to allow the approximation to hold, an approximate solution could be found for the Warwick model which holds for any shell depth.

The applications of mathematical models within industry don't lie solely in tumour growth models however; models are used to describe numerous different biological and chemical systems, and similarly to growth models ensuring that the models used correctly represent the systems they're simulating is crucial to the validity of results which are made using those models. An example of an unsuitable model was the model widely used to simulate the rate of phosphorylation of a protein in the presence of kinase and phosphatase inhibitors; the most commonly used model failed to take into account many of the factors which affected the system, so was assumed to be a bad representation of the system. A new model was derived from the mechanics of the system, which shared two parameters with the original model and had a selection of new ones. By varying the new parameters and fitting the original model to the new one it was shown that the original model failed to change to certain changes in the system, such as changes to the level of saturation of the phosphatase, showing it wasn't capable of representing the system completely.

Whilst the majority of tumour models assume a tumour is spherical, it has been observed in many cases that overtime the tumours have grown to be more ellipsoidal overtime. Due to this, seeing if growth models can be adapted to fit an ellipsoidal shape could provide useful insight into how tumours grow in these instances. Through assuming a tumour holds ellipsoidal shape, and using polar coordinates, a general formula for growth in each axis of a 3D Cartesian plane can be found by manipulating the fundamental equation of an ellipsoid and the equation for the volume of an ellipsoid. This has potential

applications into describing specific types of tumour growth, and whilst very primitive and limited could be further developed into a larger model.

If this project was carried on further, the most prominent next step would be to analyse a wider range of functions in order to strengthen the conclusion to the comparison from Chapter 2. Considering functions with different mechanics to allometric scaling processes would allow for a stronger conclusion. With regards to the protein phosphorylation model, fitting both the linear and the revised models to observed clinical data would better justify the results, and help prove or disprove the assumption that the linear model is unsuitable at representing the system.

7 ACKNOWLEDGEMENTS

I'd like to thank James W T Yates, my manager at AstraZeneca, and Dr Carina Dunlop, my supervisor at the University of Surrey, for supporting and guiding me for the duration of this project.

8 BIBLIOGRAPHY

Angelo. L. S, K. R. (2009). Target validation using Phase 0 clinical trials: Promises and pitfalls. *Cancer Biol Ther, 8*(21), 2010-2012.

Anisimov. V. N, U. S. (2005). Cancer in rodents: does it tell us about cancer in humans? *Nat Rev Cancer, 5*(10), 807-819.

Aranujo, R., & McElwain, L. (2004). A History of the Study of Tumour Growth: The Contribution of Mathematical Modelling. *Bulletin of Mathematical Biology, 66*(1), 1039-1091.

Begley. C. G, E. L. (2012). Drug Development: Raise standards for preclinical cancer research. *Nature, 483*, 531-533.

Burnham. K, A. D. (2011). AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behav Ecol Sociobiol, 65*, 23-35.

Cook. N, J. D. (2012). Predictive in vivo animal models and translation to clinical trials. *Drug Discovery Today, 17*, 253-260.

Evans, N., Dimelow, R., & Yates, J. (2014). Modelling of tumour growth and cytotoxic effect of docetaxel in xenographs. *Computer Methods and Programs in Biomedicine, 114*(1), 3-13.

Gan. H. K, e. a. (2012). Assumptions of Expected Benefits in Randomized Phase III Trials Evaluating Systemic Treatments for Cancer. *J Natl Cancer Inst, 104*(8), 590-598.

Gatenby, R., & Maini, P. (2003). Cancer summed up. *Nature, 421*, 321.

Glimelius. B, L. M. (2011). Window-of-opportunity trials to evaluate clinical activity of new molecular entities in oncology. *Ann Oncol, 22*(8), 1717-1725.

H, M. S. (1983). Model for Human Carcinogenesis: Action of Environmental Agents. *Environmental Health Perspectives, 50*, 283-291.

Hahn. W. C, W. R. (2002). Modelling the molecular circuitry of cancer. *Nat Rev Cancer, 2*(5), 331-341.

Johnson, L. (2012). Cancer: Clinical trials unite mice and humans. *Nature, 483*, 546-548.

Johnson. J, e. a. (2001). Relationships between drug activity in NCI preclinical in vitro and in vivo models and early clinical trials. *British Journal of Cancer, 84*(10), 1424-1431.

