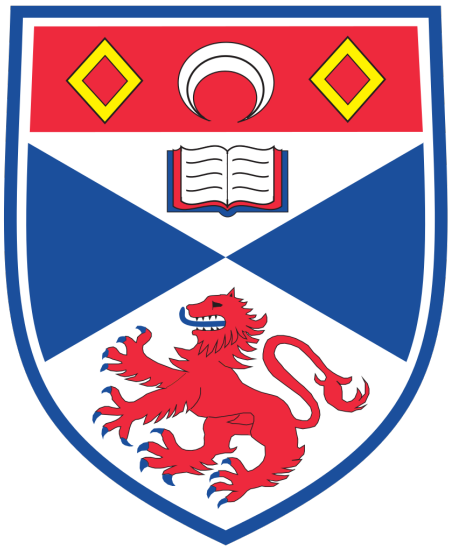
Qualitative and numerical study of some ordinary differential equations and integro-differential equations arising from the mathematical modelling of cancer cell populations

Ilaria Maresi*

130008434



University of St Andrews

23rd April 2017

*School of Mathematics and Statistics, University of St Andrews, North Haugh, St Andrews, Fife, KY16 9SS, United Kingdom

For Zia Marica, who never stops fighting.
A Zia Marica, che ci insegna cosa vuol dire lottare.

Acknowledgements

I would like to thank, first and foremost, Dr. Tommaso Lorenzi, who has had the utmost patience and kindness with me throughout this entire project. Grazie mille per tutto il tuo aiuto e il tuo incoraggiamento, questo progetto non sarebbe stato lo stesso senza di te.

It is a bittersweet feeling to be completing one of my last pieces of work for this University and this department. I can undoubtedly say, this place will always hold a special place in my heart. Thank you to all the professors, tutors and peers that have been there along the way and kept me going, you make this place what it is.

Declaration

I certify that this project report has been written by me, is a record of work carried out by me, and is essentially different from work undertaken for any other purpose or assessment.

Abstract

In order to comprehend the behaviour of solid tumours, it is essential to analyse the dynamic evolution of cancer cell populations and the driving ecological and evolutionary mechanisms. Of equal importance is an understanding of the phenotypic variants commonly found in cancer cell populations and how they evolve with time. Mathematical models are of key importance in cancer research as they create important predictions that would otherwise require lengthy experimentation.

The results in this paper are achieved via two central models. The first is an ordinary differential equation model, which describes the time evolution of the size of a well-mixed population of cancer cells. The model is considered both with and without the effect of a cytotoxic drug acting on the population. An integro-differential equation model is then introduced to account for and further study the evolution of phenotypic heterogeneity found in cancer cell populations.

Results illustrate that tumours tend to maximum population sizes in the absence of treatment. When cytotoxic treatment is incorporated, the ordinary differential equation model highlights the efficacy of bolus injections over continuous infusions. Furthermore, it is shown that in heterogeneous tumours one phenotypic trait will dominate over others in the long-term.

Introduction

Aim

This paper will explore the dynamic evolution of cancer cell populations through an ordinary differential equation (ODE) model and aim to understand temporal behaviour of phenotypic variants present in tumours via a system of integro-differential equations (IDEs). The ODE model will be considered for two cases: without the presence of treatment and under the effect of cytotoxic agents. For both ODE and IDE models, the analysis will be twofold comprising of theoretical analysis paired with numerical simulations.

Structure

The work is divided into four main sections:

- *Biological background*: This section provides a brief introduction to cancer as an evolutionary and ecological disease with a particular focus on tumour heterogeneity. Furthermore, it goes on to explain the importance of mathematical models as predictive tools in a biological realm. Lastly, treatment methods of interest – namely, bolus injections and continuous infusions – are introduced and outlined.

- *Ordinary differential equation model*: Here, the ODE model is introduced and is firstly analysed without the effect of cytotoxic treatment. This provides information on the dynamics of the cell population as a tumour grows in the absence of external agents. A term for cytotoxic treatment is then incorporated into the model, which allows for a comparison to be made between the efficacy of a bolus injection versus continuous infusions. Furthermore, the ODE can be used to elucidate the optimal time to inject a bolus treatment – a result that holds important biological implications for treatment optimisation. This section is largely based on the work of Benoît Perthame, see ref. 10.

- *Integro-differential equation model*: The IDE model allows for a more detailed look inside the dynamics of cancer cell populations by incorporating the presence of phenotypic heterogeneity in tumours. The model is based on the temporal evolution of the density of cells in specific phenotypic states. Finding solutions to this model allows us to understand long-term dynamic behaviour of cell populations with a wide variety of phenotypes and how certain phenotypes may dominate over others. This section does not include the effect of cytotoxic treatment on heterogeneous cancer cell populations, as this would involve technical difficulties that range beyond the scope of this project. This section is largely based on the work of Benoît Perthame, see ref. 12.

- *Conclusion*: The project concludes with some final remarks on the research conducted and the potential for future research as an extension of work carried out here. There are some suggestions for future studies that integrate the mathematical models in a more biological framework.

Results

Analytical and numerical simulations of the ODE model point to the same results: in the absence of cytotoxic treatment cancer cell populations will tend to a maximum tumour size, irrespective of the initial population size. Upon the inclusion of the treatment term in the model, the bolus injection

transpires as more efficacious than a continuous infusion, as it kills a greater number of cancer cells after injection. Moreover, the model discerns the best time to administer a bolus injection is at the end of the treatment time period. These results allow for robust biological conclusions under a broad structural stability, as parameter values can freely be changed to fit specific biological scenarios.

Results from the IDE model reinforce the statement that in the long-term cancer cell population sizes – even heterogeneous ones – tend to maximum values. Furthermore, it is proven both numerically and analytically, that the density of cells in specific phenotypic states concentrate as Dirac masses in the long-term. Biologically this means, in the limit of large times, cells of one specific phenotypic state are selected, whilst others die out.

Contents

1 Biological background

1.1 Cancer through the ages

Cancer, one of the most prevalent causes of death in the developed world, has affected organisms for millions of years. The discovery of metastatic cancer in a dinosaur bone allows us to trace the condition back as far as the Mesozoic era – see Fig. 1. Eons later, ancient Egyptians noted in hieroglyphs the presence of benign and malignant tumours and their respective treatments. Although, unlike the ancient Egyptians we no longer treat cancer with pig ears and other animal parts, we are yet to find a cure to this ancient disease[1].

Figure 1: Evidence of metastatic cancer found in dinosaur bone [11]

1.2 The power of mathematical models and the evolution of cancer

The existence of cancer can be attributed to somatic evolution. Somatic cells, that is to say any cells of an organism bar reproductive ones, will divide and die in a lifetime as part of a natural biological process. Over time, somatic cells may acquire alterations, which could lead to a population of premalignant cells. Fig. 2 is an illustration of these concepts of cell division, death and mutation. Even though other non-genetic and genetic alterations are possible, cancer cells are evolutionarily favoured over others because of their high proliferation rates and high chances of survival[3,4].

As a population of cancer cells progresses, new phenotypes arise; these phenotypic variants emerge as a consequence of changes in the cells' gene expressions that are passed on from parent cell to new cell, *i.e.* these are heritable changes. We can distinguish between two types of cancer cell populations: homogeneous and heterogeneous, as depicted in Fig. 3. The effect of the tumour microenvironment and anti-cancer treatments subject the phenotypes to natural selection, causing them survive, reproduce and die accordingly. Whilst populations of cancer cells can go through periods of time with a sole dominant phenotype common to all the population's cells, there are also periods of heterogeneity. Fig. 4 is a depiction of this evolution in homogeneous and heterogeneous neoplasms. The variation in phenotype from cell-to-cell in heterogeneous tumours is a major cause for therapeutic resistance and thus has an important role when considering the success of treatment

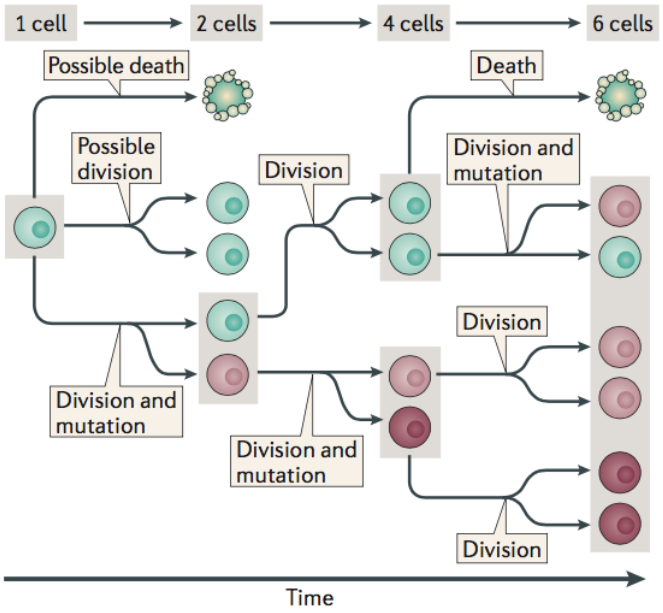


Figure 2: A visual representation of a cell's possible path to a malignant mutation. As the cell initially divides there are three outcomes: the first is cell death, the second is division without mutation, and the last is division and mutation. Mutated cells continue to divide, meanwhile there is a chance the non-mutated cells also develop mutations with time. These divisions and mutations could eventually lead to a population of premalignant cells[3].

therapies[7].

