

POLITECNICO DI TORINO



**Politecnico
di Torino**

**Mathematical modelling of
intermittent radiotherapy protocols for
recurrent high-grade glioma**

Master Degree in Physics of Complex Systems

Supervisor:

Prof. Marcello Edoardo Delitala

Candidate:

Francesco Albanese

Academic year 2021/2022

Abstract

Cancer is one of the world's deadliest diseases. Number of researchers have put time and efforts in finding an effective treatment, improving efficiency of current treatments and finding the ways to help the patients develop their immune system that enable them to fight it.

Being a leading cause of death, cancer's mechanisms of growth and destruction are widely investigated. Mathematical models explaining these mechanisms significantly help to predict the behaviour of cancer cells proliferation.

In this thesis, the main focus is on what's referred to as high-grade glioma, a brain cancer whose prognosis is rarely positive. Despite there has been a continuous progress, standard treatments often lead to a poor outcome, suggesting that innovative approaches should be considered.

A possibility in this sense is represented by deviations from traditional radiotherapy protocols. Here, different fractionation schedules are explored by means of a genetic algorithm. Starting from a certain mathematical model, fitted using a dataset of magnetic resonance images showing longitudinal tumour volumes during a hypofractionated stereotactic radiotherapy, the aim is to predict specific features for each patient and use them to provide a personalized radiotherapy protocol capable of improving the final therapy outcome, for example by further extending the patient survival time.

A final analysis concerning the model reliability has been done, providing a concrete measure about how much accurate is the parameters' prediction over time, which is crucial since personalized medicine requires promptness in the treatment choices as the disease advances.

Not only, the problem of a raising therapy resistance due to the remarkable cancer heterogeneity is a constant challenge in the clinical path. Multicompartmental models are thus proposed to quantify the growth of a cancer resistant population, in order to exploit the genetic algorithm in designing effective countermeasures.

Results of *in silico* experiments are promising: in the limit of an extremely simplified picture of a complex system as that of tumour dynamics, virtual improvements have been achieved. This suggest that further efforts in the direction of an ever more personalized medicine may be worth it, hopefully leading to a less dismal scenario.

Contents

1	Introduction	4
1.1	Modelling tumour	4
1.2	Deterministic mathematical models for cancer dynamics	5
1.2.1	Exponential model	5
1.2.2	Logistic model	5
1.2.3	Interacting models	6
1.2.4	Why ODE models	7
1.3	The need for an adaptive therapy	8
1.4	Glioma	9
1.5	Radiotherapy and fractionation	10
1.6	Cancer is a complex system	12
2	Genetic Algorithms	13
2.1	Overview	13
2.2	Structure of a Genetic Algorithm	15
2.2.1	Chromosomes encoding	15
2.2.2	Fitness function	15
2.2.3	Selection	16
2.2.4	Recombination	17
2.2.5	Evolution	19
2.3	GA parameters	20
2.4	A brief insight into GAs theory	21
2.4.1	The Schema Theorem	21
2.4.2	The Building Block Hypothesis	23
2.5	Why choosing the GA method	23
3	Monoclonal description of recurrent high-grade glioma	25
3.1	The mathematical model	26
3.2	Linear Stability Analysis	29
3.3	Parameters fitting through GA	31
3.3.1	Fitting results	32
3.4	Intermittent Radiotherapy Protocol	42
3.4.1	Results	44
3.4.2	Data Analysis	53

3.5	Hyperfractionation	56
3.5.1	Results	56
3.5.2	Data Analysis	65
3.6	Model reliability	69
4	Radiation resistance: polyclonal modelling of recurrent high-grade glioma	72
4.1	Multicompartmental model	74
4.1.1	Linear Stability Analysis	75
4.1.2	Model fitting	77
4.1.3	Data Analysis & Clustering	86
4.2	Maximizing Clonal Inversion Time	89
4.2.1	IRT	90
4.2.2	Hyperfractionation	99
4.3	Competitive models	103
4.3.1	Including intraspecific competition	103
4.3.2	Including interspecific competition	104
4.3.3	Model fitting	104
4.3.4	IRT	114
	Conclusion	117
A	GA template for model parameters' fitting	118
B	GA template for the optimal IRT protocol search	127
	Bibliography	134
	Acknowledgements	136

Chapter 1

Introduction

1.1 Modelling tumour

Cancer is one of the world's deadliest diseases, second leading cause of death in the world after cardiovascular diseases [16]. Number of researchers have put time and efforts in finding an effective treatment, improving efficiency of current low cost treatment and finding the ways to help the patient develop their immune system that enable them to fight cancer.

Cancer is typically initiated by genetic mutations that lead to enhance the abnormal of proliferation rate and cell growth. After a certain size, cancer cell starts spreading to the other parts of body and this process is called metastasis. Approximately 90% of cancer deaths are due to cancer metastasis.

Mathematical models explaining cancer's dynamics are crucial to predict the behaviour of cancer cells proliferation.

Literature dealing with mathematical modelling of cancer initiation, proliferation and metastasis is abundant. Mathematical models to simulate the growth rate of the cancer cells have been derived from both deterministic and stochastic considerations. Early model of tumor growth by diffusion was first introduced and then set the scene for many later mathematical models.

Many deterministic models have been used to describe the behaviour of cancer cell growth and proliferation. Cancer cell growth and proliferation is subjected to the uncontrolled factors or environmental noise which includes for example cellular metabolism, hormonal oscillations and individual characteristics such as body mass index, genes, smoking and stress impact.

Deterministic models in fact are inadequate to explain in-depth the dynamical process of the cancer cell proliferation. In such a case, research has started to

extend the deterministic models to their stochastic counterpart. Nevertheless, deterministic models are still widespread in the scientific community, since they allow to capture at least many important macroscopic features of cancer developing and treatment.

This thesis massively relies on these models, aiming to explore new modelling behaviours without claiming to replace the solid scientific background of reference, but rather with the purpose of combining them with other study techniques such as Genetic Algorithms (presented in chapter 2) and numerical optimization (presented in chapter 3) in order to step forward from a generic description to what's called precision medicine, which is essential to adapt therapies meeting the needs of every single patient.

1.2 Deterministic mathematical models for cancer dynamics

1.2.1 Exponential model

Exponential model is the natural description of early stages of cancer growth. In exponential model, each cancer cell split into two daughter cells in the affected area with a constant rate. The exponential model is given by:

$$\frac{dV(t)}{dt} = \lambda V(t) \quad (1.1)$$

where, λ is the growth parameter and $V(t)$ is the volume of the cancer cells. The cancer cell growth in exponential model is proportional to the population of the cancer. However, at the last stages, the exponential model fails to predict the angiogenesis process and reduction of the nutrient. Extension of the exponential model thus is required.

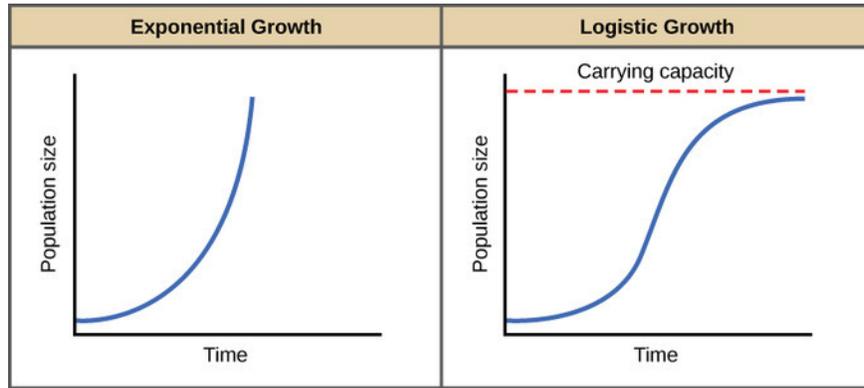
1.2.2 Logistic model

Angiogenesis is a complex process in which there is growth of new blood vessels from the pre-existing ones and is an essential phenomenon for the growth and survival of solid neoplasms. Tumour angiogenesis is the proliferation of blood vessels penetrating the cancerous growth for the supply of nutrients and oxygen. Angiogenesis is a requisite not only for continued tumour growth, but also for metastasis [18].

The exponential model has limitations to predict the long-term growth rate of cancer cell proliferation and the metastasis phenomenon. To overcome these problems, a logistic model was introduced to explain the behaviour of cancer cell growth and proliferation.

The logistic model equation describes that the growth is proportional linearly with size until the growth of the cells reach the carrying capacity, K . Logistic equation produces S-shape curve for the volume of cancer cell. This model can interpret the mutual competition between the cells (intraspecific competition). The generalized logistic equation is

$$\frac{dV(t)}{dt} = \lambda V(t) \left(\frac{K - V(t)}{K} \right) \quad (1.2)$$



1.2.3 Interacting models

In chapter 4 the starting point will be the assumption about the existence of a heterogeneous scenario within cancer. Experimental evidence suggests that tumour cells respond to therapies in many complex ways, in particular they mutate in order to survive the dangerous effects they undergo. It is a fact that cancer, in a long-time perspective, evolves in such a way to neutralize almost completely the effectiveness of the treatment.

The simplest way to take into account this phenomenon is introducing a two-population (“two compartments”) model for tumour cells: the first sensible to the treatment and the second resisting it. At some point it turns out natural to consider some sort of interaction between these two kinds of cell.

When species interact the population dynamics of each species is affected [13]. There are three main types of interaction:

- 1) If the growth rate of one population is decreased and the other increased the populations are in a predator–prey situation.
- 2) If the growth rate of each population is decreased, it is competition.
- 3) If each population’s growth rate is enhanced, then it is called mutualism or symbiosis.

The competitive scenario suits cancer behaviour: cells with different features compete for the same limited resources or in some way inhibit each other’s growth. A description of this struggling necessitates for a refinement of the previous models, mathematically through a non-linear term (or more) in order

to include what's called interspecific competition.

Consider two species V_s and V_r having logistic growth in the absence of the other. Inclusion of logistic growth in the system makes them much more realistic, but sometimes it is not strictly necessary. The dynamics of these two populations can be described as:

$$\frac{dV_s(t)}{dt} = \lambda_s V_s(t) \left(\frac{K_1 - V_s(t) - b_{sr} V_r(t)}{K_1} \right) \quad (1.3)$$

$$\frac{dV_r(t)}{dt} = \lambda_r V_r(t) \left(\frac{K_2 - V_r(t) - b_{rs} V_s(t)}{K_2} \right) \quad (1.4)$$

where $\lambda_s, K_1, \lambda_r, K_2, b_{sr}$ and b_{rs} are all positive constants and, as before, the λ 's are the linear birth rates and the K 's are the carrying capacities. The b_{sr} and b_{rs} measure the competitive effect of V_r on V_s and V_s on V_r respectively: they are generally not equal.

1.2.4 Why ODE models

Mathematically speaking, models of this kind, using one or several ODEs, are naive and oversimplified compared to various other kinds of models, reaction-diffusion PDEs are one example among many used in current mathematical and computer biology and medicine.

However, models used routinely by experimental biologists and clinicians are the simplest ODE models, which form the foundation of applied biological modelling in practice. The most important case in point is the equation of exponential growth itself; discussions of the value of the Malthusian growth parameter under various conditions still form a major portion of practical tissue culture and tumour modelling. More generally, it is usually the specification of parameter values for simple ODE models that is the crux of the discussion.

Such models aim to capture key features using a small number of adjustable parameters and, equally important, aim to neglect peripheral features judiciously. Often a mathematically quite sophisticated model is basically a marginal elaboration of some simple ODE model.

The physics of tumorigenesis is heavily complex, at the point that it is practically impossible to take into account every possible interaction, but useful insight can be obtained through a smart usage of main physical considerations combined with the tools of applied mathematics.

1.3 The need for an adaptive therapy

A number of successful systemic therapies are available for treatment of disseminated cancers.

However, tumour response is often transient, and therapy frequently fails due to emergence of resistant populations. The latter reflects the temporal and spatial heterogeneity of the tumor microenvironment as well as the evolutionary capacity of cancer phenotypes to adapt to therapeutic perturbations. Although cancers are highly dynamic systems, cancer therapy is typically administered according to a fixed, linear protocol.

In [6], Gatenby et al. employ a mathematical model which finds that when resistant phenotypes arise in the untreated tumor, they are typically present in small numbers because they are less fit than the sensitive population. This reflects the “cost” of phenotypic resistance such as additional substrate and energy used to upregulate xenobiotic metabolism (i.e. the metabolism of a foreign and unusual substance in the tissue, the cancer itself) and therefore not available for proliferation, or the growth inhibitory nature of environments (i.e., ischemia or hypoxia) that confer resistance on phenotypically sensitive cells. Thus, in the Darwinian environment of a cancer, the fitter chemosensitive cells will ordinarily proliferate at the expense of the less fit chemoresistant cells.

The models show that, if resistant populations are present before administration of therapy, treatments designed to kill maximum numbers of cancer cells remove this inhibitory effect and actually promote more rapid growth of the resistant populations.

Talking about a general therapy (usually a combination of chemotherapy and radiotherapy interventions), the goal of adaptive therapy is to enforce a stable tumour burden by permitting a significant population of sensitive cells to survive so that they, in turn, suppress proliferation of the less fit but resistant subpopulations. Computer simulations in this thesis will illustrate that this could be achieved in some cases, overcoming the static treatment schedules.

It is worth mentioning that adaptive therapies aim not only to eradicate the resistance phenomenon, but to improve the patient’s quality of life, avoiding undesired side effects. For example, many current chemotherapy regimens have, as a fundamental strategy, the goal of killing maximal numbers of tumour cells. Usually, this is achieved through application of the highest drug dose that results in acceptable patient toxicity. This limit is “dynamic”, in the sense that knowing a priori the maximum dosage is hard, because many factors during therapy influence the patient’s health status. Adaptive therapy then considers these aspects and design a precise treatment in response to this high dynamicity. In this thesis, the adaptive traits are restricted to the previously discussed raising resistance feature.

1.4 Glioma

This thesis concerns brain tumours, and in particular gliomas, which make up about half of all primary brain tumours diagnosed; they are particularly nasty tumours with a depressingly dismal prognosis for recovery. Gliomas are highly invasive and infiltrate the surrounding tissue. The impressive increased detection capabilities (but unfortunately still woefully inadequate) in computerized tomography (CT) and magnetic resonance imaging (MRI) over the past years have resulted in earlier detection of glioma tumours. Despite this progress, the benefits of early treatment have been minimal [12].