Jong. M, M. T. (2019). Of mice and humans: are they the same? Implications in cancer translational research. *J Nucl Med, 51*(4), 501-504.

Jumbe, N., & et al. (2010). Modeling the efficacy of trastuzumab-DM1, an antibody drug conjugate, in mice. *Journal of Pharmacokinetics and Pharmacodynamics, 37*, 221-242.

Kleiman. R. J, E. M. (2016). Data gaps limit the translational potential of preclinical research. *Sci Trans Med, 8*, 1-5.

L, J. (2012). Cancer: Clinical trials unite mice and humans. *Nature, 483*, 546-548.

Mager. D. E, W. S. (2009). Scaling pharmacodynamics from in vitro and prelinical animal studies to humans. *Drug Metab Pharmacokinet, 24*(1), 16-24.

Mak. I. W, E. N. (2014). Lost in translation: animal models and clinical trials in cancer treatment. *Am J Transl Res, 6*(2), 114-118.

Marusic, M. (1994). Tumor growth in vivo and as multicellular spheroids compared by mathematical models. *Bull Math Biol, 56*(4), 617-631.

Marusic, M. (2006, April 5). *Mathematical Communications, Vol. 1, No. 2*. Retrieved from hrcak: Portal of Scientific Journals of Croatia: http://hrcak.srce.hr/1837?lang=en

Marusic, M., & Bajzer, Z. (1993). Generalized Two-Parameter Equation of Growth. *Journal of Mathematical Analysis and Applications, 179*, 446-462.

Mestas. J, H. C. (2004). Of mice and not men: differences between mouse and human immunology. *J Immunol, 172*(5), 2731-2738.

Moolgavkar, S. (1983). Model for Human Carcinogenesis: Action of Environmental Agents. *Environmental Health Perspectives, 50*, 285-291.

Quaranta , V. (2005). Mathematical modeling of cancer: The future of prognosis and treatment. *Clinica Chimica Acta, 357*, 173-179.

Rangarajan. A, W. R. (2003). Opinion: Comparative biology of mouse versus human cells: modelling human cancer in mice. *Nat Rev Cancer, 3*(12), 952-959.

Sakai. I. S, Y. Y. (2014). Species Differences in the Pharmacokinetic Parameters of Cytochrome P450 Probe Substrates between Experimental Animals, such as mice, rats, dogs, monkeys, and microminipigs, and humans. *J Drug Metab Toxicol, 5*(6), 1-12.

Sena. E. S, e. a. (2010). Publication bias in reports of animal stroke studies leads to major overstatement of efficacy. *PLoS Biol, 8*(3), 1-8.

Singh. B, G. R. (1985). Species-specific differences in the toxicity and mutagenicity of the anticancer drugs mithramycin, chromomycin A3, and olivomycin. *Cancer Res, 45*(6), 2813-2820.

Tsoularis, A., & Wallace, J. (2002). Analysis of logistic growth models. *Mathematical Biosciences, 179*, 21-55.

West, G., Brown, J., & Enquist , B. (1997). A general model for the origin of allometric scaling laws in biology. *Science, 276*, 122-126.

Wong. H, e. a. (2012). Antitumor activity of targeted and cytotoxic agents in murine subcutaneous tumor models correlates with clinical response. *Clin Cancer Res, 18*(14), 3846-3855.

9 APPENDIX

9.1 PARAMETER ESTIMATION

9.1.1 Model Parameters

9.1.1.1 GCATR1434

	Gompertz (with death)			Generic			
Mouse	a	b	d	β	K	n	p
1	0.6708	0.0206	0.8151	0.1296	0.0046	0.2259	0.1352
4	0.9355	0.0535	1.1582	0.1490	0.0161	0.2901	0.1017
5	0.0656	0.0005	0.1200	0.1289	0.0098	0.2243	0.1355
6	0.5372	0.0194	0.6524	0.1292	0.0082	0.2255	0.1350

	Generalised Two-Parameter				Autostimulation				
Mouse	a	b	α	β	a	b	α	β	ω
1	15.6334	1.9343	2.7849	1.8093	0.0832	1.1097	0.0527	1.1121	0.1318
4	0.3953	0.5115	0.5983	0.6656	1.5025	0.4986	0.2302	4.7730	0.3224
5	0.0026	0.0182	0.1914	0.4744	2.1514	2.7053	1.2818	0.7167	1.5407
6	0.1282	0.1122	0.0794	0.0193	0.3238	4.3680	0.0206	0.9356	0.0778