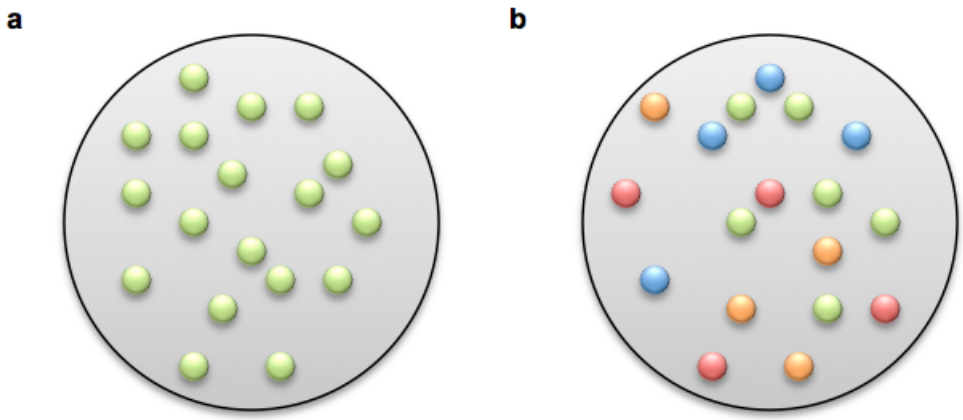


Figure 3: **a)** This is a population of homogeneous cancer cells, *i.e.* one phenotype dominates the population. **b)** Here is a heterogeneous population of cancer cells, where phenotypes vary from cell to cell. Various colours indicate different phenotypic variants.

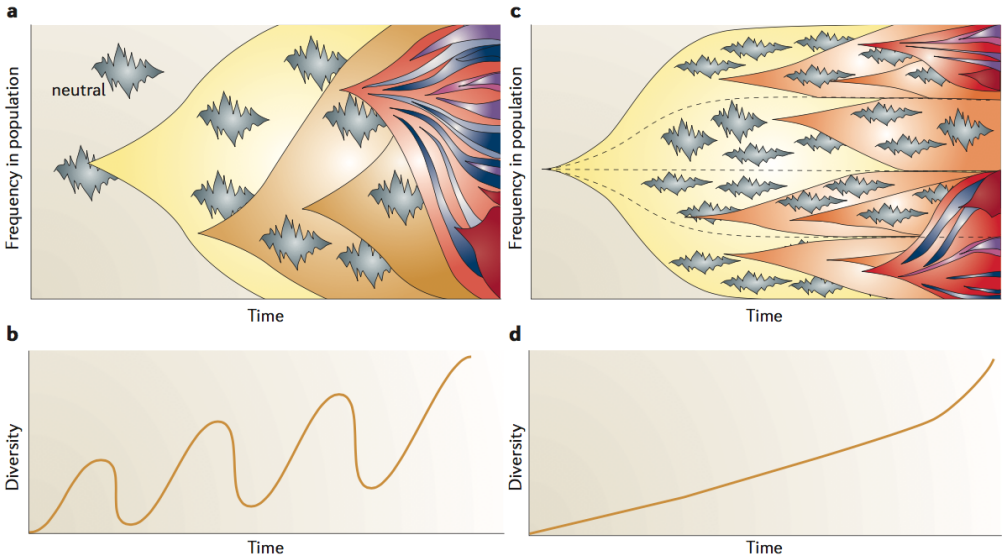


Figure 4: **a)** This is a population of homogeneous cells, *i.e.* the neoplasm is essentially a single population of cells with a common phenotype. Over time various phenotypes emerge and dominate the population **b)** Here we see the genetic diversity of the population fluctuating with time. The diversity will increase as new phenotypes emerge and decrease as these variations dominate the population. **c)** Here the neoplasm has various sub-populations with different phenotypes and hence is a heterogeneous population. Phenotypes will emerge and dominate different sub-populations rather than the entire population **d)** The genetic diversity has a much more clear upwards trend with no fluctuations since phenotypes do not dominate the majority of the population[8].

ODEs are extremely useful in outlining the time evolution of the size of a well-mixed population of cancer cells, but become unwieldy when describing phenotypic evolution in heterogeneous tumours. For this task, one would require a system of differential equations, one equation to describe the evolution of each phenotype. In tumours of high phenotypic heterogeneity, it seems impractical to work with such a large system. Thus, we introduce a secondary model that makes use of a new variable, x, which represents a cell's particular phenotypic state. In this model it is much more simple to work with different concentrations of various phenotypes for every tumour, as opposed to a system of ODEs. This model is an IDE model, which successfully lends itself to understanding phenotypic evolution in solid tumours[7].

Presently, a great deal of qualitative research from a biological standpoint has been carried out on various types of cancers and tumours and has answered many important questions. Yet queries still remain: What factors govern tumour growth? Why do certain tumour phenotypes thrive whilst others die out? These issues would call for many clinical trials that may not even necessarily replicate the in vivo behaviour of cancer. This is the power of mathematical modelling: prediction[2,3].

Mathematical models use existing data to create novel qualitative conclusions to better understand cancer from an evolutionary standpoint. These models can further act as tools to disprove

or affirm biological hypotheses, such as, is initial decay in cancer cell population a good predictive sign for future behaviour? The results of these models can later be applied to in vivo experiments[3].

Throughout this paper we have chosen to focus solely on proving and demonstrating, via numerical and analytical methods, the models' mathematical results. The applications of these models to biological situations have not been extensively explored in the interest of the project's limited duration. However, for the interested reader, analogous models and their applications can be found in various papers, such as *"The mathematics of cancer: integrating quantitative models"*[3] and *"Dissecting cancer through mathematics: from the cell to the animal model"*[2].

1.3 Cancer treatments of interest

The research conducted in this project will focus solely on the effect of treatment in the framework of the ODE model. It is possible to include the influence of treatment in the IDE model but this is rather complicated and includes technical difficulties that are beyond the scope of this project.

This paper will focus on specific treatments that will be restricted to those attained through cytotoxic drugs. Cytotoxic drugs target cells' abilities to replicate and grow. They have therefore proved effective in stopping tumour growth but inadvertently result in harming healthy cells. In treatment therapies cytotoxic drugs are administered a number of ways; in particular, we will distinguish two types: bolus injections versus continuous infusion.

Bolus injections are given to patients in short time intervals over a certain time period. For example, a bolus treatment could comprise of a daily dose of a cytotoxic drug given over a time span of 15 minutes administered once a week. Whereas, continuous infusions are, much like the name suggests, given to patients continuously over an often longer period of time, such as 48 hours. For a visual representation of methods of administration performed in a specific study[3], see Fig. 5. Both treatments hold advantages and disadvantages. As an example, bolus treatments are limited by cardiotoxicity, damage created to the muscles of the heart, as opposed to continuous infusions, which have been shown to have lower rates of cardiotoxicity[5]. However, in the section on ordinary differential equations in this paper we focus on the advantages of bolus as a treatment therapy and demonstrate it can be more effective than continuous treatment.

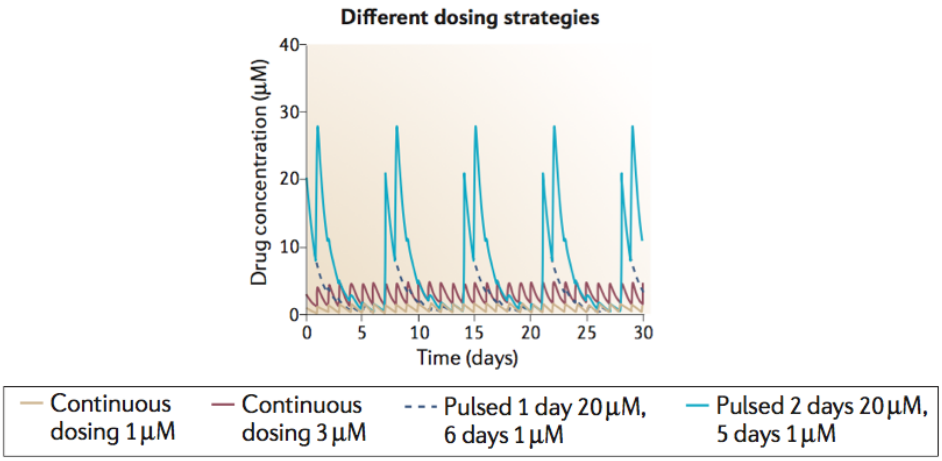


Figure 5: Various methods of cytotoxic drug administration. The pulsed strategies correspond to bolus injections whilst the continuous dosing is the aftorementioned continuous infusion. These concentrations and frequencies are based off treatment for the epidermal growth factor receptor inhibitor, erlotinib, in patients with lung cancer[3].