Gliomas are neoplasms of glial cells (neural cells capable of division) that usually occur in the upper cerebral hemisphere, but which can be found throughout the brain and nervous system.

Tumour grade indicates the level of malignancy and is based on the degree of anaplasia (or deformity in behaviour and form) seen in the cancerous cells under a microscope. Gliomas often contain several different grade cells with the highest or most malignant grade of cells determining the grade, even if most of the tumour is lower grade.

Generally, the higher-grade cancer cells are more capable of invading normal tissue and so are more malignant. However, even with their invasive abilities, gliomas very rarely metastasize outside the brain (and the nervous system in general). Here, the focus will be on high-grade gliomas, which are particularly hard to counter.

An enormous amount of experimental and some theoretical work has been devoted to trying to understand why gliomas are so difficult to treat. Unlike many other tumours, gliomas can be highly diffuse. Experiments indicate that within 7 days of tumour implantation in a rat brain, glioma cells can be identified throughout the central nervous. A locally dense tumour growth remains where the cancerous tissue was initially implanted but there are solitary tumour cells throughout the central nervous system.

There are various, regularly used, treatments for gliomas, mainly chemotherapy, radiation therapy and surgical intervention. Chemotherapy essentially uses specialized chemicals to poison the tumour cells. The brain is naturally defended from these and other types of chemicals by the intricate capillary structure of the blood–brain barrier.

Many chemotherapeutic treatments are cell-cycle-dependent: the drugs are triggered by certain phases of the cell cycle.

Silbergeld and Chicoine [4] have observed that the motile cells distant from the bulk tumour do not appear to enter mitosis so cell-cycle specific drugs and stan-

standard radiation therapy have limited effectiveness. Not only that, gliomas are often heterogeneous tumours. Those drugs that do reach the cancerous cells are hindered by drug resistance commonly associated with cancer cell heterogeneity. While one cell type is responsive to treatment and dies off, other types are waiting to dominate. This phenomenon requires a model which includes cell mutation to resistant cells, in other words a polyclonal model.

The biological complexity of gliomas makes treatment a difficult undertaking. For planning effective (or seemingly so) treatment strategies, information regarding the growth rates and invasion characteristics of tumours is crucial.

The use of mathematical modelling can help to quantify the effects of chemotherapy and radiation on the growth and diffusion of malignant gliomas. However standard treatments for high-grade glioma have a very poor record of success. There is a pressing need for a totally different approach to the treatment of gliomas, several of which are currently being investigated, for example the intermittent strategies proposed in this thesis.

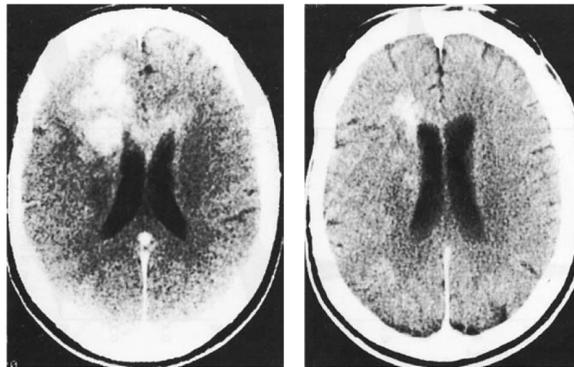


Figure 1.1: CT scans during the terminal year of a patient with anaplastic astrocytoma (another type of brain cancer) who was undergoing chemotherapy and radiation treatment. The image on left was taken approximately 180 days after the image on the right, showing how the treatment didn't manage to prevent the tumour's wild growth. (Figures from Tracqui et al. 1995)

1.5 Radiotherapy and fractionation

Radiation therapy is the use of high-energy radiation to damage cancer cells' DNA and destroy their ability to divide and grow. It may be delivered using machines called linear accelerators or via radioactive sources placed inside the patient on a temporary or permanent basis.

Preparation for radiation therapy is focused on targeting the radiation dose to the cancer as precisely as possible to minimize side effects and avoid damaging normal cells.

In the following, the main focus will be on radiotherapy applications, in particular about the possibility to explore new fractionation schedules which could possibly prevent the resistant cells development which arises as the therapy goes on.

The need for radiotherapy personalization is now widely recognized, however, it would require considerations not only on the probability of control and survival of the tumour, but also on the possible toxic effects, on the quality of the expected life and the economic efficiency of the treatment [2].

Among treatment options, radiotherapy is the most applied because it can be used either to reduce the extent of the tumour before proceeding with surgery or to irradiate the resection margins post-surgery. It can be also used as a palliative therapy and as an elective treatment alone or concomitantly with adjuvant chemotherapy).

External Beam Radiation Therapy (EBRT) is usually delivered to the patients by means of multiple fractions characterized by the nominal dose that must be conveyed to the region of interest including visible tumour and micro-lesions. Conventional treatment consists of 1.8–2 Gy fractions delivered 5 days a week, a therapeutic regimen established in early radiobiological studies to maximize the curative effect while reducing toxicity.

Recently, the identification of patient-specific genomics and radiomics biomarkers has suggested the possibility of exploiting altered regimens. An accurate and personalized approach to EBRT planning would require at least two steps:

- 1) definition of the most suitable fractionation program, including the nominal dose value per fraction, based on patient-specific characteristics.

- 2) accurate delivery of the nominal dose taking into account the anatomical-pathological changes between fractions and intra-fraction organ movement. Regarding dose delivery, irradiation is usually carefully planned by optimizing the beam entry and activation strategies to administer the tumour with the prescribed amount of dose while sparing the organ at risks (OAR). The dose profile can be adjusted according to slow morphological changes (inter-fraction) using a plan-of-the-day approach.

In the light of this clinical context, mathematical models of tumour evolution and response to treatment could play an important role allowing the customization of radiotherapy simulating different irradiation protocols and thus supporting the selection of the most effective one.

In this thesis, therapy adaptation is limited only to the seek of few (but fundamental) features about how tumour volume response is managed, without considering dose quantity and OAR feedback. Tools such as GAs and numerical optimization will be exploited to propose several different radiotherapy fractionations with a mathematical modelling background of increasing complexity.

1.6 Cancer is a complex system

In recent years, there have been a number of advances in understanding the dynamics of cancer that have been achieved through a combination of in vivo or in vitro experiment and computational simulation using quantitative theoretical models, sometimes referred to as in silico experimentation.

However, cancer is a complex system: mathematical models need to be sufficiently realistic to provide predictive understanding for cancer behaviour and this implies complexity. At the same time, complex simulations whose properties are not well understood risk merely replacing one intractable problem with another.

A typical approach is to use low-dimensional systems of coupled differential equations to study the dynamics of immune systems. In the following, multi-compartmental models are developed to study the ability of cancer to evolve and resist to the applied therapies. Typical aims of this sort of study are to identify intercompartmental transfer rates, to reproduce and explain phenomena observed in vivo and to explore scenarios under which particularly beneficial or pathological responses can occur.

It is often difficult to achieve a realistic compartmental model. Solutions to the differential equations that faithfully capture observable behavioural phenomena depend crucially on the values of intercompartmental flow rates and other system parameters. These are often difficult or impossible to measure directly and so values are estimated from available data using parameter estimation techniques.

Together with finding optimal treatment schedules, this is the main challenge of the whole thesis.

Chapter 2

Genetic Algorithms

2.1 Overview

Genetic Algorithms (GAs) are a family of search and optimization heuristic techniques inspired by natural (Darwinian) evolution [10]. They make use of a chromosome-like data structure [17], which encodes a potential solution to a specific problem. Starting from a certain initial population of chromosomes (typically random), the GA assigns a fitness to each of them, namely a score that quantifies how much “optimal” the solution is.

Then chromosomes are ranked and the ones with larger fitness have a higher chance of being selected for mating and reproducing. Just like real chromosomes, they pass through the steps of recombination and (random) mutation of the genes, thus generating an offspring with new genomes which are expected to have a higher fitness.

The procedure is iterated with the next generations over and over to improve the quality of the solutions encoded by the population. Once certain termination criteria are satisfied, the GA stops, and among the final population, one hopefully extracts the optimal solution for the initial problem.

GA theory is an active and growing area [9], widely used by researchers in many contests and applications such as software engineering, neural networks, image processing, speech recognition, healthcare, machine learning, etc. . .

In this thesis, GAs represent the main tool employed in order to explore a definitely complex search space of solutions, for example that one related to the parameters estimation of highly non-linear and non-differentiable functions, where deterministic gradient-based methods may easily fail.

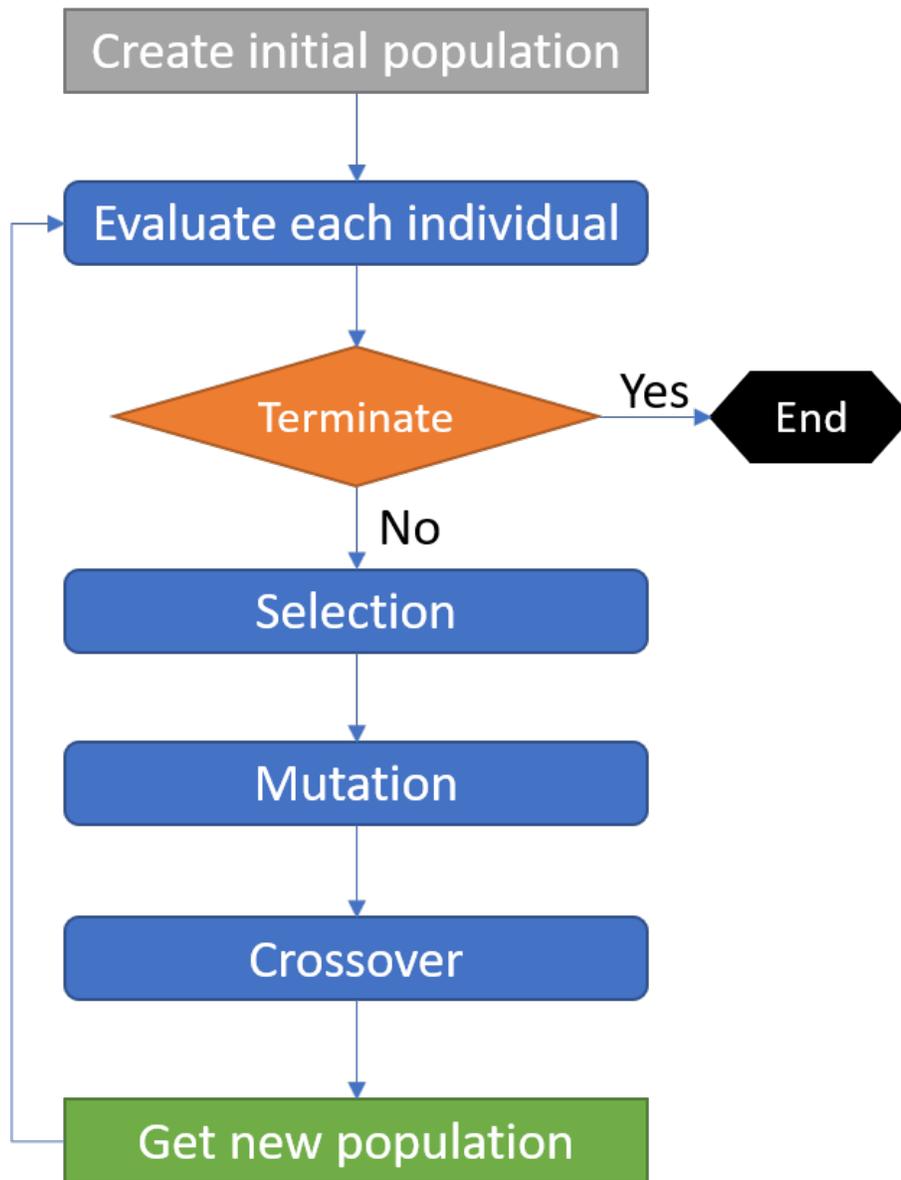


Figure 2.1: Flow chart of a generic GA

2.2 Structure of a Genetic Algorithm

A GA is constructed from a number of distinct components. This is a particular strength because it means that standard components can be re-used, with trivial adaptation in many different GAs, thus easing implementation. As anticipated, the main components are the chromosome encoding, the fitness function, selection, recombination, and the evolution scheme. The particular choices of these components allow to identify several kinds of GAs; In this thesis, the choices are in line with what's referred to as the *Canonical Genetic Algorithm*.

2.2.1 Chromosomes encoding

A GA manipulates populations of chromosomes, which are string representations of solutions to a particular problem. A chromosome is an abstraction of a biological DNA chromosome, which can be thought of as a string of letters from the alphabet A, C, G, T. A particular position or *locus* in a chromosome is referred to as a *gene* and the letter occurring at that point in the chromosome is referred to as the *allele*.

The Canonical GA uses a bit-string representation to encode solutions. Bit-string chromosomes consist of a string of genes whose allele values are characters from the alphabet $\{0,1\}$. The interpretation of these strings is entirely problem dependent. For example, a bit string of length 20 might be used to represent a single integer value (in standard binary notation) in one problem, whereas, in another, the bits might represent the choice of a particular dose fractionation in a therapy. It is a strength of GAs that common representations can be used in this way for a multiplicity of problems, making it faster and easier to apply GAs to new situations.

On the other hand, the consequence is that the chromosome encoding alone will contain only limited problem-specific information. For example a range of values can be represented with some discretization step.

In Figure 2.2, some examples of chromosome string encodings. The Binary representation is the one employed in the Canonical GA.

2.2.2 Fitness function

The fitness function is a numeric value that evaluates the quality of the chromosome as a solution to a particular problem. By analogy with biology, the chromosome is referred to as the *genotype*, whereas the solution it represents is known as the *phenotype*.

This is reminiscent of biological evolution, where the chromosomes in a DNA molecule are a set of instructions for constructing the phenotypical organism. A

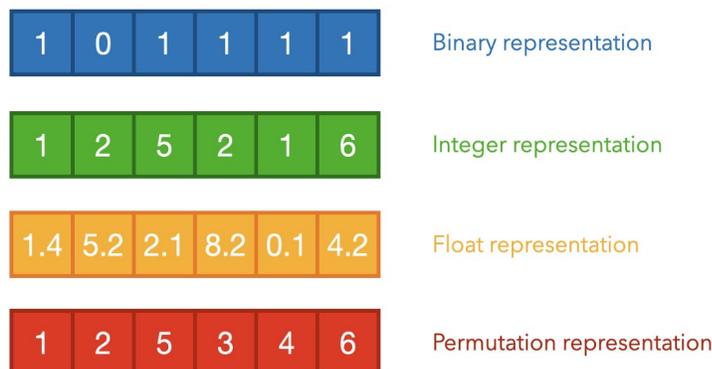


Figure 2.2: Examples of chromosome string encodings

complex series of chemical processes transforms a small collection of cells containing the DNA into a full-grown organism, which is then “evaluated” in terms of its success in responding to a range of environmental factors and influences.