	Generalised Bertalanffy-Logistic		
Mouse	a	b	α
1	0.9087	1.0580	1.0219
4	0.8727	1.1061	1.0563
5	1.0163	1.0726	0.9995
6	0.9765	1.0930	1.0187

9.1.1.2 GCERK1420

	Gompertz (with death)			Generic			
Mouse	a	b	d	β	K	n	p
1	1.546	0.0606	1.4445	1.0981	3.3371	0.0351	-0.2669
2	1.6032	0.0553	1.4831	0.3462	3.3287	0.3236	-0.2516
3	1.5784	0.0540	1.4669	0.3678	2.5698	0.1859	-0.4599
5	1.5594	0.0582	1.4722	0.4435	3.8302	0.1802	0.0636
6	1.2656	0.0558	1.1547	0.1125	2.9370	2.6466	0.0575
7	1.5368	0.0421	1.4178	0.3142	2.5945	0.3722	-0.5193
8	1.3875	0.0459	1.2708	0.2159	3.7567	0.8398	-0.1394
9	1.4071	0.0721	1.3031	0.7602	3.2283	0.0916	-0.1515
10	1.4232	0.0390	1.3409	0.3989	6.5310	0.1313	0.0410

	Generalised Two-Parameter				Autostimulation				
Mouse	a	b	α	β	a	b	α	β	ω
1	1.0972	0.9943	0.9945	1.0539	0.9694	0.2961	0.7097	1.3737	0.6489
2	1.0673	0.9456	1.0129	1.0742	0.8819	0.7934	0.4491	0.8138	0.3442
3	1.0791	0.9664	1.0062	1.0638	1.1002	0.5162	1.0484	0.8678	1.1376
5	1.0434	0.9547	1.0523	1.1234	1.9489	3.4969	0.3463	0.9309	0.2173
6	1.0757	0.9637	1.0139	1.0738	0.399	0.5623	0.3496	0.9778	0.1854
7	1.0703	0.9500	1.0244	1.0728	1.1348	0.3059	0.9912	0.9643	1.1487
8	1.1322	0.9999	1.4086	1.4967	2.5347	1.21	0.2266	0.9481	0.1497
9	1.0719	0.9594	1.1792	1.2784	0.1490	2.3414	0.9843	0.2552	0.8382
10	1.4013	1.3198	0.9145	0.9375	1.4796	0.9962	0.3203	0.9485	0.2692

	Generalised Bertalanffy-Logistic		
Mouse	a	b	α
1	1.8319	1.7293	1.0348
2	1.1229	1.0019	1.0561
3	1.0853	0.9728	1.0562
5	1.0990	1.0113	1.0590
6	1.1056	0.9940	1.0565
7	1.1342	1.0147	1.0420
8	1.1218	1.0045	1.0456
9	1.0898	0.9847	1.0739
10	2.9185	2.8361	1.0138

9.1.1.3 GCGLS1404

	Gompertz (with death)			Generic			
Mouse	a	b	d	β	K	n	p
1	0.0452	0.0028	0.0007	0.0614	3.1330	1.7431	-0.0841
2	0.1194	0.0004	0.0672	0.0092	15.6643	0.0840	-2.7827
3	0.0435	0.0028	0.0007	0.4290	2.2201	0.0023	-0.6051
4	0.1214	0.0005	0.0643	0.1909	2.1789	2.4821	-0.3742
6	0.0629	0.0029	0.0006	0.5178	4.8554	0.0000[8]	-0.8406
7	1.4218	0.0245	1.3550	0.1925	3.8659	0.3361	-0.2404
8	1.5245	0.0088	1.4648	3.7075	1.9515	0.0000	-0.7269
9	1.4123	0.0195	1.3386	2.6290	1.8796	0.0002	-0.6042
10	1.7161	0.0006	1.6489	0.1252	3.8018	0.0002	-0.9243
11	0.1221	0.0008	0.0558	1.8825	3.2386	0.0000	-0.7893

[8] 0.0000 is the value to 4d.p.