2 Ordinary differential equation model

In this section, we consider a mathematical model formulated in terms of an ordinary differential equation (ODE) which describes the time evolution of the size of a well-mixed population of cancer cells. Cells within the population proliferate and die due to intrapopulation competition. Moreover, a cytotoxic drug can be present, which we assume to act by increasing the death rate of cancer cells.

2.1 Model description

We model the number of cancer cells in the population at time $t \geq 0$ by means of the function $N(t) \geq 0$, whose evolution is governed by the following initial value problem

$$\begin{cases} \dfrac{dN}{dt}(t) = R\big(N(t)\big)\, N(t) - c(t)\, N(t), & t \in (0,\infty) \\ \\ N(t=0) = N^0 > 0. \end{cases} \tag{2.1}$$

In the equation $(4.1)_1$, the function $c(t) \geq 0$ stands for a suitably rescaled version of the dose of cytotoxic drug to which cancer cells are exposed to at time t. The function $R\big(N(t)\big)$ models the net proliferation rate of cancer cells (*i.e.,* the net growth rate of the cell population) at time t in the absence of the cytotoxic drug under the environmental conditions identified by the cell number $N(t)$. Throughout this work, we assume $c(t)$ to be given and we let the function R satisfy the following assumptions

$$R(0) = r > 0, \quad \frac{\partial R}{\partial N}(\cdot) < 0, \quad R(K) = 0 \quad \text{and} \quad -\infty < R(\cdot) < \infty. \tag{2.2}$$

The parameter $r \in \mathbb{R}_+$ stands for the maximal cell proliferation rate in the presence of large concentrations of nutrients (*e.g.*, oxygen and glucose) and the parameter $K \in \mathbb{R}_+$ models the maximal size

(*i.e.*, the carrying capacity) of the cell population. Assumption $(2.2)_2$ translates into mathematical terms the observation that higher cell numbers lead to more intense intrapopulation competition for nutrients, which corresponds to a smaller net growth rate of the cell population.

2.2 Dynamics of the cell population without cytotoxic drug

In this section we study the dynamics of the cell population in the absence of a cytotoxic drug, *i.e.*, we assume

$$c(\cdot) = 0. \tag{2.3}$$

In this case, a characterisation of the time evolution of the number of cancer cells in the population is provided by the following proposition:

Proposition 2.1 *Under assumptions* (2.2) *and* (2.3) *the solutions to the initial value problem* (4.1) *are such that*

$$\operatorname{sgn}\left(\frac{dN}{dt}(t)\right) = \operatorname{sgn}\big(R\big(N^0\big)\big) \quad \textit{for all} \;\; t \in \mathbb{R}_+^* \tag{2.4}$$

and

$$N(t) \longrightarrow K \quad \textit{as} \;\; t \to \infty. \tag{2.5}$$

Proof. We divide the proof of Proposition 2.1 into two parts.

Part 1: proof of (2.4)**.** Firstly, we introduce the notation $u(t) = \dfrac{dN}{dt}(t)$ and compute

$$\frac{du}{dt} = \frac{dN}{dt}\, R + N\, \frac{dR}{dN}\frac{dN}{dt} = \frac{dN}{dt}\left[R + N\, \frac{dR}{dN}\right]. \tag{2.6}$$

Secondly, we define

$$g(t) = R(N) + N(t)\, \frac{dR}{dN}(N)$$

and solve equation (2.6) to obtain

$$\frac{du}{dt}(t) = u(t)g(t) \Longrightarrow u(t) = u_0\, e^{\int_0^t g(s)ds},$$

that is,

$$\frac{d}{dt}\Big(\frac{dN}{dt}\Big) = R\big(N_0\big)\, N^0\, e^{\int_0^t g(s)ds}. \tag{2.7}$$

Due the positivity of the exponential and the fact that $N^0 > 0$, equation (2.7) allows us to conclude that

$$\operatorname{sgn}\left(\frac{dN}{dt}(t)\right) = \operatorname{sgn}\big(R\big(N^0\big)\big) \quad \text{for all} \;\; t \in \mathbb{R}_+^*.$$

Part 2: proof of (2.5)**.** Introducing the following notation for the right-hand side of equation $(4.1)_1$

$$F(N) = R(N)\, N$$

and solving the equation $F(N^*) = 0$ for N^* we find that, under assumption $(2.2)_3$, the equation $(4.1)_1$ admits two steady-states, that is,

$$N_1^* = 0 \quad \text{and} \quad N_2^* = K.$$

Differentiating $F(N)$ with respect to N yields

$$\frac{dF}{dN}(N) = R(N) + N\,\frac{dR}{dN}(N),$$

from which, using assumptions $(2.2)_1$, we achieve

$$\frac{dF}{dN}(N_1^*) = R(N_1^*) + N_1^*\,\frac{dR}{dN}(N_1^*) = R(0) = r > 0$$

and

$$\frac{dF}{dN}(N_2^*) = R(N_2^*) + N_2^*\,\frac{dR}{dN}(N_2^*) = K\,\frac{dR}{dN}(K) < 0.$$

Therefore, the steady state N_1^* is unstable whereas the steady state N_2^* is asymptotically stable. This allows us to conclude that the result given by (2.5) holds.

□

The analytical results established by Proposition 2.1 are illustrated by the numerical results presented in Figure 6. These results have been obtained by constructing numerical solutions of the initial value problem (4.1) for $t \in [0, 50]$,

$$c(t) = 0 \quad \forall\, t \in [0, 50] \tag{2.8}$$

and

$$R\big(N\big) := r\Big(1 - \frac{N}{K}\Big) \quad \text{with} \quad r = 0.5 \quad \text{and} \quad K = 10^8. \tag{2.9}$$

Note that the definition given by equation (2.9) satisfies assumptions (2.2). We consider both the case in which $N^0 = 10^6$, so that $R\big(N^0\big) > 0$, and the case where $N^0 = 2 \times 10^8$ so that $R\big(N^0\big) < 0$. Further details of numerical simulations are provided in Appendix 1.

In agreement with the results of Proposition 2.1, the condition $\operatorname{sgn}\big(\frac{dN}{dt}(t)\big) = \operatorname{sgn}\big(R\big(N^0\big)\big)$ is met for all time instants t and $N(t) \to K$ as $t \to \infty$.

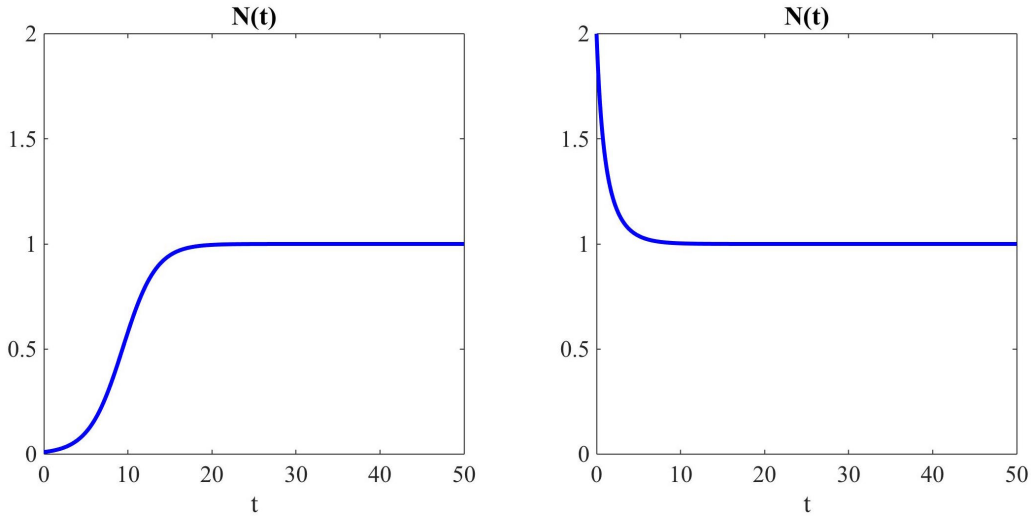


Figure 6: Sample dynamics of the number of cells $N(t)$ in the absence of the cytotoxic drug for the net proliferation rate $R(N)$ given by equations (2.9). The left panel refers to the case in which $R\big(N^0\big) > 0$, while the right panel corresponds to the case where $R\big(N^0\big) < 0$. The values of N are in units of K.

2.3 Cell dynamics under the action of the cytotoxic drug

We let the cancer cell population be exposed to the cytotoxic drug over the time interval $[0, t_f]$. We introduce an additional parameter $M > 0$ to model the total dose of cytotoxic drug delivered over this interval of time, that is, we assume

$$\int_0^{t_f} c(t)\, dt = M. \tag{2.10}$$

Focussing on a biological scenario in which

$$N^0 < K \quad \text{so that} \quad R(\cdot) \geq 0, \tag{2.11}$$

we compare the efficacy between two widely used administration protocols: continuous infusion and bolus injection. In particular, we consider the case where

$$c(t) = c_c(t) := \frac{M}{t_f} \quad \text{for all} \quad t \in [0, t_f] \qquad \text{(continuous infusion)} \tag{2.12}$$

and the case in which

$$c(t) = c_b(t) := M\, \delta(t - \tau) \quad \text{with} \quad \tau \in [0, t_f] \quad \text{(bolus injection)}, \tag{2.13}$$

where $\delta(\cdot)$ stands for the Dirac delta distribution.