As anticipated, the evaluation process consists of choosing individuals for mating based on their fitness value. In the following chapters, the chosen fitness function for a particular problem will be specified.

2.2.3 Selection

A GA uses fitness as a discriminator of the quality of solutions represented by the chromosomes in a GA population. The selection component of a GA is designed to use fitness to guide the evolution of chromosomes by selective pressure. Chromosomes are therefore selected for recombination on the basis of fitness. Those with higher fitness should have a greater chance of selection than those with lower fitness, thus creating a selective pressure towards more highly fit solutions.

The selection method used in this work is *Roulette Wheel* (or fitness proportional) selection. This allocates each chromosome a probability of being selected proportional to its relative fitness, which is its fitness as a proportion of the sum of fitnesses of all chromosomes in the population.

Suppose a population contains M chromosomes, with fitnesses f_1, f_2, \dots, f_M respectively. We then select $M' \leq M$ times from the population according to the

following scheme:

- 1) Calculate the total fitness: $f = \sum_{i=1}^M f_i$
- 2) Calculate the proportion $p_i = \frac{f_i}{f}$ of total fitness for each chromosome.
- 3) Divide the unit interval $[0, 1]$ into M subintervals $[t_0, t_1], (t_1, t_2], \dots, (t_{M-1}, t_M]$ where $t_0 = 0$ and: $t_i = \sum_{k=1}^i t_k$
- 4) (Repeat M' times) Calculate a random number, r , in $[0, 1]$; r will lie in a subinterval containing precisely one t_i . Select chromosome i^{th} for reproduction.

Note that $[t_{i-1}, t_i]$ has length p_i , $1 \leq i \leq M$, so that $[0, 1]$ is partitioned proportionately according to the relative fitness of the M chromosomes in the population. This process results in composition of a *mating buffer* (MB), which is a subset of the original population and consists of the M' chromosomes selected for reproduction. Since the solutions are marked proportionally to their fitness, a solution with a higher fitness is likely to receive more copies than an inferior solution.

Although easy to implement, roulette wheel selection shows a major drawback: if a population contains a solution whose fitness substantially exceeds those of the rest of the solutions, this 'supersolution' very soon will dominate the population, which will inevitably lose its diversity and the GA converges prematurely (excess of selective pressure). In the GAs used during this study the possibility of having more than one copy while building the MB has been neglected, in order to prevent massive dominance of supersolutions, at a cost of a larger time of convergence.

2.2.4 Recombination

Recombination is the process by which chromosomes from MB are recombined to form members of a successor population. The idea is to simulate the mixing of genetic material that can occur when organisms reproduce. Since selection for recombination is biased in favour of higher fitness, the balance of probabilities (hopefully) is that more highly fit chromosomes will evolve as result. There are two main components of recombination, the genetic operators *crossover* and *mutation*. Genetic operators are nondeterministic in their behaviour. Each occurs with a certain probability and the exact outcome of the crossover or mutation is also nondeterministic.

The crossover operator represents the mixing of genetic material from two selected parent chromosomes to produce a child chromosome. After two parent chromosomes have been selected for recombination, a random number in

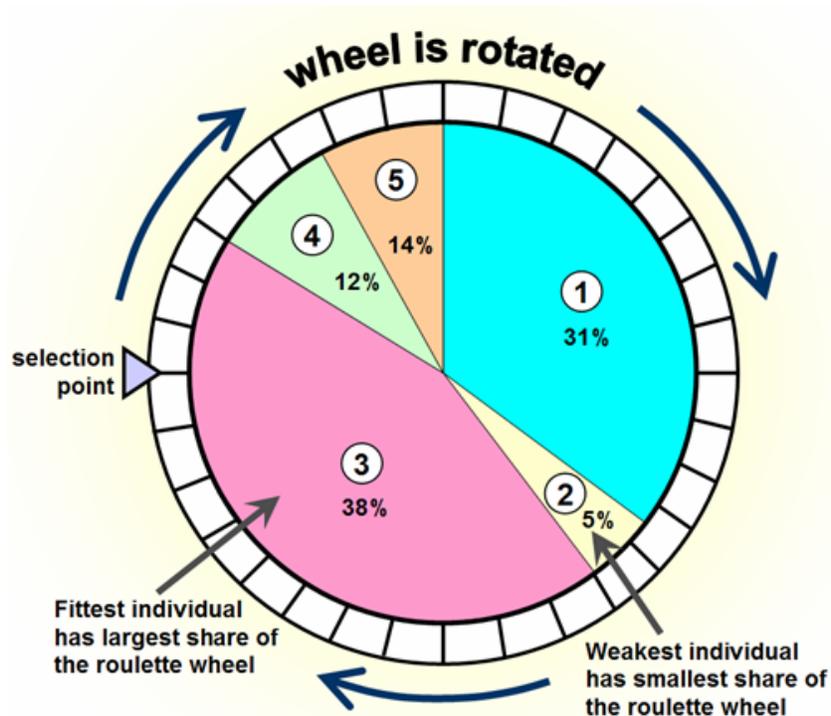


Figure 2.3: The roulette-wheel selection

the interval $[0,1]$ is generated with uniform probability and compared to a predetermined “crossover rate”. If the random number is greater than the crossover rate, no crossover occurs and one or both parents pass unchanged on to the next stage or recombination. If the crossover rate is greater than or equal to the random number, then the crossover operator is applied. One commonly used crossover operator is *one-point crossover*: a crossover point between 0 and L (with L being the length of each chromosome) is chosen with uniform probability. Child chromosomes are then constructed from the characters of the first parent occurring before the crossover point and the characters of the second parent occurring after the crossover point.

Mutation operators act on an individual chromosome to flip one or more allele values. In the case of bit-string chromosomes, the normal mutation operator is applied to each position in the chromosome. A random number in the interval $[0,1]$ is generated with uniform probability and compared to a predetermined “mutation rate”. If the random number is greater than the mutation rate, no mutation is applied at that position. If the mutation rate is greater than or equal to the random number, then the allele value is flipped from 0 to 1 or vice

versa. Mutation rates are typically very small.

Mutation creates new solutions to avoid local optima: using the crossover operator alone to produce an offspring makes the GA stuck in the local optima, thus, the good parts of the parents survive in each generation, and the local optimal ones are to be found. This problem is called as the local-optima problem. The mutation operator is used to alleviate this problem by proving new offspring different from parents, and this encourages diversity in the population. However, an abuse of mutation events may drive the GA to lose important information essential to reach optimal solutions. The trade-off between recombination and mutation probabilities is one the main concerns when the algorithm is designed.

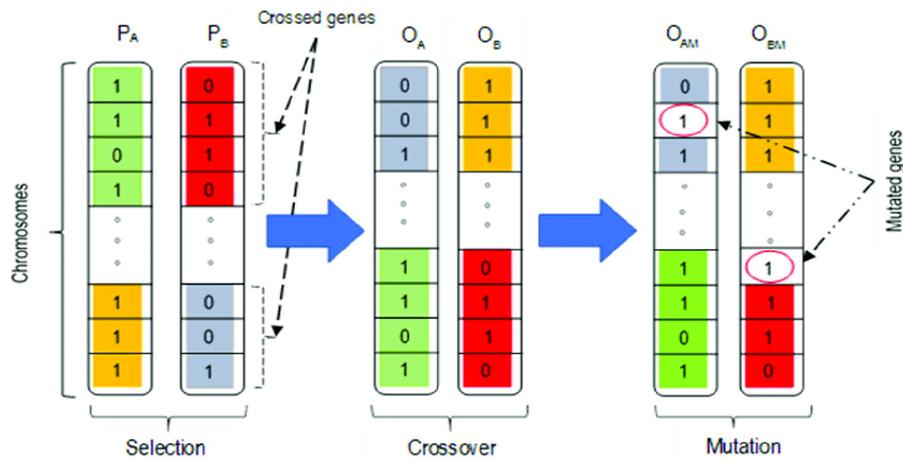


Figure 2.4: Example of one-point recombination

2.2.5 Evolution

After recombination, resultant chromosomes are passed into the successor population. The processes of selection and recombination are then iterated until a complete successor population is produced. At that point the successor population becomes a new source population (the next generation).

The GA is iterated through a number of generations until appropriate stopping criteria are reached. These can include a fixed number of generations having elapsed, observed convergence to a best-fitness solution, or the generation of a solution that fully satisfies a set of constraints.

There are several evolutionary schemes that can be used. These range from *complete replacement*, where all members of the successor population are generated through selection and recombination to the so called *steady state*, where

the successor population is created by generating one new chromosome at each generation and using it to replace a less-fit member of the source population. The choice of evolutionary scheme is an important aspect of GA design and will depend on the nature of the solution space being searched.

In the present work a stochastic evolution scheme similar to steady state is used: a member from the old population is selected randomly and its fitness is compared with one new chromosome from the MB, then choosing the one with highest fitness. The process is repeated for every chromosome in the MB.

2.3 GA parameters

Determining the interactions that occur among different GA parameters has a direct impact on the quality of the solution, and keeping parameters values "balanced" improves the solution of the GA. There are four basic and important parameters used by GA, those include:

1) Crossover rate (probability): it determines how often crossover processes occur for chromosomes in one generation (i.e., the chance that two chromosomes exchange some of their parts). 100% crossover rate means that all offspring are made by crossover. If it is 0%, then the complete new generation of individuals is to be exactly copied from the older population, except those resulted from the mutation process. Crossover rate is in the range of $[0, 1]$.

2) Mutation rate (probability): this rate determines how many chromosomes should be mutated in one generation. The purpose of mutation is to prevent the GA from converging to local optima, but if it occurs very often, GA is changed to random search (Recall the crossover-mutation trade-off). Also mutation rate is in the range of $[0, 1]$.

3) Population size: the size of the population indicates the total number of the population's inhabitants. Selection of population size is a sensitive issue; if the size of the population (search space) is small, this means little search space is available, and therefore it is possible to reach a local optimum. although, if the population size is very large, the area of search is increased and the computational load becomes high. Therefore, the size of the population must be reasonable.

4) Number of generations: It refers to the number of cycles before the termination. In some cases, hundreds of loops are sufficient, but in other cases we might need more, this depends on the problem type and complexity. Depending on the design of the GA, sometimes this parameter is not used, particularly if the termination of the GA depends on specific criteria.

These GA parameters are determined whether the GA finds an optimal or near-optimal solution, and whether it finds an efficient solution. Any change in the value of these parameters (increasing or decreasing) affects the result of GA negatively or positively. Tuning the right parameters is a nontrivial task.

2.4 A brief insight into GAs theory

Although GAs have enjoyed significant practical success, attempts to establish a theoretical account of their precise operation have proved more difficult. There are two goals for a satisfactory theory of GAs. The first is to explain which classes of problem GAs are particularly suitable for and why. The second is to provide techniques and approaches for optimal design and implementation of GAs, as there are many choices of structure and parameters to be made.

2.4.1 The Schema Theorem

The Schema Theorem is a central result of GA theory. It attempts to explain how the evolutionary processes in a GA can locate optimal or near-optimal solutions, even though they only ever sample a tiny fraction of the set of all possible solutions. A schema is a pattern within a chromosome defined by fixing the values of specific chromosome loci. A schema defines a set of chromosomes, namely all those containing the pattern.

A schema is a string of symbols from the alphabet $0, 1, *$. For example, a schema for 10-bit chromosomes might be specified using the string

01***100**

Chromosomes which belong to this schema include

0110110000
0100010001
0111110011

Chromosomes that do not match the pattern do not belong to the schema and include

1010110000
1111100000

All chromosomes belong to the schema

For a particular schema H , we define the length l_H to be the difference of the allele positions of the first and last defined bits of H . The order of H , indicated with $o(H)$ is the number of defined bits. For example, the schema

01***100**

has 5 defined bits and so has order $o(H) = 5$. The last defined bit is in position 8 and the first is in position 1, and so the schema has length $l_H = 8 - 1 = 7$.

The Schema Theorem describes how schemata featuring in highly fit chromosomes have a greater expectation of propagating through successive populations as a GA evolves and is stated as follows.

Theorem (Schema). *Let H be a schema and let $m_H(i)$ be the number of chromosomes belonging to H present in population i of an evolving GA. Then the expectation of the number of chromosomes belonging to H in population $i + 1$, denoted $\langle m_H(i + 1) \rangle$, is given by the formula*

$$\langle m_H(i + 1) \rangle = F_H(i) \cdot m_H(i) \left[1 - p_c \frac{l_H}{l - 1} \right] (1 - p_m)^{o(H)} \quad (2.1)$$

where $F_H(i)$ is the relative fitness of H , defined to be the average fitness of all those chromosomes in the population belonging to H divided by the average fitness of all chromosomes in the population; p_c is the crossover probability; p_m is the mutation probability.

The formula assumes that the GA uses fitness proportional selection and takes into account the possibility that the genetic operators of crossover and mutation can act to disrupt schema H . It is worth noting that the formula does not include a term for new instances of schema H being introduced to the population by genetic operators and so really gives a lower bound for the expectation.

The Schema Theorem allows one to reason about the ways in which particular patterns are likely to propagate and so gain an understanding about how the performance of a particular GA is affected by the various design choices one must make. The theorem shows that, all else being equal, those schemata with relative fitness greater than 1 will be increasingly represented in the successor population, whereas those with relative fitness less than 1 will be decreasingly represented. This is purely the effect of selective pressure implemented in the fitness proportional selection process. However, in the presence of genetic operators, other factors in the formula can become very significant. In particular, as

the length of H becomes close to l_1 , the term relating to crossover, $1-p_c \frac{l_H}{l-1}$, approaches $1-p_c$. In other words, schemata with a length approaching the length of the chromosome are almost certain to be disrupted by crossover when it occurs. Also, as the order of H increases, the term relating to mutation, $(1-p_m)^{o(H)}$, decreases exponentially. In other words, schemata involving many defined bits are quite likely to be disrupted by mutation when it occurs.

Consequently, we can conclude that short length, low order, above average fit schemata are increasingly sampled in subsequent generations. Such schemata are known as building blocks and can be thought of as “partial solutions” in those chromosomes containing them tend to be of higher fitness. Intuitively, a chromosome containing many building blocks will be a near-optimal solution.