	Generalised Two-Parameter				Autostimulation				
Mouse	a	b	α	β	a	b	α	β	ω
1	1.0128	0.9674	1.0397	1.0419	0.0115	1.7485	0.5355	0.1137	0.4707
2	1.0496	0.9674	1.0642	1.0218	1.0851	1.2338	0.9232	0.4346	1.1620
3	1.0172	0.9805	1.0343	1.0491	1.3159	1.1391	0.5676	0.8577	0.6279
4	1.0658	0.9854	1.0829	1.0419	1.9976	1.7325	0.6377	0.3636	0.8712
6	1.6579	1.5220	1.0376	0.9853	4.1226	3.2977	0.4875	0.2833	0.6757
7	2.1477	2.0672	1.2665	1.2914	1.1535	1.1195	1.0643	0.5454	1.2505
8	1.0371	0.9775	1.0402	1.0524	1.1311	1.2069	0.8775	0.5627	1.0366
9	1.0774	1.0036	1.0218	1.0437	0.0675	2.5203	1.6599	0.1419	1.5978
10	1.0342	0.9671	1.0459	1.0500	1.6927	1.4601	0.4410	0.4860	0.5350
11	1.0654	0.9850	1.0585	1.0339	0.9985	1.1481	0.5743	0.3631	0.7201

	Generalised Bertalanffy-Logistic		
Mouse	a	b	α
1	1.0396	0.9942	0.9997
2	1.0754	1.0232	1.0000
3	1.1301	1.0933	1.0115
4	1.1469	1.0898	1.0000
6	1.1106	1.0463	1.0000
7	1.0190	0.9421	1.0262
8	1.1039	1.0442	1.0085
9	1.0936	1.0198	1.0193
10	1.0927	1.0254	1.0006
11	1.0403	0.9740	1.0000

9.1.1.4 GCERK1421

	Gompertz (with death)			Generic			
Mouse	a	b	d	β	K	n	p
1	1.4354	0.0247	1.3562	0.1974	2.7520	0.4084	-0.1174
2	1.3057	0.0336	1.2027	5.9083	3.4200	0.0002	-0.5173
3	1.7951	0.0398	1.7369	0.6665	3.3072	0.0683	-0.0332
5	1.6897	0.0477	1.6049	0.4197	3.2308	0.1627	-0.0983
6	1.3631	0.1072	1.2543	0.1864	2.6535	0.8782	0.2067
8	1.5612	0.0301	1.4623	0.5679	4.1417	0.0239	-0.4933
9	1.4674	0.0740	1.3923	0.0574	9.2129	0.5415	0.7213
10	1.1360	0.0547	1.0255	1.3624	6.6534	0.0474	0.0233
11	1.6435	0.0602	1.5732	0.0760	2.8510	1.2161	0.2565
12	0.9155	0.0404	0.8380	0.2387	1.6889	0.5199	-0.2580
13	1.3974	0.0481	1.3051	2.0808	1.6818	0.0017	-0.5541
14	1.6536	0.0244	1.5426	6.1818	2..9870	0.7854	-0.0169
15	1.5306	0.0523	1.4480	0.3544	1.0543	0.7208	-0.3461
16	1.1955	0.0865	1.1114	0.4914	3.0043	0.2336	0.2338
17	1.4913	0.0583	1.4300	0.5834	1.8928	0.2292	-0.0486
19	1.5096	0.0476	1.4126	0.3478	3.9189	0.2292	-0.0486

	Generalised Two-Parameter				Autostimulation				
Mouse	a	b	α	β	a	b	α	β	ω
1	1.6021	1.5962	1.7373	1.9218	0.3787	9.6244	0.9585	0.5435	0.8551
2	1.0530	0.9500	0.9929	1.0280	1.0292	1.3167	0.9853	0.6315	1.0050
3	1.0438	0.9860	0.9908	1.0316	0.6498	2.0064	0.9752	0.6477	0.8788
5	0.3216	0.2384	1.0483	1.3509	1.6019	0.7993	0.3756	1.6274	0.2618
6	0.9859	0.8735	1.0243	1.1521	0.2866	1.0977	0.5629	1.3259	0.2690
8	0.7857	0.6888	0.7226	0.7226	1.0286	0.8245	0.9990	0.6700	1.1063
9	0.5162	0.4432	0.3345	0.3345	1.3160	1.7836	0.8844	0.731	0.8546
10	1.5609	1.4525	0.4882	0.4882	2.3135	0.8854	0.2242	1.6074	0.1009
11	0.6856	0.6381	1.2967	1.5915	1.0799	1.4496	0.6232	1.1629	0.4874
12	0.8601	0.7838	1.3213	1.4515	1.2486	1.5236	0.2596	1.2654	0.1414
13	0.7797	0.6831	1.5525	1.9692	1.1361	1.3210	0.2693	0.8520	0.7036
14	1.0076	0.8974	1.0250	1.0584	1.3117	1.4139	0.4620	0.9295	0.3706
15	0.7624	0.7067	1.3318	1.6170	1.6364	1.1087	0.8154	1.0671	0.7167
16	1.0511	0.9669	0.9628	1.0492	1.3611	0.3397	0.5208	5.5691	0.1447
17	0.9804	0.257	1.1001	1.1952	0.9484	0.9835	1.0435	1.1352	0.9667
19	1.0770	0.9801	0.9677	1.0124	2.5522	0.7623	0.5464	1.2761	0.5619