In order to study the time evolution of the number of cancer cells under the bolus injection, first we have to find an equivalent smooth version of the definition given by (2.13). In fact, in this case the function $N(t)$ can be discontinuous and it makes no sense multiplying distributions by discontinuous functions. Therefore, we introduce the following smooth approximation of definition (2.13)

$$c_b(t) \approx \lim_{\varepsilon \to 0} \frac{M}{\varepsilon} \mathbb{1}_{\{\tau - \varepsilon < t < \tau\}}, \tag{2.14}$$

where $\mathbb{1}$ denotes the indicator function, and we show how the initial value problem (4.1) reads when $c(t) = c_b(t)$ and $c_b(t)$ is given by approximation (2.14):

Lemma 2.2 *Under assumptions* (2.2) *and* (2.11)*, if* $c(t) = c_b(t)$ *and* $c_b(t)$ *is given by approximation* (2.14) *then the initial value problem* (4.1) *reads as*

$$\begin{cases} \dfrac{dN}{dt}(t) = R\big(N(t)\big)\, N(t), \\ \\ N(t=0) = N^0 > 0, \\ \\ N(\tau^+) = N(\tau^-)e^{-M}. \end{cases} \tag{2.15}$$

Proof. We introduce the notations

$$U(t) = \ln N(t) \quad \text{and} \quad G(U) = R\big(N\big).$$

Since

$$\frac{dU}{dt}(t) = G(U)(t) - c_b(t),$$

when $c_b(t)$ is given by approximation (2.14) we have

$$\frac{dU}{dt}(t) = G(U)(t) - \frac{M}{\varepsilon} \quad \text{for } \tau - \varepsilon < t < \tau.$$

Integrating the above differential equation over the interval $(\tau - \varepsilon, \tau)$ we find

$$U(\tau) = U(\tau - \varepsilon) + \int_{\tau-\varepsilon}^{\tau} G(U)(t)\, dt - M. \tag{2.16}$$

Since $\operatorname{sgn}\Big(\dfrac{dG}{dU}\Big) = \operatorname{sgn}\Big(\dfrac{dR}{dN}\Big)$, assumptions $(2.2)_4$ and (2.11) guarantee that

$$0 < \int_{\tau-\varepsilon}^{\tau} G(U)(t)\, dt < \varepsilon \sup G(U) < \infty.$$

Hence, estimating equation (2.16) from below and from above we find

$$U(\tau - \varepsilon) - M \leq U(\tau) \leq U(\tau - \varepsilon) + \varepsilon \sup G(U) - M.$$

Therefore, in the limit $\varepsilon \to 0$ we achieve

$$U(\tau^-) - M \leq U(\tau^+) \leq U(\tau^-) - M,$$

that is,

$$U(\tau^+) = U(\tau^-) - M \quad \Longrightarrow \quad N(\tau^+) = N(\tau^-)e^{-M}. \tag{2.17}$$

This concludes the proof of Lemma 2.2.

□

A first natural question when using the bolus injection is: what is the injection time that maximises the cytotoxic effect of the drug (*i.e.*, the injection time that allows one to kill the highest possible number of cancer cells)? In the mathematical framework of problem (2.15), addressing this question corresponds to finding the injection time $\tau^* \in [0, t_f]$ such that

$$N(\tau^{*-}) = \max_{\tau \in [0,t_f]} N(\tau^-).$$

This question is answered by the following proposition:

Proposition 2.3 *Under assumptions* (2.2) *and* (2.11)*, if* $c(t) = c_b(t)$ *and* $c_b(t)$ *is given by approximation* (2.14) *then* $\tau^* = t_f$*.*

Proof. Assumptions (2.2) and (2.11) imply that [see Proposition 2.1] the solutions to the problem (2.15) will be monotonically increasing over the time interval $[0, \tau)$. Therefore, $\max\limits_{\tau\in[0,t_f]} N(\tau^-) = N(t_f^-)$.

□

Biologically, the result established by Proposition 2.3 indicates that the time instant t_f is the injection time that maximises the cytotoxic effect of the drug.

Now we compare the efficacy between continuous infusion over the time interval $[0, t_f]$ and bolus injection at time t_f. We assess the efficacy of the two protocols in terms of the number of cells that are still alive after therapy. In the mathematical framework under consideration, this can be done by considering the following problems – where $t \in (0, \infty)$ and $c_c(t)$ are given by equation (2.12) –

$$\begin{cases} \dfrac{dN_c}{dt}(t) = R\big(N_c(t)\big)\, N_c(t) - c_c(t)\, N_c(t), \\ \\ N_c(t=0) = N^0 > 0 \end{cases} \qquad \text{and} \qquad \begin{cases} \dfrac{dN_b}{dt}(t) = R\big(N_b(t)\big)\, N_b(t), \\ \\ N_b(t=0) = N^0 > 0, \\ \\ N_b(t_f^+) = N_b(t_f^-)e^{-M} \end{cases} \tag{2.18}$$

and comparing the value of $N_c(t_f)$ with the value of $N_b(t_f^+)$. The result established by the following proposition demonstrates that a bolus injection at time t_f is more effective than the continuous infusion over the time interval $[0, t_f]$:

Proposition 2.4 *Under assumptions* (2.2) *and* (2.11)*, the solutions to the problems* (2.18) *are such that*

$$N_c(t_f) > N_b(t_f^+). \tag{2.19}$$

Proof. Substituting

$$U_{b,c}(t) = \ln N_{b,c}(t) \quad \text{and} \quad G(U_{b,c}(t)) = R\big(N_{b,c}(t)\big)$$

and the definition given by (2.12) into equations (2.18) we find

$$\begin{cases} \dfrac{dU_c}{dt}(t) = G(U_c(t)) - \dfrac{M}{t_f}, \\ \\ U_c(t=0) = U^0 = \ln\big(N^0\big) \end{cases} \qquad \text{and} \qquad \begin{cases} \dfrac{dU_b}{dt}(t) = G(U_b(t)), \\ \\ U_b(t=0) = U^0, \\ \\ U_b(t_f^+) = U_b(t_f^-) - M. \end{cases}$$

Firstly, we note that

$$U_c(t) < U_b(t) \quad \text{for} \quad t \in (0, t_f). \tag{2.20}$$

In fact, assumptions (2.2) and (2.11) guarantee that there exists a positive real constant A such that

$$\frac{d}{dt}\big[U_c(t) - U_b(t)\big] = G(U_c(t)) - G(U_b(t)) - \frac{M}{t_f} \leq A\,\big[U_c(t) - U_b(t)\big].$$

The above differential inequality implies that

$$\frac{d}{dt}\big[U_c(t) - U_b(t)\big]_+^2 \leq 2\,A\big[U_c(t) - U_b(t)\big]_+^2,$$

where $\big[U_c(t) - U_b(t)\big]_+$ denotes the positive part of the function $U_c(t) - U_b(t)$. Therefore, since $\big[U_c(0) - U_b(0)\big]_+^2 = 0$, we can conclude that $\big[U_c(t) - U_b(t)\big]_+^2 = 0$ for all $t \in [0, t_f)$.

Since

$$\operatorname{sgn}\Big(\frac{dG}{dU}\Big) = \operatorname{sgn}\Big(\frac{dR}{dN}\Big) \leq 0,$$

the result given by (2.20) implies that

$$G(U_c(t)) - G(U_b(t)) \geq 0 \;\text{ for all }\; t \in [0, t_f).$$

Hence,

$$\frac{d}{dt}\big[U_c(t) - U_b(t)\big] = G(U_c(t)) - G(U_b(t)) - \frac{M}{t_f} \geq -\frac{M}{t_f} \;\text{ for all }\; t \in [0, t_f).$$

Integrating the above differential inequality over the interval $[0, t_f]$ we find

$$U_c(t_f) - U_b(t_f^-) \geq -M \quad \Longrightarrow \quad U_c(t_f) \geq U_b(t_f^-) - M.$$

Using this inequality together with the result given by (2.17) we achieve

$$U_c(t_f) - U_b(t_f^+) = U_c(t_f) - U_b(t_f^+) + U_b(t_f^-) - U_b(t_f^-) \geq U_b(t_f^-) - M + M - U_b(t_f^-) = 0,$$

that is,

$$U_c(t_f) \geq U_b(t_f^+) \quad \Longrightarrow \quad e^{U_c(t_f)} \geq e^{U_b(t_f^+)} \Longrightarrow \quad N_c(t_f) \geq N_b(t_f^+).$$

This concludes the proof of Proposition (2.4).