2.4.2 The Building Block Hypothesis

The Building Block Hypothesis (BBH) states that GAs achieve near-optimal performance by preferentially selecting and juxtaposing building blocks. Ideally, the notion of building blocks should be of particular utility when choosing an encoding for a GA. Knowledge of the application domain could be used to ensure that the values of factors known to interact are encoded in neighbouring loci. Also, genetic operators could be refined to disrupt a chromosome only at locations where problem-specific knowledge indicates that it makes sense, thereby increasing the probability that the evolutionary process will fruitfully juxtapose building blocks and quickly arrive at an optimal solution.

In practice, however, problem knowledge is rarely sufficient to do this. Usually, GAs are applied to problems where solution sets are very poorly understood. There is no guarantee that, for any given problem, an encoding can be found, or even exists, that contains building blocks. Moreover, it is known that, even where building blocks do exist, the BBH breaks down for many theoretical and practical problems because other factors impede the evolution.

2.5 Why choosing the GA method

Among a virtually infinite range of problems, here the main focus is about parameters estimation and optimization of objective functionals representing therapy schedules. A cancer therapy schedule is usually subjected to many restriction (e.g. limit on dose due to its toxicity).

The distinctive attributes that make genetic algorithms potentially suitable for dealing with problems intractable to many traditional mathematical methods of optimization of a certain objective functional may be summarised as follows [14]:

- GAs implement multidirectional search by sustaining a population of candidate solutions;
- GAs explore the search space using stochastic processes rather than deterministic rules.
- GAs exploit the valuable information obtained so far by being biased towards the selection of "good" solutions.
- GAs require very little from the objective functional. The objective functional must unambiguously define the payoff of each solution, but may be multimodal, discontinuous and may allow a certain degree of imprecision (GAs are "robust").

Multidirectional search has two advantages. Firstly, it obviously increases the efficiency of the search by looking for the optimum in many directions simultaneously. Secondly, if the search in one direction gets stuck at a local optimum, there is still a chance to find the absolute optimum by approaching it from another direction. In addition to that, 'stochastic wandering' through the search space unburdens Genetic Algorithms from strong dependence on additional properties of the objective functional such as the gradient.

The last advantage, i.e. the robustness of GAs with respect to the type of the objective functional, raises Genetic Algorithms to the position of a very versatile method, which can be applied to a large number of real life optimisation tasks. This versatility can be explained by the fact that GAs are not concerned with finding the optimal solution *per se*. Instead, their main goal is to find better solutions than those that are already known.

An underlying purpose of this thesis is that GAs may have a role to play in supplementing traditional approaches to parameter estimation.

Chapter 3

Monoclonal description of recurrent high-grade glioma

In this chapter, a model used to describe the growth dynamics of glioma is developed. Starting from a preselected database of longitudinal tumour volumes, Genetic Algorithms (GAs) and a numerical optimization approach are used to explore several aspects about the model as well as consequences of choosing a particular fractionation protocol for the radiation therapy.

In particular, the focus is on recurrent high-grade glioma (rHGG), which remains incurable with inevitable evolution of resistance and high inter-patient heterogeneity in time to progression (TTP), namely the length of time from the date of diagnosis or the start of treatment for a disease until the disease starts to get worse or spread to other parts of the body. In a clinical trial, measuring the time to progression is one way to see how well a new treatment works. Here, we evaluate if early tumour volume response dynamics can calibrate a mathematical model to predict patient-specific resistance to develop opportunities for treatment adaptation for patients with a high risk of progression.

Predicting progression prior to radiographic manifestation would allow clinicians to modify therapy before selection for and proliferation of treatment-resistant tumour subpopulations.

Patients with rHGG included in this modelling study (16 patients in total) were treated at the Moffitt Cancer Center (Tampa, Florida), between August 2015 and March 2018 as part of a phase I clinical trial (NCT02313272, 05/12/2014) [3].

The reference dataset consists of a total of 95 T1-weighted contrast-enhanced (T1post) magnetic resonance images (MRIs) from those patients treated in a phase I clinical trial with hypo-fractionated stereotactic radiation (HFSRT; 6

Gy x 5) plus a standard chemotherapy protocol consisting of: pembrolizumab (100 or 200 mg, every 3 weeks) and bevacizumab (10 mg/kg, every 2 weeks).

Without getting into details about the selection of patient cohort, it is however important to stress that patients with no recurrency (namely those who didn't show tumour regrowth) are excluded from this study since the standard protocol achieved the desired result, making unnecessary further considerations about deviating from the adopted strategy.

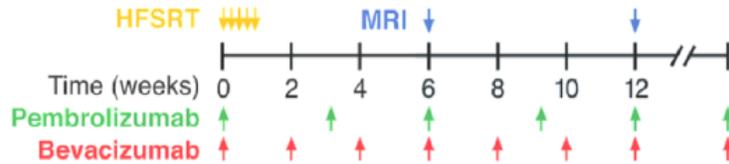


Figure 3.1: Schedule of the adopted HFSRT

3.1 The mathematical model

The aim of this study is to provide a simple mathematical framework to: 1) fit the observed tumor growth response data to HFSRT given in five daily fractions, and, based on this description, to 2) simulate intermittent radiation treatment (IRT) schedules. The presented model captures only the key mechanisms of treatment response to limit the mathematical complexity of the model to be able to obtain high confidence fit parameters estimates.

Glazar et al. [8] suggest a simple tumour growth inhibition (TGI) model to describe tumour volume dynamics $V_s(t)$:

$$\frac{dV_s(t)}{dt} = \lambda V_s(t) - \gamma(t)V_s(t) \quad (3.1)$$

where λ [day^{-1}] is the net tumour growth rate in the absence of therapy and $\gamma(t)$ [day^{-1}] is the rate at which the tumour volume decays in response to therapy. An exponential growth is assumed, as the tumour volume is likely far from carrying capacity after surgery and HFSRT, supported by observed dynamics.

Bevacizumab and pembrolizumab were administered every 2 and 3 weeks until progression; However, an approximated continuous and constant drug concentration, ignoring pharmacokinetics and pharmacodynamics, is assumed for simplicity. To simulate the evolution of resistance to therapy, is assumed the decay

rate to be exponentially declining with time, such that:

$$\frac{d\gamma(t)}{dt} = -\epsilon\gamma(t) \quad (3.2)$$

where ϵ [day^{-1}] is the rate at which resistance develops.

The analytic solution of the coupled system of Equations (3.1) and (3.2) is:

$$V_s(t) = V_{s,0} \cdot \exp \left[\lambda(t - t_0) + \frac{1}{\epsilon}(\gamma(t) - \gamma_0) \right] \quad (3.3)$$

$$\gamma(t) = \gamma_0 \cdot \exp[-\epsilon(t - t_0)] \quad (3.4)$$

with initial conditions $[V_0, \gamma_0]$ evaluated at $t = t_0$.

Brüningk et al. [3] extended the present model in order to account for the contribution of HFSRT to tumour volume reduction. To model radiotherapy effects, at each treatment fraction delivery (t_{RT}), a proportion of $(1 - S)$ of the viable tumour (V_s) is transferred to a dying compartment, V_d , i.e. tumour cells incapable of further proliferate due to the radiation effect: when attempting to divide, these cells die (mitotic catastrophe). Thus it makes sense to assume this mechanism having a time scale of λ^{-1} . The surviving fraction S is here used as a model parameter in itself, rather than as a function of radiation dose and patient-specific radiation sensitivity, as described by the linear-quadratic model, thus minimizing the modelling complexity of radiation therapy.

$$V_s(t_{RT}^+) = S \cdot V_s(t_{RT}^-) \quad (3.5)$$

$$V_d(t_{RT}^+) = V_d(t_{RT}^-) + (1 - S) \cdot V_s(t_{RT}^-) \quad (3.6)$$

Here, t_{RT}^- denotes time immediately before delivery of a radiation fraction, t_{RT}^+ the time immediately after treatment delivery. Observe that the presented model provides a worst-case estimate of no explicit consideration of radiation-induced immune stimulation.

We model radiation induced cell death as mitotic catastrophe which is a proliferation-dependent process. Hence, we describe the volume change as an exponential reduction of $V_d(t)$. The reduction of $V_d(t)$ is assumed to be at rate identical to

the growth rate to restrict the number of free parameters. As anticipated, this assumption is motivated on the possibility of cell death upon the attempt of cell division:

$$\frac{dV_d(t)}{dt} = -\lambda V_d(t) \quad (3.7)$$

whose solution is:

$$V_d(t) = V_{d,0} \cdot \exp[-\lambda(t - t_0)] \quad (3.8)$$

Hence, the total, observed tumor volume $V(t)$ comprises a proliferating $V_s(t)$ and dying $V_d(t)$ population.

$$V(t) = V_s(t) + V_d(t), V(0) = V_0 \quad (3.9)$$

In conclusion, this model comprises five parameters $(V_0, \lambda, \gamma_0, \epsilon, S)$.

At the very beginning, $\lambda = \gamma_0$ is assumed. Looking at (3.1), it is possible to straightforwardly realize the meaning of this pessimistic assumption: chemotherapy can only slow down the exponential growth of viable tumour, at most stopping it when $\gamma = \lambda$.

In [3], λ is considered to be uniform over the whole population. Here λ is assumed to vary for each patient in a certain range, the latter obtained from previous literature (as well as in the case of ϵ), in order to move toward a more personalized clinical picture.

In contrast, while in [3] a 30% deviation from the original V_0 in the fitting stage, here V_0 is fixed as the value provided from the first MRI scan for each patient, thus introducing a constraint in the fitting and reducing the space of feasible solutions and, consequently, shortening the computation time of the applied GA. It sounds reasonable since treatment strategies are decided taking into account the very first measured volume.

A last remark about the pre-treatment behaviour of the model is necessary: the tumour growth is malthusian for $t < t_0$, according to:

$$\frac{dV_s(t)}{dt} = \lambda V_s(t) \Rightarrow V_s(t) = V_0 \cdot e^{\lambda t} \text{ for } 0 \leq t < t_0$$

where $t = 0$ is the time of the first volume measure.

Parameter	Unit	Meaning	Fit bounds
λ	day ⁻¹	Tumor growth rate in the absence of treatment	[0.017, 0.14]
ε	day ⁻¹	Evolution of resistance rate	[0, 0.1]
S		Radiation surviving fraction	[0, 1]

Figure 3.2: Patient specific model parameters to be fitted and related boundaries

3.2 Linear Stability Analysis

Before running computer simulations to find model's parameters (*quantitative information*), it is customary to perform a Linear Stability Analysis [15] to get some insight into the system stability and long-term behaviour (*qualitative information*).

The presented model, after the instantaneous action of radiotherapy described in (3.5) and (3.6), consists of 3 ODEs: (3.1), (3.2), (3.7). (3.7) is decoupled from the rest, therefore representing a one-dimensional dynamical system:

$$\dot{V}_d = f(V_d) = -\lambda \cdot V_d \quad (3.10)$$

By imposing the steady-state condition $\dot{V}_d = 0$ we find the unique fixed point $V_d^* = 0$. Since it is $\lambda > 0$, the sign of the derivative is always negative and from the phase portrait in Figure 3.3 it is straightforward to see that every trajectory, determined by the initial condition $V_{d,0}$, move towards $V_d^* = 0$, which turns out to be a stable fixed point.

Physically, this simple behaviour makes sense since V_d represents doomed cells.

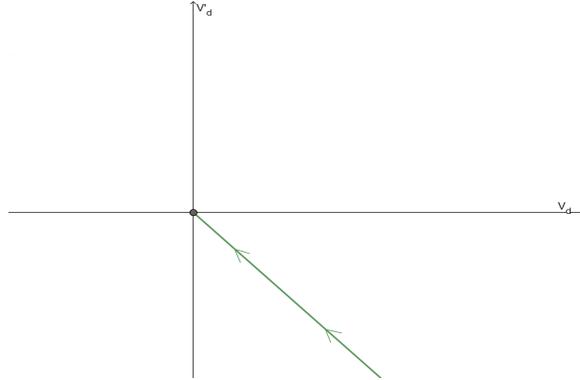


Figure 3.3: phase portrait of (3.10)

Equations (3.1) and (3.2) form a two-dimensional non-linear system of coupled ODEs, which can be rewritten in the compact form:

$$\dot{V}_s(t) = f(V_s, \gamma) \quad (3.11)$$

$$\dot{\gamma}(t) = g(V_s, \gamma) \quad (3.12)$$

The unique fixed point is the origin of the phase space: $(V(t)^*, \gamma^*) = (0, 0)$. Linearization procedure around this fixed point leads to the following Jacobian matrix of the system

$$J(V_s, \gamma) = \begin{bmatrix} \frac{\partial f}{\partial V_s} & \frac{\partial f}{\partial \gamma} \\ \frac{\partial g}{\partial V_s} & \frac{\partial g}{\partial \gamma} \end{bmatrix} = \begin{bmatrix} \lambda - \gamma & -1 \\ 0 & -\epsilon \end{bmatrix}$$

which must be evaluated in the origin:

$$J(0,0) = \begin{bmatrix} \lambda & -1 \\ 0 & -\epsilon \end{bmatrix}$$

Being two-dimensional, the linear behaviour of the system around the fixed point is totally determined by trace τ (sum of Jacobian eigenvalues) and determinant Δ (product of Jacobian eigenvalues) of $J(0, 0)$:

$$\tau = \lambda - \epsilon \Delta = -\lambda \epsilon < 0$$

According to Figure 3.4, this is a saddle-point, thus we do expect an unstable behaviour of the system. This is in agree with the analytical solution of the system, which in this simple case is known: the viable tumour volume tends to evolve exponentially, moving away form the fixed point, the latter representing a non-sense case, since it describe a situation where cancer is absent and no chemotherapy is applied.