	Bertalanffy		
Mouse	a	b	α
1	1.0036	0.9249	1.0277
2	1.0182	0.9142	1.0378
3	1.0101	0.9522	1.0438
5	1.0040	0.9195	1.0545
6	1.0526	0.9415	1.1151
8	0.0785	0.8804	1.0346
9	4.3528	4.2777	1.0175
10	1.0402	0.9289	1.0588
11	1.0410	0.9719	1.0667
12	0.2645	0.1878	1.2587
13	0.1848	0.0974	2.2662
14	1.0839	0.9733	1.0259
15	1.0723	0.9909	1.0566
16	1.0353	0.9480	1.0934
17	1.0264	0.9667	1.0651
19	1.0824	0.9711	0.9981

9.1.1.5 GCGLS1433

	Gompertz (with death)			Generic			
Mouse	a	b	d	β	K	n	p
1	1.3552	0.0550	1.2569	0.6237	4.7051	0.1279	0.1133
2	1.6284	0.0608	1.4916	1.0554	3.1834	0.0603	-0.2616
3	1.3953	0.0255	1.1621	3.9864	39.3035	0.0009	-0.5088
4	1.3058	0.0059	1.1297	0.3695	0.5310	0.3958	-0.3002
6	1.3782	0.0833	1.1343	1.6754	3.2523	0.0256	-0.4683
8	1.6256	0.0214	1.3556	6.5860	4.1767	6.3317	-0.2544
9	1.5754	0.0604	1.3871	0.1691	6.0019	1.0091	0.1632
10	1.5053	0.0875	1.3379	0.1119	4.8696	2.3695	0.1265
17	1.3779	0.0435	1.2576	0.2381	2.0272	1.8646	-0.0241
18	1.2068	0.0429	1.0869	0.4187	8.9480	0.1641	0.0483

	Generalised Two-Parameter				Autostimulation				
Mouse	a	b	α	β	a	b	α	β	ω
1	1.0401	0.9436	0.9841	1.0444	0.5298	1.0856	0.9364	1.3635	0.6433
2	1.0700	0.9323	0.9883	1.0566	0.6075	1.6167	0.4092	1.3822	0.0379
3	1.1713	0.9416	0.9352	0.9450	0.5848	1.7332	0.3747	0.6348	0.0676
4	0.5502	0.3535	1.2474	1.4929	0.6632	1.2950	0.2191	0.5299	0.0518
6	0.1561	0.9108	0.9843	1.0772	1.6402	0.0342	1.5432	7.3076	0.7167
8	1.1432	0.8723	0.9299	0.9299	0.5022	1.3903	0.3839	0.9479	0.0572
9	1.8977	1.7020	0.7849	0.7849	0.0401	1.3093	0.6110	0.7280	0.2335
10	0.5172	0.3402	0.6931	0.6931	0.0631	1.9126	1.2556	0.4807	0.8334
17	0.8250	0.6660	1.3181	1.5887	1.1294	0.9155	1.1292	0.8455	1.0675
18	0.8873	0.7367	1.3081	1.5380	1.964	0.8661	1.1155	0.8879	1.0445

	Bertalanffy		
Mouse	a	b	α
1	1.0863	0.9924	1.0641
2	1.0591	0.9215	1.0749
3	1.1108	0.8786	0.9761
4	1.1381	0.9620	1.0061
6	1.0206	0.7732	1.1217
8	1.2886	1.183	1.0218
9	2.6124	2.4298	1.0259
10	1.7662	1.6005	1.0595
17	1.0776	0.9574	1.0486
18	0.5193	0.3981	1.1230