□

3 Integro-differential equation model

ODE models like that studied in the previous section can support a better understanding of the ecological and evolutionary mechanisms that drive the dynamics of cancer cell populations. However, such models do not account for the high degree of phenotypic heterogeneity which is experimentally observed in cancer cell populations. This can be done by introducing an additional independent variable $x \in \mathbb{R}^{d \geq 1}$, which characterises the cellular phenotypic state, and a new dependent variable $n(t,x)$, which stands for the density of cells in the phenotypic state x at time $t \geq 0$ within the population. For the sake of simplicity, here we will neglect the effects of phenotypic modifications and study the dynamics of cancer cells in the presence of proliferation and competition phenomena only in the absence of cytotoxic drugs. In this case, the evolution of the population density function $n(t,x)$ can be described by means of integro-differential equations (IDEs) like that considered in this section, which represents a natural extension of the ODE model analysed in the previous section.

3.1 Model description

We focus on the case in which the evolution of the population density function $n(t,x)$ is governed by the following initial value problem

$$\begin{cases} \dfrac{\partial n}{\partial t}(t,x) = R\big(x, I(t)\big)\, n(t,x), & (t,x) \in (0,\infty) \times \mathbb{R}^d \\ \\ I(t) = \displaystyle\int_{\mathbb{R}^d} n(t,x)\, dx \\ \\ n(t=0) = n^0(x) \in L^1 \cap L^\infty(\mathbb{R}^d), \qquad n^0(x) \geq 0 \ \text{ for a.e. } x \in \mathbb{R}^d. \end{cases} \tag{3.1}$$

In the equation $(4.4)_1$, the function $R\big(x, I(t)\big)$ models the net proliferation rate of cancer cells in the phenotypic state x at time t under the environmental conditions identified by the total cell number $I(t)$. The first condition given by equation $(4.4)_3$ means that the number of cells within the population at time $t = 0$ is finite, while the second condition given by equation $(4.4)_3$ translates into mathematical terms the idea that, due to phenotypic heterogeneity, there are some cancer cells in every phenotypic state at time $t = 0$.

Throughout this section we will make use of the following assumptions:

- the function $R\big(x, I(t)\big)$ is such that

$$R\big(x, I(t)\big) \in C\big(\mathbb{R}^d \times \mathbb{R}^*_+\big); \tag{3.2}$$

- there exist some real constants $0 < I_m < I_M < \infty$ and $0 < \overline{R} < \infty$ such that

$$\min_{x \in \mathbb{R}^d} R(x, I_m) = 0 \quad \text{and} \quad \max_{x \in \mathbb{R}^d} R(x, I_M) = 0 \tag{3.3}$$

and

$$\max_{x \in \mathbb{R}^d} R(x, I) \leq \overline{R} \quad \forall I \in [I_m, I_M]; \tag{3.4}$$

- there exist some real constants $0 < K_- < K_+ < \infty$ such that

$$-K_+ \leq \frac{\partial R}{\partial I}(x, I) < -K_- < 0 \quad \forall x \in \mathbb{R}^d \ \text{ and } \ \forall I \in [I_m, I_M]. \tag{3.5}$$

Moreover, we will assume that the initial condition $n^0(x)$ satisfies the following additional assumption

$$I_m \leq \int_{\mathbb{R}^d} n^0(x)dx \leq I_M. \tag{3.6}$$

Assumption $(3.3)_1$ means that if the population size is equal to I_m then the net cellular proliferation rate is either positive or zero for all phenotypic states. On the other hand, assumption $(3.3)_2$ frames mathematically the idea that the net cell proliferation rate is either negative or zero for every phenotypic state in the case when the population size is equal to I_M. Furthermore, assumption (3.4) means that the net proliferation rate is bounded for all phenotypic state and all values of the populations size between I_m and I_M. Finally, assumptions (3.5) translate into mathematical terms the observation that larger total numbers of cells lead to more intense intrapopulation competition for nutrients and thus cause a decrease in the proliferation rate for all phenotypic states. Taken together, assumptions (3.3)-(3.5) model a biological scenario in which if the population size I is larger than I_M then the environmental conditions are such that the net proliferation rate is negative for all phenotypic states, whereas if $I < I_m$ then there will be enough free resources for cells in all phenotypic states to proliferate.

3.2 Well-posedness of the initial value problem (4.4)

We begin our analysis by recalling the results established by the following theorem, which guarantees that the initial value problem (4.4) is well-posed in the sense of Hadamard (*i.e.*, the solution exists, it is unique and depends continuously on the initial data):

Theorem 3.1 *Under assumptions* (3.2)-(3.4) *and* (3.6) *there exists a unique non-negative solution* $n \in C\big(\mathbb{R}_+; L^1(\mathbb{R}^d)\big)$ *to problem* (4.4)*, and it satisfies the following a priori estimates*

$$I_m \leq I(t) \leq I_M \qquad \forall\, t \geq 0 \tag{3.7}$$

and

$$\int_0^T \int_{\mathbb{R}^d} R\big(x, I(t)\big)_- \, n(t,x)dx\, dt \leq \overline{R}\, I_M\, T + \int_{\mathbb{R}^d} n^0(x)dx \qquad \forall\, T > 0. \tag{3.8}$$

Here we provide a proof for the *a priori* estimates (3.7) and (3.8) only.

Proof. Integrating equation $(4.4)_1$ over $\mathbb{R}^d$ and estimating from above we obtain

$$\frac{d}{dt}\int_{\mathbb{R}^d} n(t,x)dx = \int_{\mathbb{R}^d} n(t,x)\, R\big(x, I(t)\big)dx \leq \int_{\mathbb{R}^d} n(t,x)\; dx\; \max_{x\in\mathbb{R}^d} R\big(x, I(t)\big),$$

that is

$$\frac{dI}{dt}(t) \leq I(t)\, \max_{x\in\mathbb{R}^d} R\big(x, I(t)\big),$$

Since $\max\limits_{x\in\mathbb{R}^d} R(x, I_M) = 0$ [*vid.* assumption $(3.3)_2$], the above differential inequality allows us to conclude that if $I(t) = I_M$ then $\dfrac{dI}{dt}(t) \leq 0$. This result together with assumption (3.6) yields $I(t) \leq I_M$ for all $t \geq 0$. Similarly, integrating equation $(4.4)_1$ over $\mathbb{R}^d$ and estimating from below we find

$$\frac{dI(t)}{dt} \geq \int_{\mathbb{R}^d} n(t,x)\; dx\; \min_{x\in\mathbb{R}^d}\, R\big(x, I(t)\big),$$

which means that

$$\frac{dI}{dt}(t) \geq I(t) \min_{x\in\mathbb{R}^d} R\big(x, I(t)\big).$$

Since $\min_{x\in\mathbb{R}^d} R(x, I_m) = 0$ [*vid.* assumption $(3.3)_1$], the above differential inequality allows us to conclude that if $I(t) = I_m$ then $\dfrac{dI}{dt}(t) \geq 0$. This result together with assumption (3.6) yields $I(t) \geq I_m$ for all $t \geq 0$. Taken together these results conclude the proof of the results given by equation (3.7).

On the other hand, integrating equation $(4.4)_1$ over $[0, T] \times \mathbb{R}^d$ we obtain

$$\begin{aligned}\int_{\mathbb{R}^d} \big[n(T,x) - n^0(x)\big]dx &= \int_0^T \int_{\mathbb{R}^d} R\big(x, I(t)\big)\, n(t,x)dx\, dt \\ &= \int_0^T \int_{\mathbb{R}^d} R\big(x, I(t)\big)_+ n(t,x)dx\, dt - \int_0^T \int_{\mathbb{R}^d} R\big(x, I(t)\big)_- n(t,x)dx\, dt\end{aligned}$$

Under assumption (3.4), using this result in combination with the upper bound for $I(t)$ given by equation (3.7) we achieve

$$\int_{\mathbb{R}^d} \big[n(T,x) - n^0(x)\big]dx \leq \overline{R}\, I_M\, T - \int_0^T \int_{\mathbb{R}^d} R\big(x, I(t)\big)_- n(t,x)dx\, dt.$$

This implies that the following relation hold

$$\int_0^T \int_{\mathbb{R}^d} R\big(x, I(t)\big)_- n(t,x)dx\, dt \leq \overline{R}\, I_M\, T + \int_{\mathbb{R}^d} n^0(x)dx - \int_{\mathbb{R}^d} n(T,x)dx$$

from which, using the lower bound for $I(t)$ given by equation (3.7), we can conclude that

$$\int_0^T \int_{\mathbb{R}^d} R\big(x, I(t)\big)_- n(t,x)dx\, dt \leq \overline{R}\, I_M\, T + \int_{\mathbb{R}^d} n^0(x)dx,$$

i.e., the result given by (3.8) holds.