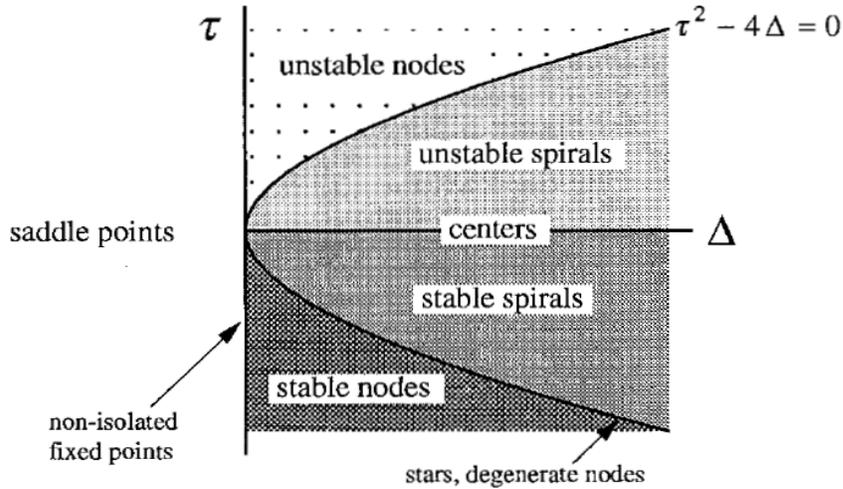


Figure 3.4: Trace-determinant and related fixed point stability

3.3 Parameters fitting through GA

Once the model has been defined, the fitting stage took place in order to determine patient specific parameters from which further considerations will be made. As anticipated in chapter 2, a Genetic Algorithm inspired on the basis of the canonical one has been employed in parameters estimation (see Appendix A). Tuning the parameters of the GA has been far from an easy task, since in this sense no precompiled scheme exists. Moreover, GA parameters (such as crossover and mutation rates) must be set specifically for every patient and simulation in order to drive the algorithm toward an optimal solution. However, here (as well as in the following) details about the laborious implementation of the GA will be neglected. Nevertheless, it should be useful to illustrate few features, starting on how parameters have been encoded in the chromosome-like representation of the algorithm, giving an idea about the size of search space, which reflects the limited numbers of digits used to represent parameters.

So, each parameter could take, in the relevant individual ranges previously shown in Figure 3.5, $2^7 = 128$ discrete values. Each chromosome represents one of $2^{21} = 2'097'152$ possible solutions of the search space.

Concerning the adopted fitness function, following the work path of [3], the fitting goodness of every output solution proposed by the GA is assessed through the root mean squared error (RMSE), defined as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (V_{measured}(t_i) - V_{sim}(t_i))^2}{N}} \quad (3.13)$$

Where $V_{measured}(t)$ is the measured value of tumour volume at a certain time t through the MRI imaging, N is the total number of those available measures (it varies from patient to patient) and $V_{sim}(t)$ is the simulated tumour volume at time t , related to the particular choice of parameters in the solution.

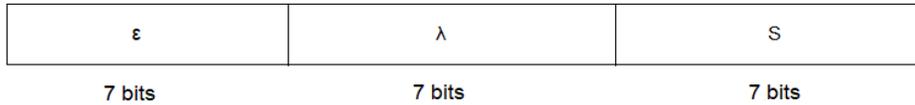


Figure 3.5: Partition of the bit-string chromosome during the GA fitting

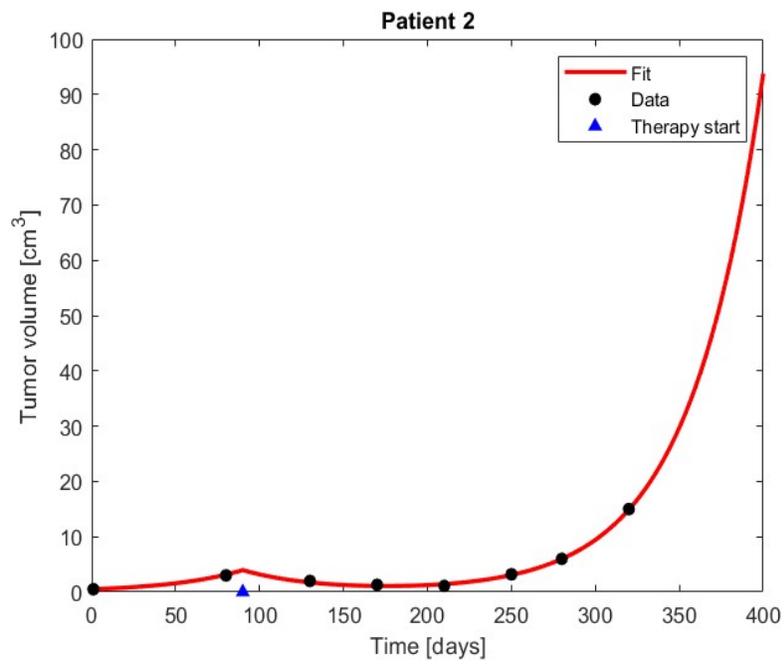
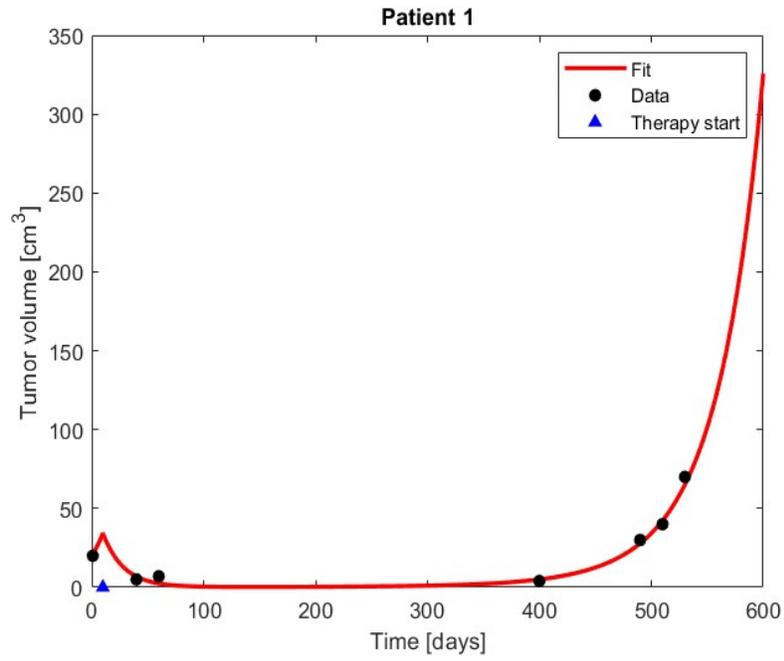
3.3.1 Fitting results

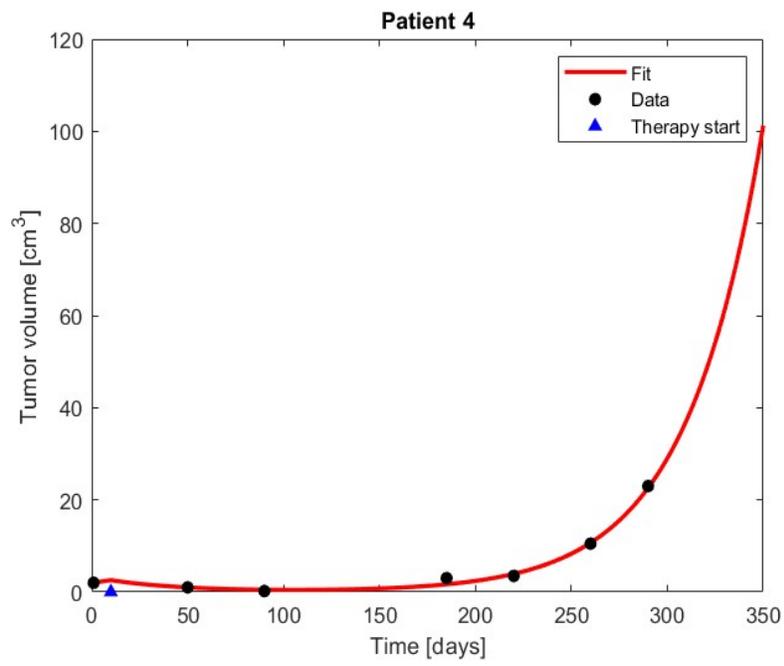
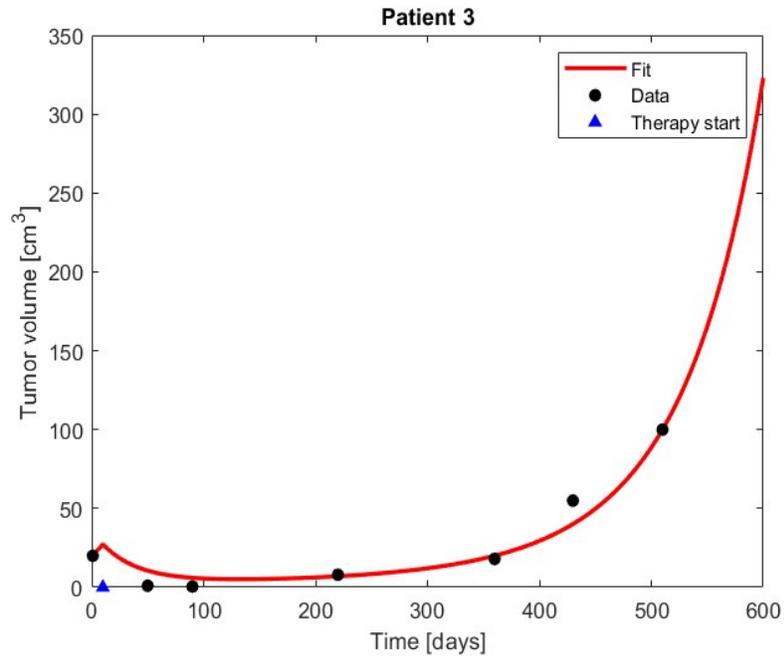
Fitting quality is at least slightly better than the original one (where a deterministic non-linear programming technique was used) in every case, except for patient 11. It is however evident that patient 11 shows an anomalous behaviour in comparison to the remaining patient cohort; this is an indication that other model external variables strongly influenced the evolution of the response. This is not a surprise: it's important to stress that cancer kinetics is extremely complex and, in the case of patient 11 whose tumour volume is relatively small, it may not be misleading to hypothesize that the stochastic component abruptly emerges.

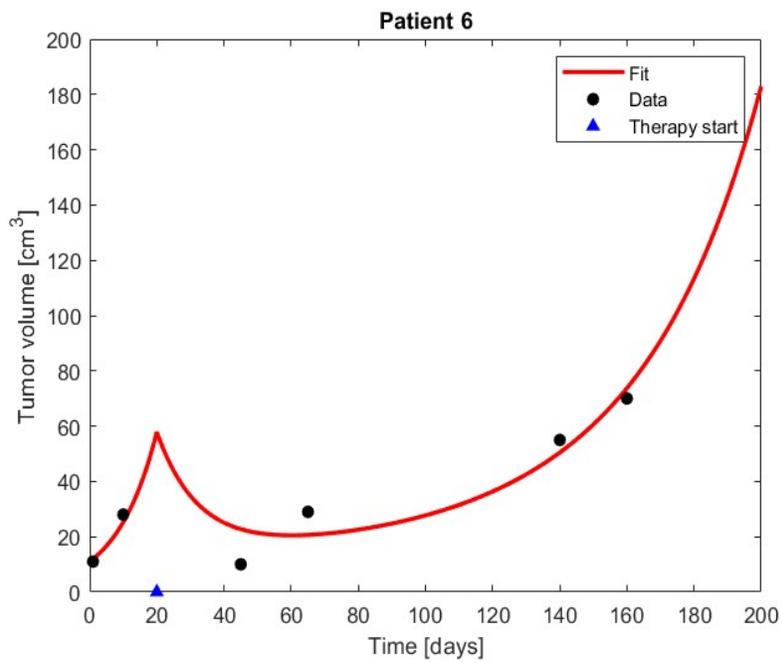
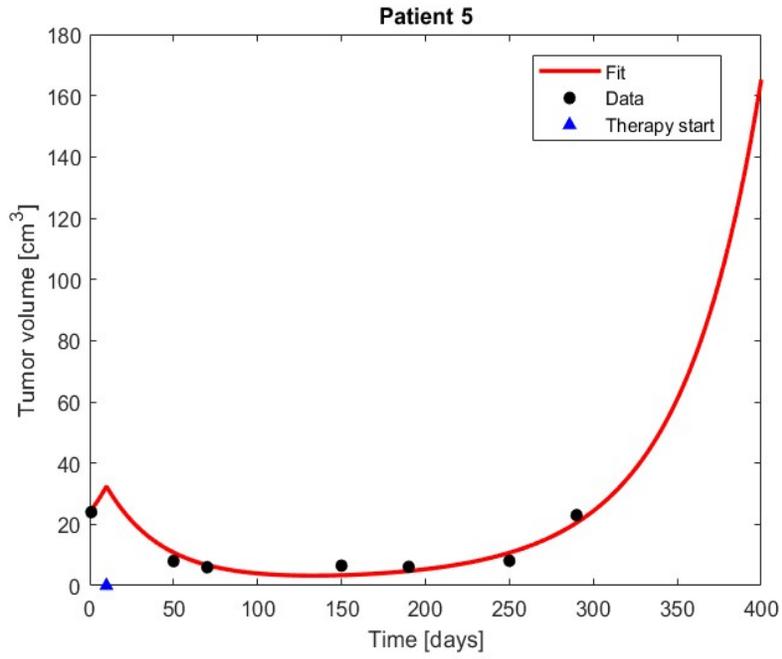
Parameters' fitting values and fitting comparison between the proposed model and the one in [3] are reported in the tables below, followed by the GA fitting plots.