9.2 WARWICK-JUMBE

Here is the proof that $\boldsymbol{r_{diff}}^* \geq \boldsymbol{r}\left(\mathbf{1}-\left(\mathbf{1}-\frac{\boldsymbol{K}}{\boldsymbol{rk_{grow}}}\left(\frac{\mathbf{3}}{\mathbf{4\pi}}\right)^{\frac{1}{3}}\right)^{\frac{1}{3}}\right)$ and $\left.\frac{\boldsymbol{dE}}{\boldsymbol{dt}}\right|_{\boldsymbol{r_{diff}}=\boldsymbol{r_{diff}}^*} < 0$;

For generality, let $\frac{K}{k_g}\left(\frac{3}{4\pi}\right)^{\frac{1}{3}} = \theta$, an arbitrary constant. Consider the Taylor expansion of $\left(1-\theta\frac{1}{r}\right)^{\frac{1}{3}}$ within $r\left(1-\left(1-\theta\frac{1}{r}\right)^{\frac{1}{3}}\right)$;

$$r\left(1-\left(1-\theta\frac{1}{r}\right)^{\frac{1}{3}}\right) = r\left(1-\left(1-\theta r^{-1}\frac{1}{3}+\theta^2 r^{-2}\frac{1}{3}\left(\frac{-2}{3}\right)-\theta^3 r^{-3}\frac{1}{3}\left(\frac{-2}{3}\right)\left(\frac{-5}{3}\right)+\cdots\right)\right)$$

$$= r\left(\theta r^{-1}\frac{1}{3}+\theta^2 r^{-2}\frac{1}{3}\left(\frac{2}{3}\right)+\theta^3 r^{-3}\frac{1}{3}\left(\frac{2}{3}\right)\left(\frac{5}{3}\right)+\cdots\right)$$

$$= \theta\frac{1}{3} + \theta^2 r^{-1}\frac{1}{3}\left(\frac{2}{3}\right) + \theta^3 r^{-2}\frac{1}{3}\left(\frac{2}{3}\right)\left(\frac{5}{3}\right) + \cdots$$

$$= \theta\frac{1}{3} + \sum_{n=1}^{\infty}\theta^{(n+1)} r^{-n}\frac{1}{3}\prod_{i=1}^{n}\frac{(3n-1)}{3}$$

As $\frac{dV}{dt} > 0$ in the absence of drug effects, we have that r is increasing. So for $t > 0$, we have;

$$\theta\frac{1}{3} + \sum_{n=1}^{\infty}\theta^{(n+1)} r_0{}^{-n}\frac{1}{3}\prod_{i=1}^{n}\frac{(3n-1)}{3} > \theta\frac{1}{3} + \sum_{n=1}^{\infty}\theta^{(n+1)} r(t)^{-n}\frac{1}{3}\prod_{i=1}^{n}\frac{(3n-1)}{3}$$

Therefore

$$r_{diff}{}^{*} = \theta\frac{1}{3} + \sum_{n=1}^{\infty}\theta^{(n+1)} r_0{}^{-n}\frac{1}{3}\prod_{i=1}^{n}\frac{(3n-1)}{3} > \theta\frac{1}{3} + \sum_{n=1}^{\infty}\theta^{(n+1)} r(t)^{-n}\frac{1}{3}\prod_{i=1}^{n}\frac{(3n-1)}{3},$$

$$\Rightarrow r_d{}^{*} > r\left(1 - \left(1 - \frac{K}{rk_g}\left(\frac{3}{4\pi}\right)^{\frac{1}{3}}\right)^{\frac{1}{3}}\right).$$

Let $r^{*} = r(t)\left(1 - \left(1 - \frac{K}{r(t)k_g}\left(\frac{3}{4\pi}\right)^{\frac{1}{3}}\right)^{\frac{1}{3}}\right)$, so at time t > 0, $\left.\frac{dE}{dt}\right|_{r_d = r^{*}} = 0.$

$$\left(1 - \left(1 - \frac{r_{diff}{}^{*}}{r}\right)^3\right) > \left(1 - \left(1 - \frac{r^{*}}{r}\right)^3\right)$$

$$\Rightarrow KV_J{}^{\frac{2}{3}} - k_g V_W\left(1 - \left(1 - \frac{r_{diff}{}^{*}}{r}\right)^3\right) < KV_J{}^{\frac{2}{3}} - k_g V_W\left(1 - \left(1 - \frac{r^{*}}{r}\right)^3\right) = 0$$

$$\Rightarrow \left.\frac{dE}{dt}\right|_{r_d = r_d{}^{*}} = KV_J{}^{\frac{2}{3}} - k_g V_W\left(1 - \left(1 - \frac{r_d{}^{*}}{r}\right)^3\right) < 0$$