□

3.3 Long-term behaviour of the cancer cell population

In this section we will prove that assumptions (3.2)-(3.6) are sufficient to guarantee that the solutions to the initial value problem (4.4) concentrate as Dirac masses in the limit of large times. The concentration points (*i.e.*, the centres of the Dirac masses) represent the fittest phenotypic traits within the population. In this sense, such an asymptotic result represents a possible mathematical way of identifying the phenotypic characteristics that provide cancer cells within the population with the strongest competitive advantage and, therefore, it has the potential to clarify the ecological and evolutionary mechanisms that allow cancer to develop and progresses.

Theorem 3.2 *Under assumptions* (3.2)-(3.6) *the solution to the initial value problem* (4.4) *is such that*

$$I(t) \xrightarrow[t\to\infty]{} I_M. \tag{3.9}$$

Furthermore, if there is a unique point $\overline{x} \in \mathbb{R}^d$ *such that*

$$R(\overline{x}, I_M) = \max_{x \in \mathbb{R}^d} R(x, I_M) = 0, \tag{3.10}$$

the initial cell population density $n^0(x)$ *is such that*

$$n^0 \in C(\mathbb{R}^d) \quad \textit{and} \quad n^0(\overline{x}) > 0 \tag{3.11}$$

and

$$\sup_{|x| \geq A} R(x, I_M) < 0 \quad \textit{for some} \;\; A \in \mathbb{R}_+, \tag{3.12}$$

then

$$n(t, x) \xrightarrow[t \to \infty]{} I_M \delta(x - \overline{x}). \tag{3.13}$$

Biologically, Theorem (3.2) means that, in the limit of large times, the total number of cells converges to the value I_M and cells in the phenotypic state characterised by the maximal fitness, *i.e.*, the phenotypic state $\overline{x}$, are selected.

The proof of Theorem (3.2) relies on the results established by the following lemma:

Lemma 3.3 *Under assumptions* (3.2)-(3.5) *the solution to the initial value problem* (4.4) *satisfies the following bounds*

$$\frac{dI(t)}{dt} \;\geq\; -e^{-K_- I_m t} \bigg(\int n^0(x) R(x, I^0) dx \bigg)_- \tag{3.14}$$

$$\int_0^\infty \left| \frac{dI(t)}{dt} \right| dt \;\leq\; I_M + \frac{2}{K_- I_m} \bigg(\int n^0(x) R(x, I^0) dx \bigg)_- \tag{3.15}$$

Proof. We begin by integrating equation $(4.4)_1$ over $\mathbb{R}^d$ to obtain

$$\frac{dI}{dt}(t) = J(t) \quad \text{with} \quad J(t) = \int_{\mathbb{R}^d} n(t, x) \; R\big(x, I(t)\big) dx. \tag{3.16}$$

Differentiating $J(t)$ with respect to t and making use of both the product rule and the chain rule yields

$$\frac{dJ}{dt}(t) = \int_{\mathbb{R}^d} n(t, x) \; R\big(x, I(t)\big)^2 dx + \int_{\mathbb{R}^d} n(t, x) \frac{\partial R}{\partial I}\big(x, I(t)\big) dx \, \frac{dI}{dt}(t), \tag{3.17}$$

that is,

$$\frac{dJ}{dt}(t) = \int_{\mathbb{R}^d} n(t, x) \; R\big(x, I(t)\big)^2 dx + \int_{\mathbb{R}^d} n(t, x) \frac{\partial R}{\partial I}\big(x, I(t)\big) dx \, J(t). \tag{3.18}$$

Since [*cf.* assumption (3.5)]

$$\frac{\partial R}{\partial I}(x, I) < -K_-$$

we have

$$\int_{\mathbb{R}^d} n(t, x) \frac{\partial R}{\partial I}\big(x, I(t)\big) dx \leq -K_- \int_{\mathbb{R}^d} n(t, x) dx.$$

Furthermore, since $I(t) \geq I_m > 0$ for all $t \geq 0$ we can conclude that

$$\int_{\mathbb{R}^d} n(t,x)\frac{\partial R}{\partial I}\big(x, I(t)\big)dx \leq -K_- \int_{\mathbb{R}^d} n(t,x)dx \leq -K_- I_m < 0.$$

Therefore, multiplying both sides of equation (3.18) by $-\big(\mathrm{sgn}(J)\big)_-$, which is a nonpositive quantity, and estimating from above we obtain the following differential inequality

$$\frac{d}{dt}\big(J(t)\big)_- \leq -K_-\, I_m \,\big(J(t)\big)_- \quad \Longrightarrow \quad \big(J(t)\big)_- \leq \big(J(0)\big)_-\, e^{-K_-\, I_m\, t}. \tag{3.19}$$

Using the fact that $J(t) = \dfrac{dI}{dt}(t)$ and $J(0) = \displaystyle\int_{\mathbb{R}^d} n^0(x)R(x, I^0)dx$ we can then conclude that

$$\frac{dI(t)}{dt} \geq -e^{-K_-\, I_m\, t}\bigg(\int_{\mathbb{R}^d} n^0(x)R(x, I(0))dx\bigg)_-,$$

that is, the result given by (3.14) holds.

To prove (3.15) we begin by noting that the result given by equation (3.19) implies that

$$\int_0^\infty \big(J(t)\big)_- dt \leq \frac{\big(J(0)\big)_-}{K_-\, I_M} = \frac{1}{K_-\, I_M}\bigg(\int_{\mathbb{R}^d} n^0(x)R(x, I(0))dx\bigg)_-. \tag{3.20}$$

Moreover, since $J(t) = \dfrac{dI}{dt}(t)$ and $0 < I_m \leq I(t) \leq I_M$ for all $t \geq 0$, we have

$$\int_0^\infty J(t)dt \leq I_M$$

from which, using the fact that $J(t) = \big(J(t)\big)_+ - \big(J(t)\big)_-$, we obtain

$$\int_0^\infty \big(J(t)\big)_+ dt \leq I_M + \int_0^\infty \big(J(t)\big)_- dt.$$

Adding $\big(J(t)\big)_-$ to both sides of the above inequality and using the fact that $\big|J(t)\big| = \big(J(t)\big)_+ + \big(J(t)\big)_-$ we find

$$\int_0^\infty \big|\big(J(t)\big)\big|dt \leq I_M + 2\int_0^\infty \big(J(t)\big)_- dt.$$

Finally, using the result given by (3.20) to estimate from above the second term on the righthand side of the above inequality we achieve

$$\int_0^\infty \big|\big(J(t)\big)\big|dt \leq I_M + \frac{2}{K_-\, I_M}\bigg(\int_{\mathbb{R}^d} n^0(x)R(x, I(0))dx\bigg)_-,$$

that is,

$$\int_0^\infty \left|\frac{dI(t)}{dt}\right| dt \leq I_M + \frac{2}{K_-\, I_M}\bigg(\int_{\mathbb{R}^d} n^0(x)R(x, I(0))dx\bigg)_-.$$

This concludes the proof of Lemma (3.3).

□

We are now in the position to prove Theorem (3.2):

Proof. We divide the proof into five steps.

1st step: *Existence of a limit for $I(t)$ as $t \to \infty$.* The BV bound given by equation (3.15) guarantees that there exists $\overline{I} \in \mathbb{R}_+$ such that

$$I(t) \xrightarrow[t\to\infty]{} \overline{I}.$$

2nd step: *Characterisation of the limit $\overline{I}$.* Suppose by contradiction that

$$\max_{x\in\mathbb{R}^d} R(x,\overline{I}) := 2M > 0. \tag{3.21}$$

Since R is a continuous function of x and I, if the above condition holds we can conclude that, choosing $\varepsilon > 0$ small enough, for every $\hat{x} \in \arg\max_{x\in\mathbb{R}^d} R(x,\overline{I})$ there exists a ball $B(\hat{x},\varepsilon)$ such that

$$R(x,\overline{I}-\varepsilon) \geq M \quad \forall x \in B(\hat{x},\varepsilon).$$

Moreover, from the first step we can conclude that there exists T_ε such that $I(t) \geq \overline{I} - \varepsilon$ for $t \geq T_\varepsilon$. Hence, integrating equation $(4.4)_1$ over the interval $[t, T_\epsilon]$ we find

$$n(t,x) = n(T_\varepsilon,x)\exp\bigg(\int_{T_\varepsilon}^{t} R\big(x,I(s)\big)ds\bigg) \geq n(T_\varepsilon,x)\exp\big((t-T_\varepsilon)M\big) \quad \forall x \in B(\hat{x},\varepsilon). \tag{3.22}$$

This implies that $n(t,x) \to \infty$ on $B(\hat{x},\varepsilon)$ as $t \to \infty$, which in turn implies that $I(t)$ tends to infinity. This contradicts the upper bound given by equation (3.7) and allows one to conclude that $\overline{I} = I_M$.