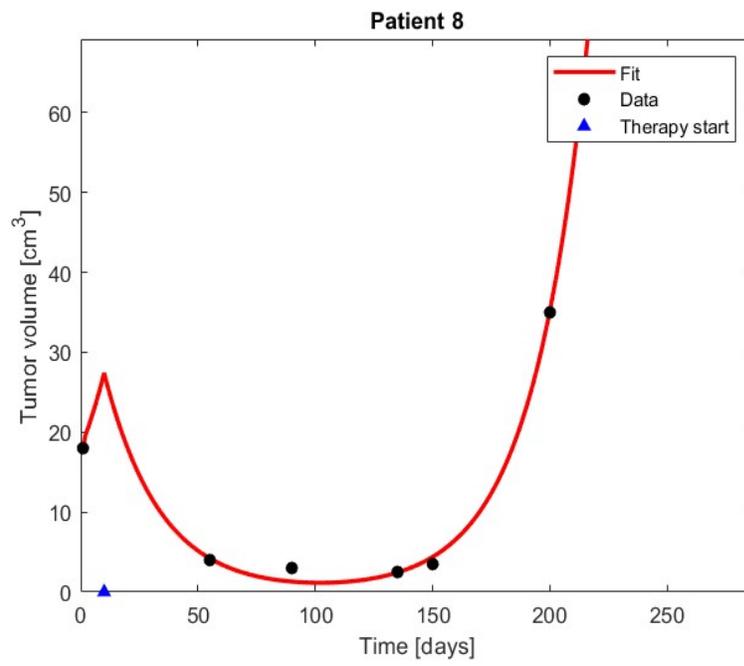
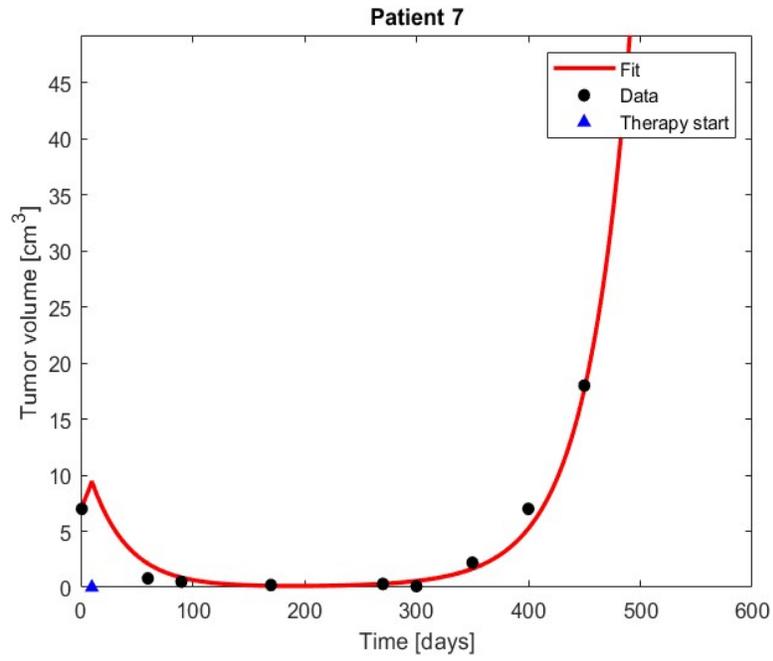
Patient	λ	ϵ	S
1	0.054	0.001	0.33
2	0.023	0.046	0.5
3	0.031	0.001	0.67
4	0.025	0.068	0.41
5	0.03	0.003	0.53
6	0.083	0.002	0.78
7	0.03	0.004	0.28
8	0.042	0.071	0.24
9	0.044	0.004	0.48
10	0.055	0.001	0.5
11	0.062	0.018	0.02
12	0.101	0.015	0.2
13	0.043	0.005	0.47
14	0.107	0.005	0.21
15	0.082	0.002	0.56
16	0.051	0.124	0.06

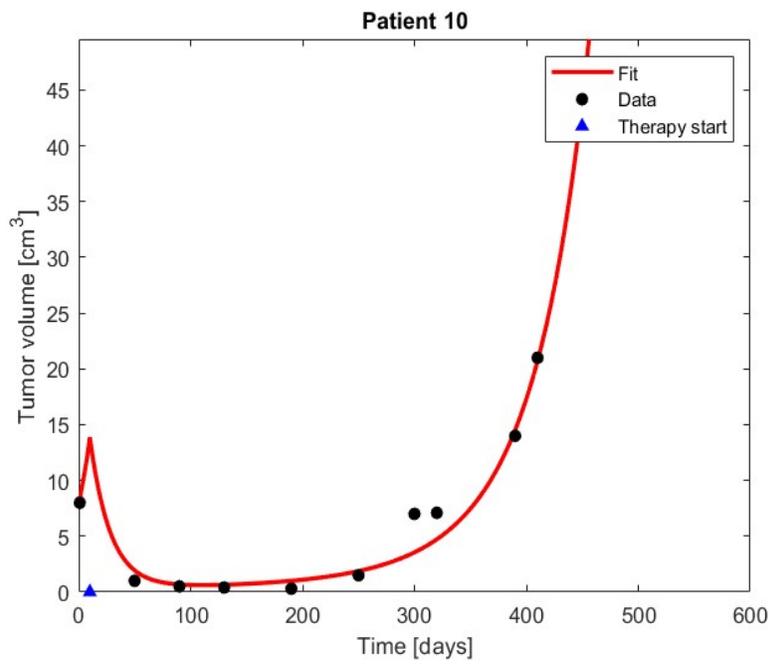
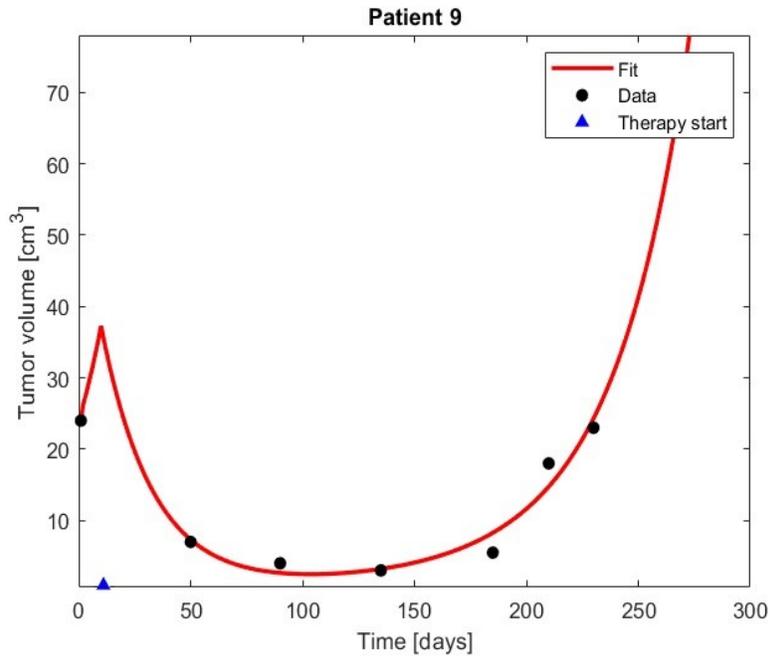
Patient	Original fitting RMSE	New fitting RMSE
1	9.7	2.9
2	1.6	0.2
3	7.2	7.0
4	1.0	0.6
5	2.5	2.1
6	8.5	6.8
7	2.4	0.8
8	2.9	0.8
9	3.3	1.8
10	3.1	1.4
11	0.6	1.4
12	1.3	0.2
13	0.2	0.1
14	0.2	0.08
15	0.3	0.1
16	0.5	0.2

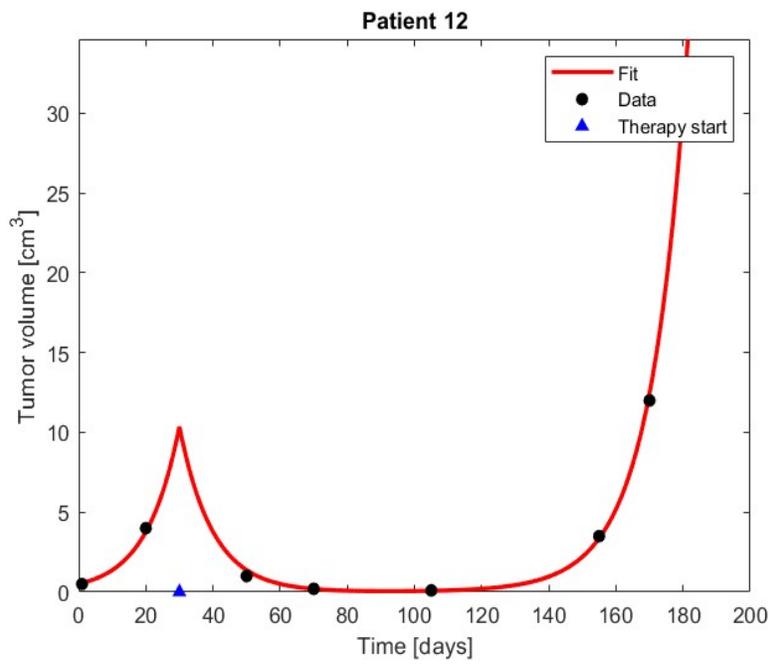
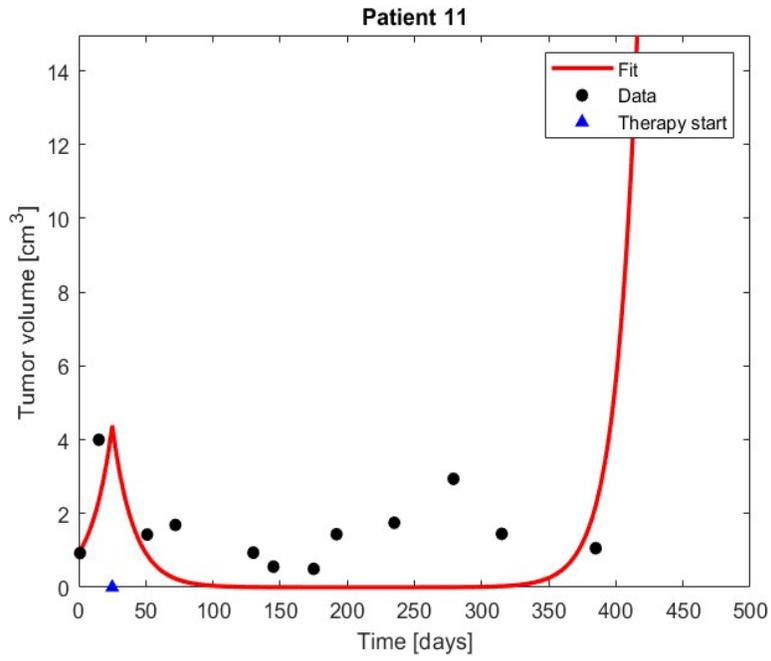


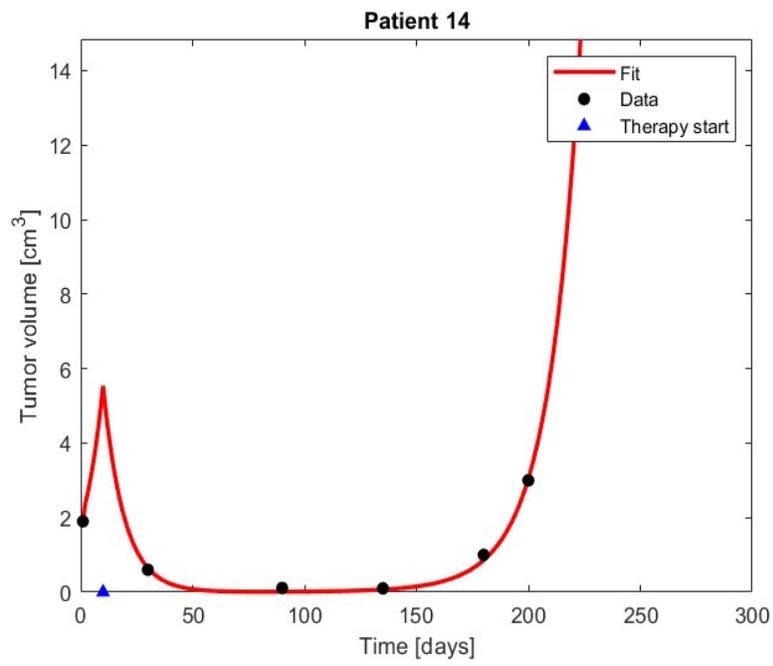
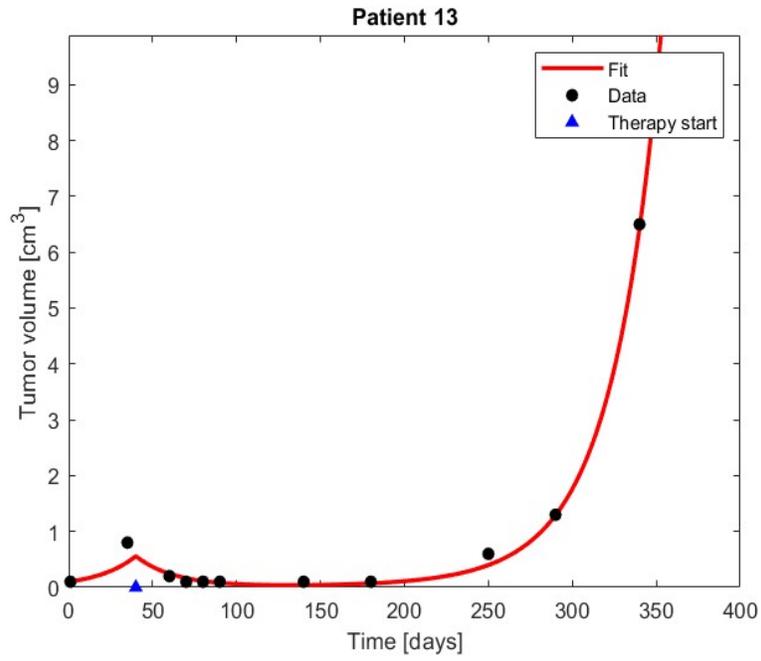


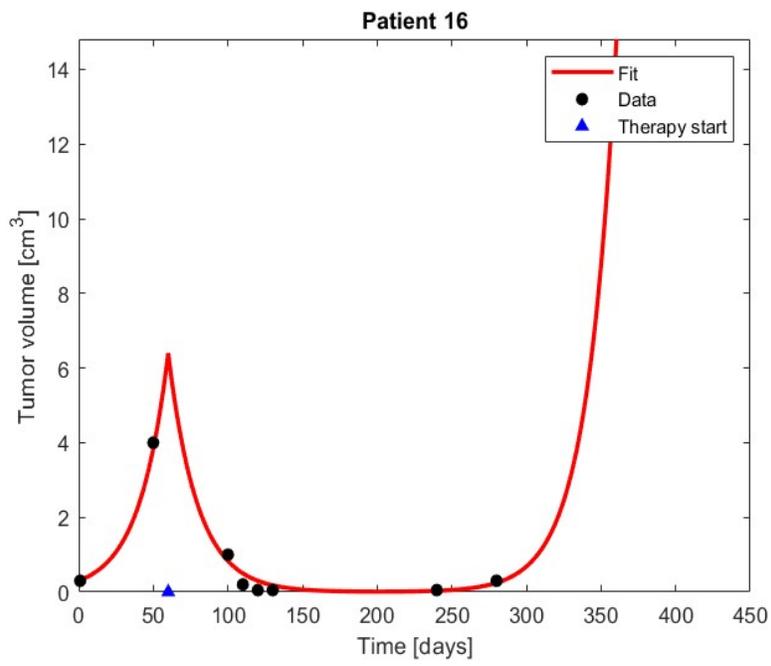
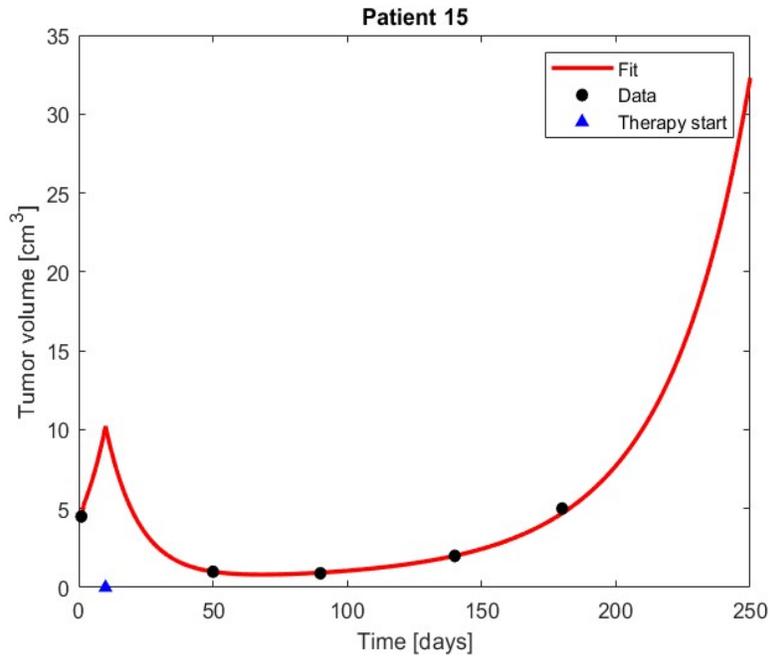












3.4 Intermittent Radiotherapy Protocol

The core of this thesis is to propose different radiotherapy schedules whose construction is based on each single patient features, aiming to improve the final outcome. Assuming that patient therapy response is based on the previously fitted parameters over the standard hypofractionated protocol, here the same amount of dose split in 5 fractions is administered in a longer time range (“intermittently”).