So for all t > 0, $\left.\frac{dE}{dt}\right|_{r_d = r_d{}^{*}} < 0$ ■

9.3 KINASE

9.3.1 Derivation of the non-linear model

Consider the following basic model a protein phosphorylation and dephosphorylation mediated by a kinase and phosphatase respectively. Here *pP* is the fraction of phosphorylated protein

$$\frac{dpP}{dt} = \frac{V_{max,1}(1-pP)}{K_{m1}+(1-pP)} - \frac{V_{max,2}\cdot pP}{K_{m2}+pP} \quad . \qquad \{1\}$$

In this case

$$pP(0) = pP_0 = \frac{\frac{V_{max,1}(1-pP_0)}{K_{m1}+(1-pP_0)}}{\frac{V_{max,2}}{K_{m2}+pP_0}} \quad \{2\}$$

to be the steady state of *pP* and so the partitioning:

$$K_{eq} = \frac{1-pP_0}{pP_0}. \quad \{3\}$$

Define the normalisation of *pP* to its steady state to be

$$\widehat{pP} = \frac{pP}{pP_0} \quad \{4\}$$

And so

$$\frac{1-pP}{1-pP_0} = \frac{1}{K_{eq}pP_0} - \frac{\widehat{pP}}{K_{eq}}. \quad \{5\}$$

If we use the following rewriting of the second term of {1}

$$\frac{V_{max,2}\cdot pP}{K_{m1}+pP} = \frac{V_{max,2}\cdot pP_0}{K_{m2}+pP_0}\frac{\frac{V_{max,2}\cdot pP}{K_{m2}+pP}}{\frac{V_{max,2}\cdot pP_0}{K_{m2}+pP_0}} = \frac{V_{max,2}\cdot pP_0}{K_{m2}+pP_0}\frac{pP}{pP_0}\frac{K_{m2}+pP_0}{K_{m2}+pP} = \frac{V_{max,2}\cdot pP_0}{K_{m2}+pP_0}\widehat{pP}\frac{K\prime_{m2}+1}{K\prime_{m2}+\widehat{pP}} \quad \{6\}$$

And similarly for the first term using {5}, then we can rewrite {1} as

$$\frac{dpP}{dt} = \frac{V_{max1}(1-pP_0)}{K_{m1}+(1-pP_0)}\left(\frac{1}{K_{eq}pP_0} - \frac{\widehat{pP}}{K_{eq}}\right)\frac{(K\prime_{m1}+1)}{\left(K\prime_{m1}+\frac{1}{K_{eq}pP_0}-\frac{\widehat{pP}}{K_{eq}}\right)} - \frac{V_{max2}\cdot pP_0}{K_{m2}+pP_0}\widehat{pP}\frac{K\prime_{m2}+1}{K\prime_{m2}+\widehat{pP}} \quad \{7\}$$

Where

$$K'_{m1} = \frac{K_{m1}}{pP_0}, K'_{m2} = \frac{K_{m2}}{pP_0} \quad \{8\}$$

are the Michaelis-Menten coefficients relative to the steady state.

If we define the time constant of the system around the steady state as

$$\frac{1}{\tau} = \frac{V_{max2}}{K_{m2}+pP_0}, \quad \{9\}$$

substituting for *pP* using {4} we can therefore write

$$\frac{d\widehat{pP}}{dt} = \frac{1}{\tau}\left[\left(\frac{1}{K_{eq}pP_0} - \frac{\widehat{pP}}{K_{eq}}\right)\frac{(K\prime_{m1}+1)}{\left(K\prime_{m1}+\frac{1}{K_{eq}pP_0}-\frac{\widehat{pP}}{K_{eq}}\right)} - \widehat{pP}\frac{K\prime_{m2}+1}{K\prime_{m2}+\widehat{pP}}\right] \quad \{10\}$$

where

$$\widehat{pP}(0) = 1 \qquad \{11\}$$