3rd step: *Existence of a limit for $n(t,x)$ as $t \to \infty$.* Thanks to the upper bound given by equation (3.7) the Banach-Alaoglu theorem allows us to conclude that it possible to extract a subsequence $\{n(t_k,x)\}_k$ such that

$$n(t,x) \xrightharpoonup[t\to\infty]{} \overline{n}(x),$$

where $\overline{n}(x)$ is a bounded non-negative measure such that

$$\int_{\mathbb{R}^d} \overline{n}(x) \leq I_M.$$

4th step: *Prove that $\overline{I} = \int_{\mathbb{R}^d} \overline{n}(x) = I_M$.* Assumption (3.5) allows us to conclude that for T sufficiently large we have

$$\sup_{t\geq T, |x|\geq A} R\big(x,I(t)\big) < 0.$$

This implies that for all $\varepsilon > 0$ we can find $A_\varepsilon > A$ such that

$$\sup_{t\geq T} \int_{|x|\geq A_\varepsilon} n(t,x)dx \leq \varepsilon.$$

In turn, this implies that for all $\varepsilon > 0$ we have

$$\int_{\mathbb{R}^d} \overline{n}(x)dx \geq \lim_{k\to\infty} \int_{|x|\leq A_\varepsilon} n(t_k, x)dx \geq I_M - \varepsilon.$$

This allows us to conclude that

$$\int_{\mathbb{R}^d} \overline{n}(x) = I_M.$$

5th step: *Characterisation of the measure* $\overline{n}(x)$. Since

$$n(t,x) = n^0(x) \exp\Big(\int_0^t R\big(x, I(s)\big)ds \Big),$$

knowing that

$$\int_{\mathbb{R}^d} \overline{n}(x) = I_M$$

we can conclude that $\text{supp}(\overline{n}) \equiv \{x \in \mathbb{R}^d : R(x, I_M) = 0\}$. Therefore, under assumptions (3.10) and (3.11) we can conclude that

$$\overline{n} = I_M \delta(x - \overline{x}).$$

□

The analytical results established by Theorem (3.2) are illustrated by the numerical results presented in Figure 7. These results have been obtained by constructing numerical solutions of the initial value problem (4.4) for $t \in [0, 50]$

$$n^0(x) := a\mathbb{1}_{[-2,2]}(x) \quad \text{with} \quad a = 10^8, \tag{3.23}$$

and

$$R\big(x, I\big) := r\Big(p(x) - d\, I\Big) \quad \text{with} \quad r = 0.5, \quad p(x) = 1 + \frac{2}{1 + (1-x)^2} \quad \text{and} \quad d = 10^{-8}. \tag{3.24}$$

Note that the definition given by equation (3.24) satisfies assumptions (3.2)-(3.5) with

$$I_m = \frac{1}{d}, \quad I_M = \frac{2}{d}, \quad K_+ = \frac{d}{2} \quad \text{and} \quad K_- = \frac{d}{4}$$

and assumption (3.10) with $\overline{x} = 1$. Further details of numerical simulations are provided in Appendix 2.

In agreement with the results of Theorem (3.2), $I(t) \to I_M$ and $n(t, x)$ concentrates at the point $x = 1$, which is the maximum point of $R(x, I_M)$.

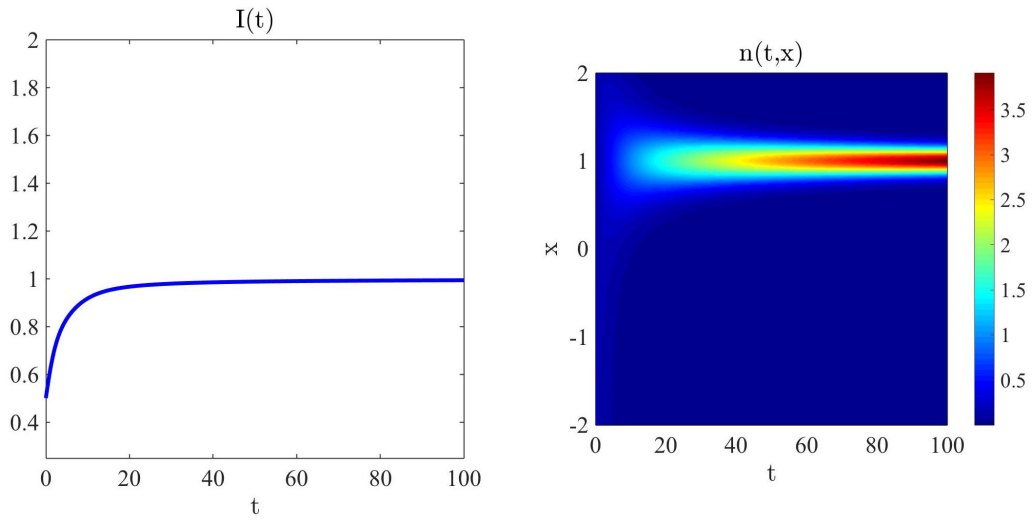


Figure 7: Sample dynamics of the total number of cells $I(t)$ (left panel) and the cell population density $n(t, x)$ (right panel) for the net proliferation rate $R(x, I(t))$ given by equation (3.24). The values of $I(t)$ and $n(t, x)$ are in units of I_M.

4 Conclusion

4.1 Future aspects

There is a great deal of work that can still be done in order to further explore the models presented in this project. The models have been studied in a rather mathematical framework with room to increase considerations of their biological applications. An idea for future research to build on this work could be to investigate how the analytical and numerical results compare with data obtained via experimentation. In turn, these results could then be incorporated into the models to increase the precision of the models' predictions. Furthermore, it would be interesting to conduct a series of experiments based on the information provided by this project to check the consistency of the models' hypotheses with biological data.

In his book, *Transport Equations in Biology*[12], Benoît Perthame dedicates a section to determining a more precise expression for the density of cells in the phenotypic state x, $n(t, x)$. The formula he provides gives a concentration profile for $n(t, x)$. This could be an interesting area of further study following this project. For example, it could be worthwhile to prove the expression for $n(t, x)$ as well as conduct some numerical analysis and see how it compares to less precise expression we have used for $n(t, x)$. The substitution of this new $n(t, x)$ in the IDE model could yield interesting and perhaps even disparate results.

Another noteworthy area of further study could be the inclusion of cytotoxic treatment in the IDE model. Albeit mathematically challenging, this could provide some important biological insights to the behavior of heterogeneous tumours under the influence of cytotoxic agents. This would pave the way to explore drug resistance in heterogeneous tumours, a very important issue in treatment optimisation. Inclusion of treatment in the IDE model would also allow for the comparison between the response of homogeneous versus heterogeneous tumours to similar treatment. This could lead

to further treatment optimisation if it was discovered that different tumours respond to treatment differently.

4.2 Final remarks

Through this project I have come to realise, more than ever, the inextricable link between mathematics and biology. I was, of course, aware of the importance of mathematical models before beginning this paper but I have been surprised at the precision, and hence importance, that is provided by mathematical models in a biological framework. This is a noteworthy consideration I will carry with me in my further studies.

There is much to be said of the merit of mathematics in cancer research, where it can be used to disprove previously made hypotheses and create new ones. Much can be understood and uncovered through the liaising of mathematics and biology, a notion I hope has been conveyed throughout this project.

Appendix 1 - Details of numerical simulations for the ODE model

The numerical simulations carried out in Section 2.2 are done using Euler's method of finding approximate numerical solutions to the ordinary differential equations (ODEs). In our case the ODEs are those describing tumour population dynamics without the presence of treatment *i.e* $c(.) = 0$ and are given by the initial value problem,

$$\begin{cases} \dfrac{dN}{dt}(t) = R\big(N(t)\big)\, N(t), \qquad t \in (0, \infty) \\ \\ N(t=0) = N^0 > 0. \end{cases} \tag{4.1}$$

Euler's method is an iterative process that relies on the following two equations,

$$t_{n+1} = t_n + \Delta t \tag{4.2}$$

$$N_{n+1} = N_n + \Delta t\, f(N_n, t_n), \tag{4.3}$$

where Δt is the time step. Relations (3.26) - (3.27) are derived by finding a Taylor series approximation for $N(t)$ around $t = t_0$. This is done in the following manner:

The Taylor series for $N(t)$ about an initial point $t = t_0$ is

$$\begin{aligned} N(t) &= N(t_0) + \left(\frac{dN}{dt}\right)_{t_0} (t - t_0) + O[(t - t_0)^2] \\ &= N_0 + f(N_0, t_0)(t - t_0) + O[(t - t_0)^2]. \end{aligned}$$

If we now let t increase in increments of Δt, the next point will be

$$t_1 = t_0 + \Delta t$$

To find N_1 we can use the Taylor approximation for $N(t)$ and setting $t = t_1$

$$N_1 = N(t_1) \approx N_0 + \Delta t\, f(N_0, t_0)$$

This process can then be repeated for as many n values as necessary for the solution to have a small enough error and a clear convergence to a solution[6].

In the case we are exploring, $f(N_n) = N_n\, R(N_n) = rN_n\big(1 - \frac{N_n}{K}\big)$. We begin the modelling by setting a numerical value for the initial condition, N_0. Two scenarios are chosen, the first is where $N_0 = 10^6$ and the second, $N_0 = 2 * 10^8$. The variable K, the maximum tumour size, must also be assigned a numerical value, which we have chosen to be $K = 10^8$. Furthermore, the intrinsic birth rate in conditions where nutrients and space are freely available, r, has been set to $r = 0.5$. This iteration is done for 500 values of n and is executed in MATLAB, yielding the results shown in the Section 2.2.