In [3], the same approach is proposed, but with a rather rigid time schedule: the interval between a dose and the next is fixed (4, 6, 8, 10 weeks), so for each patient is possible to explore only 4 possible alternative treatments.

Here a more extended space of alternatives is explored, to find the optimal intermittent timing by means of a Genetic Algorithm (see Appendix B). The fitness of a solution is simply the related patient survival time (PST). A small remark about how the PST has been defined: the available dataset usually reports a final measure of tumour volume (*cut-off volume*) which is significantly higher than the previous ones, suggesting that in the worst-case scenario that measure represents the maximal tumour burden tolerated by the patient before retiring from the therapy for some reason (e.g. decease or ineffectiveness of the treatment from then on). For a small but consistent group of patients, the final measure is small compared to the previous ones, thus the assumption about the PST definition becomes misleading. This can be due to many reasons, for example the emergence of side effects which led the patient to quit the therapy prematurely. To fix the problem, in these cases the PST is replaced by the *time to progression* of tumour (TTP) which can be non-rigorously defined as the time at which the predicted volume is twice the volume of the last measure preceding the therapy. In the following, data analyses and general considerations will neglect the distinction between PST and TTP.

The objective functional representing the PST is:

$$PST(\mathbf{c}) = \int_{t_0}^{t_{cut-off}(\mathbf{c})} d\tau \quad (3.14)$$

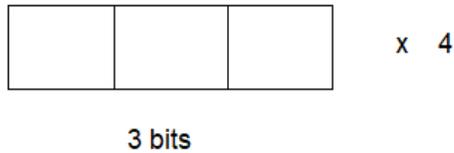
Which has to be maximized. The control vector is represented by the available time intervals between doses:

$$\mathbf{c} = [c_1, c_2, c_3, c_4, c_5] \text{ where } c_i \in \{4, 5, 6, 7, 8, 9, 10 \text{ weeks}\}$$

The intervals are not fixed anymore, so for example a possible outcome for the therapy schedule could be:

1 st dose
wait 5 weeks
2 nd dose
wait 10 weeks
3 rd dose
wait 4 weeks
4 th dose
wait 7 weeks
5 th dose

In practice, the GA encodes a potential time schedule in bit chromosomes:



A chromosome is made by 12 bits, where the first 3 bits specifies the time interval between the 1st dose and 2nd, the second group of 3 bits specifies the time between 2nd and 3rd dose and so on. The space of alternative time schedules is $2^{15} = 32768$, which definitely allows a higher grade of therapy personalization compared to the original 4 allowed choices.

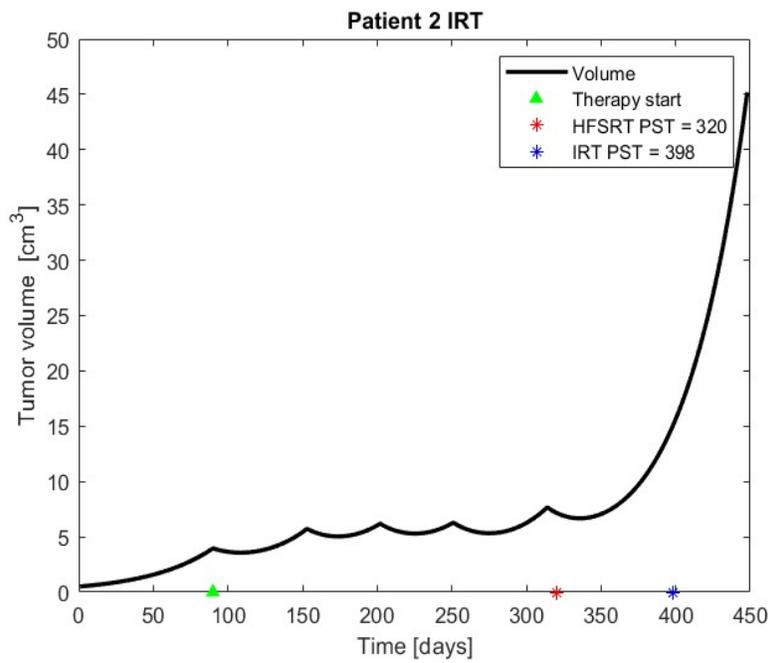
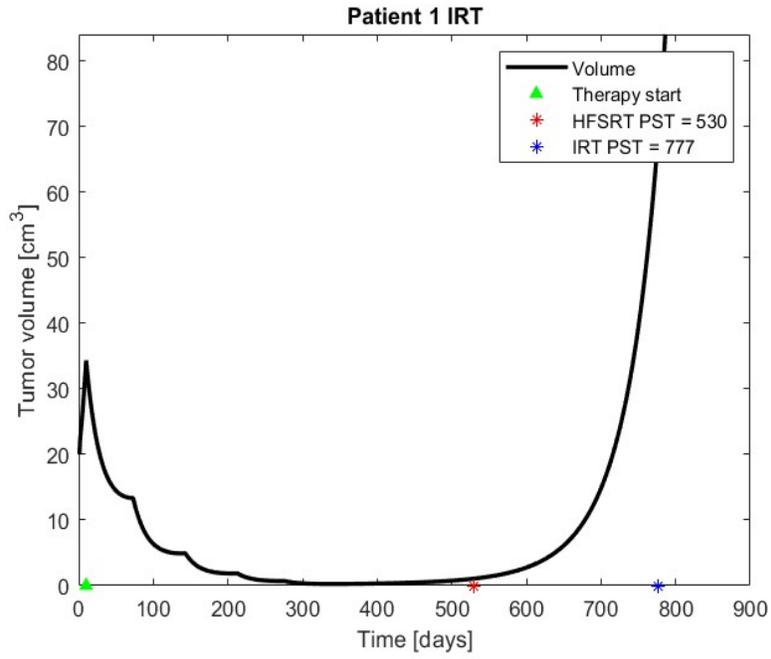
It is also important to point out that the idea of “optimal timing” is strictly related to the specific therapy goal, which is not necessarily restricted to the PST maximization. For example, the versatility of the intermittent dosage allows to modify the schedule in order to reduce the average tumour burden or to take into account emerging undesired side effects of the ongoing therapy.

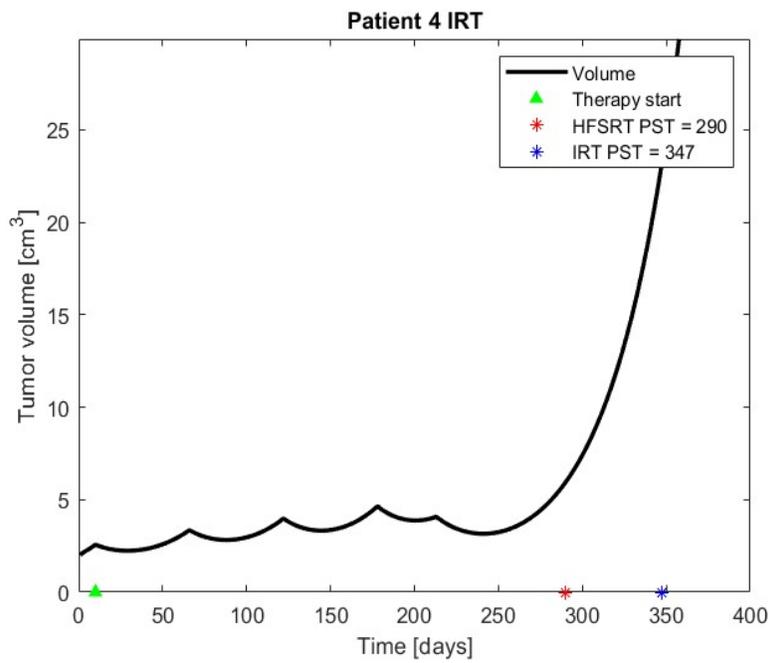
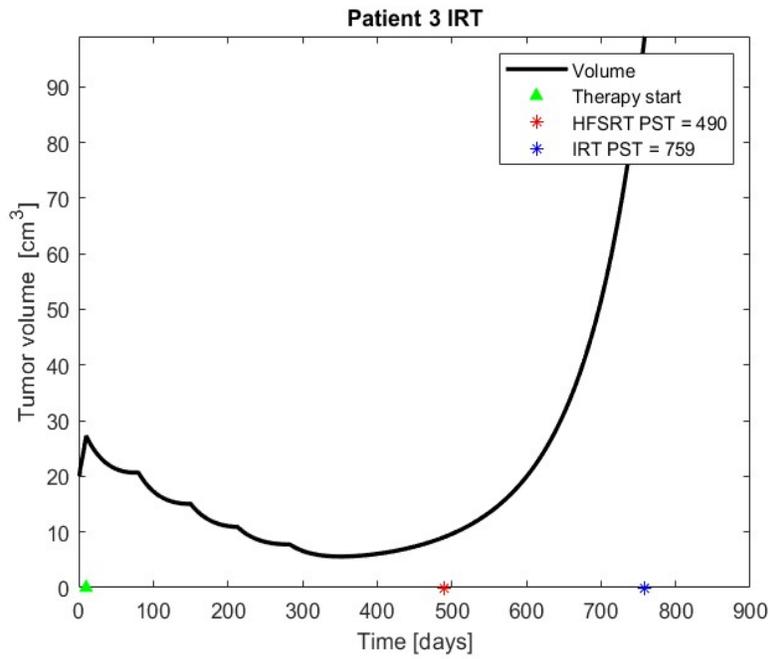
3.4.1 Results

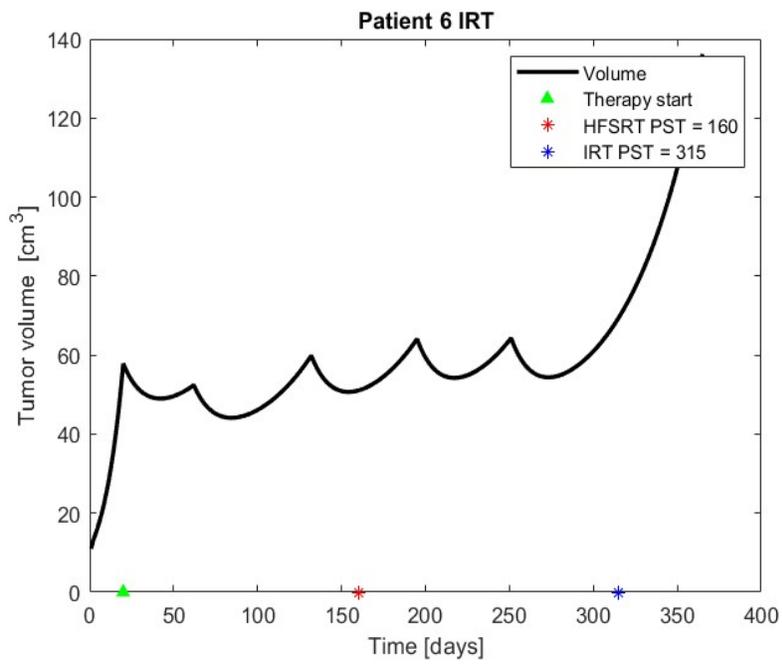
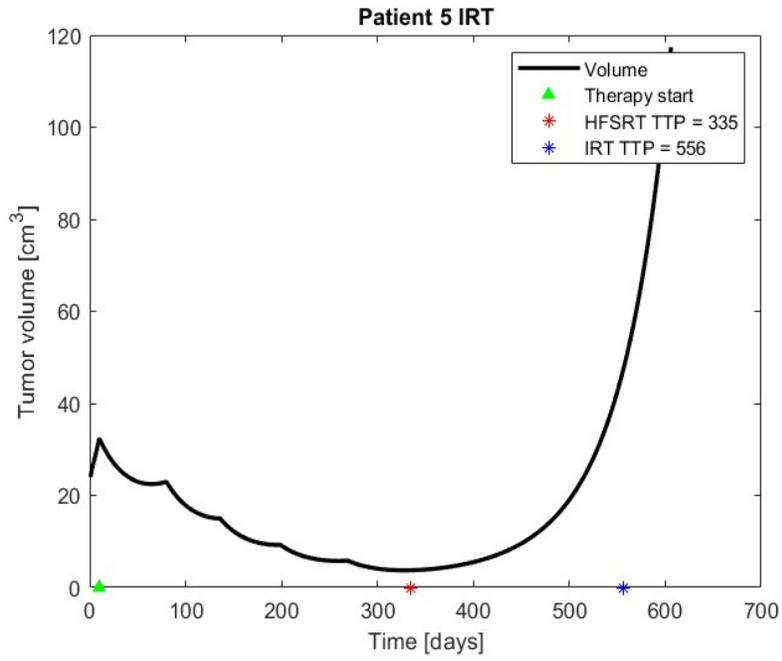
In [3], the intermittent scheme outperforms the HFSTR in 11/16 cases. Here, intermittent radiotherapy is predicted to be by far a better choice over HFSRT in every case. This means that the higher personalization grade of both model parameters and therapy makes a crucial difference in the final outcome.