This is the code used in MATLAB to produce the plots in Section 2.2,

```
clear all
clc
close all

set(0,'DefaultAxesFontName', 'Times New Roman')
set(0,'DefaultAxesFontSize', 24)
set(0,'defaultaxeslinewidth',1)
set(0,'defaultpatchlinewidth',1)
set(0,'defaultlinelinewidth',3)

% Time discretisation
tend=50;
dt=0.1; %Time step
t = 0:dt:tend;

%%%%%%%%%%%%%%%%%%%%%%%%
% Parameter functions
r=0.5;
K=1e8;

%%%%%%%%%%%%%%%%%%%%%%%%
% ICs

N(1)=1e6
%N(1)=2*1e8;

%%%%%%%%%%%%%%%%%%%%%%%%
% TIME LOOP
i=1;
while i<=tend/dt
    R=r*(1-N(i)/K);
    N(i+1) = N(i) + dt*N(i)*R;
    i=i+1;
end

figure
plot(t,K*ones(1,length(t)),'r')
hold on
plot(t,N,'b--')
axis square
xlabel('t')
title('N(t)')
axis([0 tend 0 1.1*K])

figure
subplot(1,2,1)
```

```
plot(t,K/1e8*ones(1,length(t)),'r')
hold on
plot(t,N/1e8,'b--')
axis square
xlabel('t')
title('N(t)')
axis([0 tend 0 2*K/1e8])
subplot(1,2,2)
plot(t,K/1e8*ones(1,length(t)),'r')
hold on
plot(t,N/1e8,'b--')
axis square
xlabel('t')
title('N(t)')
axis([0 tend 0 2*K/1e8])
```

Appendix 2 - Details of numerical simulations for the IDE model

The numerical simulations performed in Section 3.3 are performed using a similar method as those in Section 2.2 (see Appendix 1). An iterative method is used to iterate over values of $n(t,x)$, which are in turn used to deduce the value of the population size, $I(t)$. The initial value IDE problem that is being solved is the following,

$$\begin{cases} \dfrac{\partial n}{\partial t}(t,x) = R\big(x, I(t)\big)\, n(t,x), & (t,x) \in (0,\infty) \times \mathbb{R}^d \\ \\ I(t) = \displaystyle\int_{\mathbb{R}^d} n(t,x)\, dx \\ \\ n(t=0) = n^0(x) \in L^1 \cap L^\infty(\mathbb{R}^d), \qquad n^0(x) \geq 0 \ \text{ for a.e. } x \in \mathbb{R}^d. \end{cases} \tag{4.4}$$

To set up the iteration, t and x must be discretised. We set t in the range $t \in [0, 100]$ divided into time steps, where the time step is $\Delta t = 0.01$ – this means there are a total of 10,000 time steps in the iteration. Furthermore, we have $x \in [-2, 2]$ divided into 400 equally spaced values. The bounds on x are defined as $x_m = -2$ and $x_M = 2$. Similarly, the spatial step, $\Delta x = 0.01$. Furthermore, the continuous function $n(t, x_i)$ needs to be discretised and is done so by using an approximate function, $n_i(t_k)$ so that it is possible to iterate over a range of x and t values. The iterative process relies on the following relations,

$$n(t_{k+1}, x) = n(t_k, x) + R(t_k, x)\, n(t_k, x)\, \Delta t \tag{4.5}$$

$$I(t_{k+1}) \approx \sum_i n(t_{k+1}, x_i)\, \Delta x \approx \sum_i n_i(t_{k+1}) \Delta x \tag{4.6}$$

Equations (3.24) and (3.25) are considered over a spatial domain and iterated over a specified time interval.

Next, we set the parameter functions for R and I_M,

$$R = r\left(1 + \frac{1}{1 + (1-x)^2} - I(t)\, d\right) \tag{4.7}$$

$$I_M = \frac{2}{d}, \tag{4.8}$$

where constants $r, d \in \mathbb{R}^d$ have been set to $d = 10^-8$ and $r = 0.5$.

To begin the iteration it is necessary to specify some initial conditions for the population density, $n(t,x)$. The initial condition is for $n^0(x)$,

$$n^0(x) = \frac{I_M}{2(x_M - x_m)}. \tag{4.9}$$

Thus, at time $t = 0$ the population density is constant for every value of x at a value of 25,000,000. We also can set an initial value for the population size $I(t)$,

$$I(0) \approx \sum_i n_i(0) \Delta x \tag{4.10}$$

Now that all the initial conditions have been set, it's possible to compute the iteration. The code yields two plots: the first is a the scaled population size, i.e. $\frac{I(t)}{I_M}$, as a function of time, t, the second is a density plot of the scaled population density, i.e. $\frac{n(t,x)}{I_M}$, as a function of t and x.

```
clear all
clc
close all

set(0,'DefaultAxesFontName', 'Times New Roman')
set(0,'DefaultAxesFontSize', 24)
set(0,'defaultaxeslinewidth',1)
set(0,'defaultpatchlinewidth',1)
set(0,'defaultlinelinewidth',3)

% Time & space discretisation

T=100;
dt=0.01; % Time step to solve IDE
Nt=T/dt;
time = 0:dt:T;

Nx=400;
xm=-2;
xM=2;
x=linspace(xm,xM,Nx);
dx=abs(x(2)-x(1));

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Parameter functions

p=1 + 1./(1+(1-x).^2);
r=0.5;
d=1e-8;

IM=2/d;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% ICs

a=0.5*IM;
b=0.5;
n(1,1:Nx)=a*1/(xM-xm); % cell population density
figure
plot(x,n(1,:))
pause
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
I(1)=sum(n)*dx;  % population size

%%%%%%%%%%%%%%%%%%%%%%%%
% TIME LOOP

for i=1:Nt-1
    R = r*(p-d*I(i));
    n(i+1,:) = n(i,:) + dt*R.*n(i,:);
    I(i+1) = sum(n(i+1,:))*dx;
end

figure
subplot(1,2,1)
plot(time(1:length(I)),I/IM)
axis square
xlabel('t')
title('I (t)')
subplot(1,2,2)
pcolor(time(1:length(I)),x,n'/IM)
axis square
shading flat
xlabel('t')
ylabel('x')
title('n (t,x)')
```

References

[1]Rehemtulla, Alnawaz. "Dinosaurs and Ancient Civilizations: Reflections on the Treatment of Cancer." NEOPLASIA 12 (Dec. 2010): 957-68. N.d. Web. 29 Jan. 2017.

[2]Byrne, Helen M. "Dissecting cancer through mathematics: from the cell to the animal model." Nature Reviews 10 (Mar. 2010): 221-30. Macmillan. Web. 29 Jan. 2017

[3]Altrock, Philipp M. and Liu, Lin L. and Michor, Franziska. "The mathematics of cancer: integrating quantitative models." Nature Reviews 15 (Dec. 2010): 730-45. Macmillan. Web. 29 Jan. 2017

[4]Korolev, Kirill S., Joao B. Xavier, and Jeff Gore. "Turning Ecology and Evolution Against Cancer." Nature Reviews 14 (May 2014): 371-80. Macmillan. Web. 16 Mar. 2017.

[5]Lipshultz, Steven E. et al. "Continuous Versus Bolus Infusion of Doxorubicin in Children With ALL: Long-Term Cardiac Outcomes." Pediatrics 130.6 (2012): 1003-1011. PMC. Web. 22 Mar. 2017.

[6]Neukirch, Thomas, Prof. MT2507 Mathematical Modelling 2014-2015. N.p.: n.p., n.d. Web. 24 Mar. 2017.

[7]Lorenzi, Tommaso, Rebecca H. Chisholm, and Jean Clairambault. "Tracking the Evolution of Cancer Cell Populations through the Mathematical Lens of Phenotype-structured Equations." Biology Direct 11.43 (2016): 1-17. Web. 16 Apr. 2017.

[8]Merlo, Lauren M.F., John W. Pepper, Brian J. Reid, and Carlo C. Maley. "Cancer as an Evolutionary and Ecological Process." Nature Reviews 6 (2006): 924-35. Web. 16 Apr. 2017.

[9]"Article 6." Advanced Course on Mathematical Biology. N.p.: n.p., n.d. 189-201. Print.

[10]Perthame, Benoît. Some Mathematical Models of Tumor Growth. Paris: Universite Pierre Et Marie Curie-Paris, 2015. Print.

[11]Rothschild, Bruce M et al. "Metastatic cancer in the Jurassic." The Lancet (July 1999), Volume 354, Issue 9176, 398. Web. 19 Apr. 2017

[12]Perthame, Benoît. "Transport Equations in Biology." Benoît Perthame — Springer. Birkhaüser Basel, n.d. Web. 19 Apr. 2017.