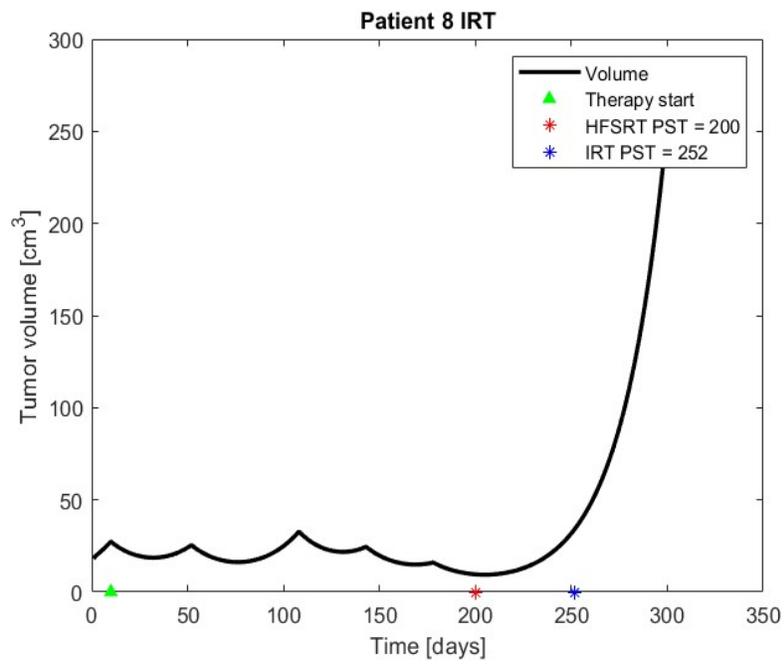
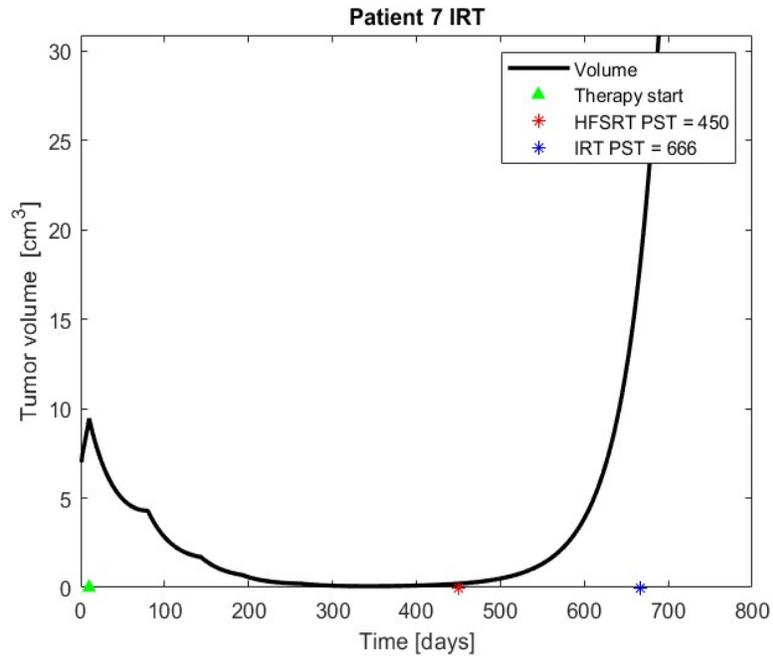
Computer simulations of best GA outcomes are reported below, together with the corresponding IRT schedules.

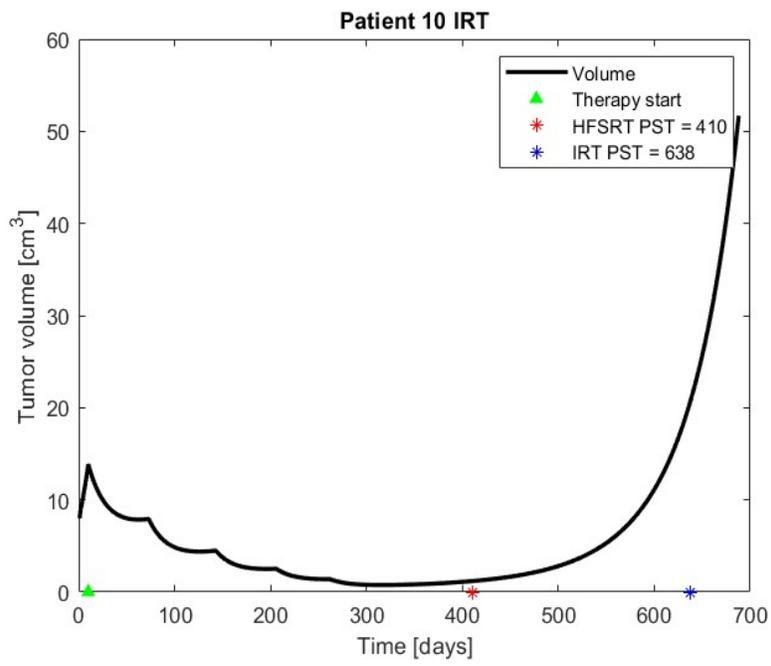
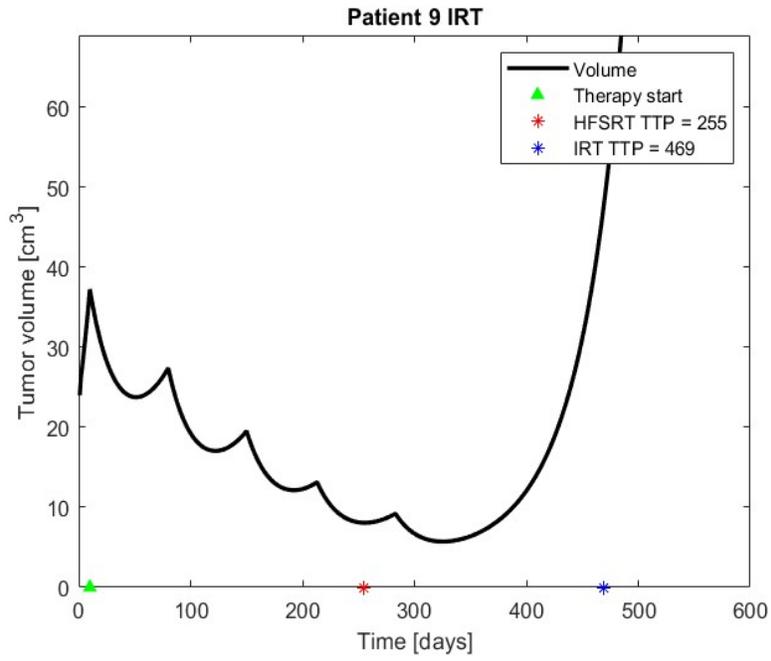
Patient	Predicted IRT timetable (days from 1 st MRI scan)				
1	10	73	143	213	276
2	90	153	202	251	314
3	10	80	150	213	283
4	10	66	122	178	213
5	10	80	136	199	269
6	20	62	132	195	251
7	10	80	143	192	262
8	10	52	108	143	178
9	10	80	150	213	283
10	10	73	143	206	262
11	25	88	158	228	298
12	30	72	135	184	226
13	40	96	159	229	299
14	10	66	129	192	248
15	10	73	129	199	248
16	60	109	158	193	221

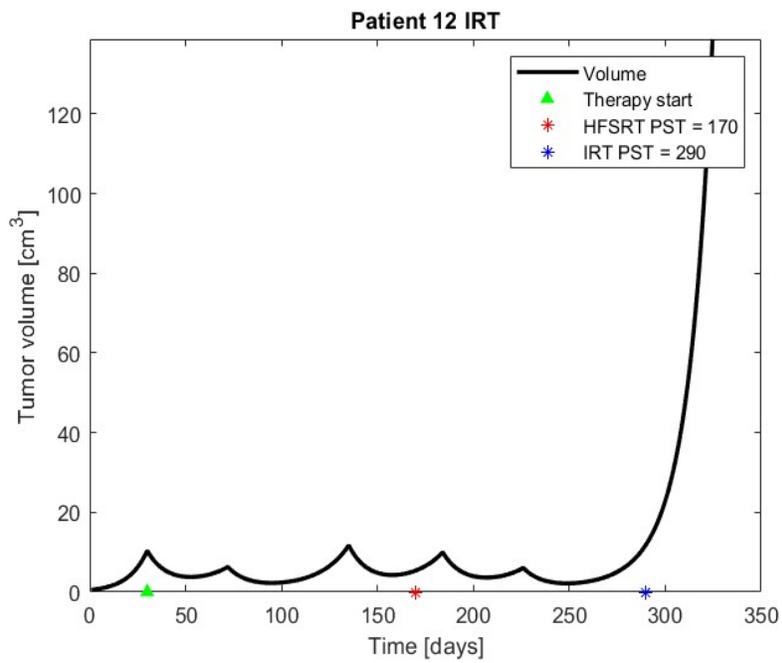
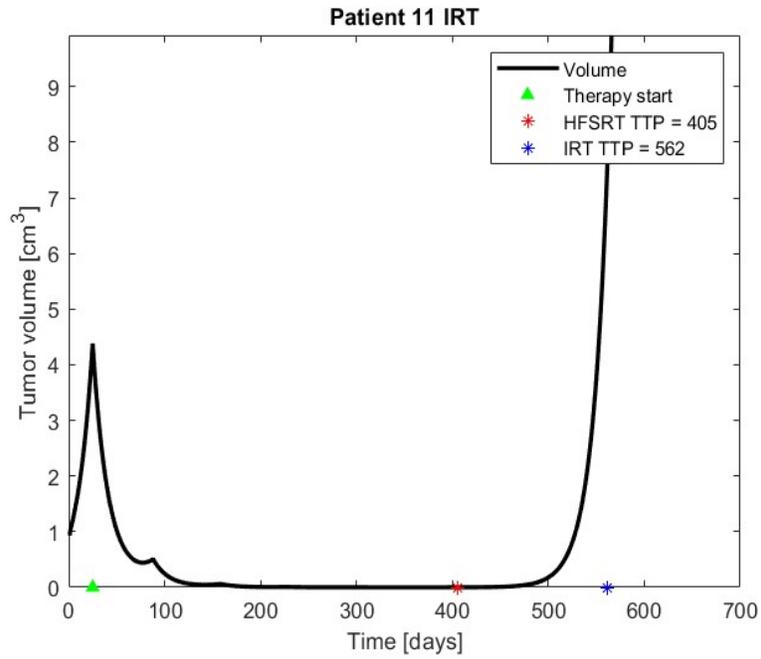


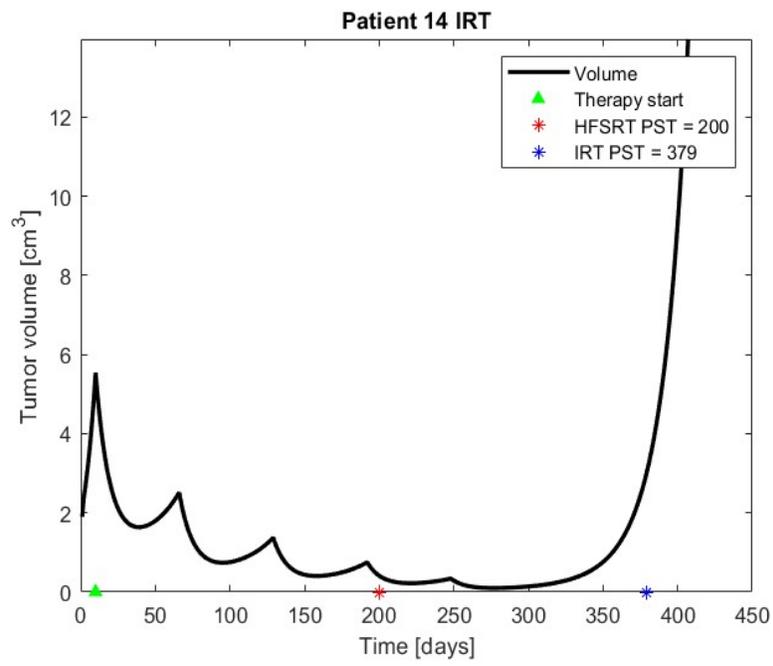
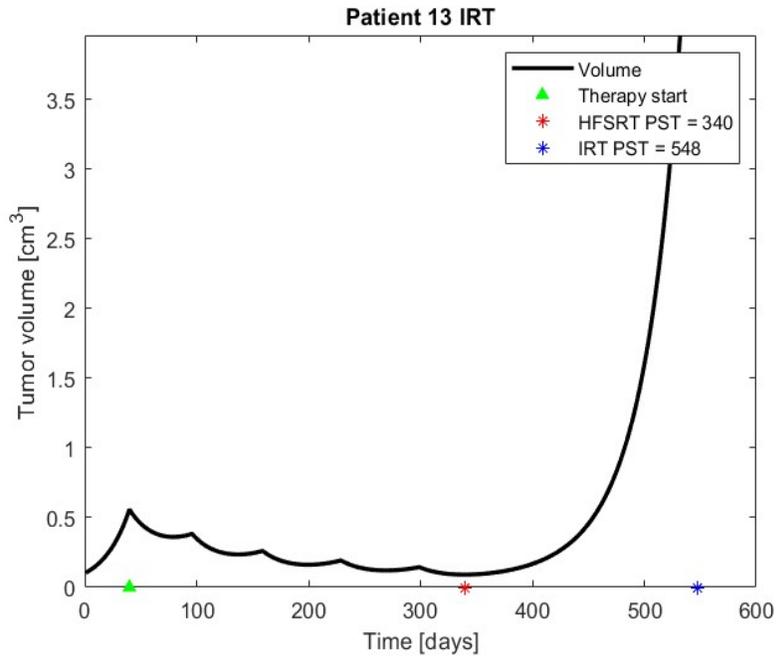


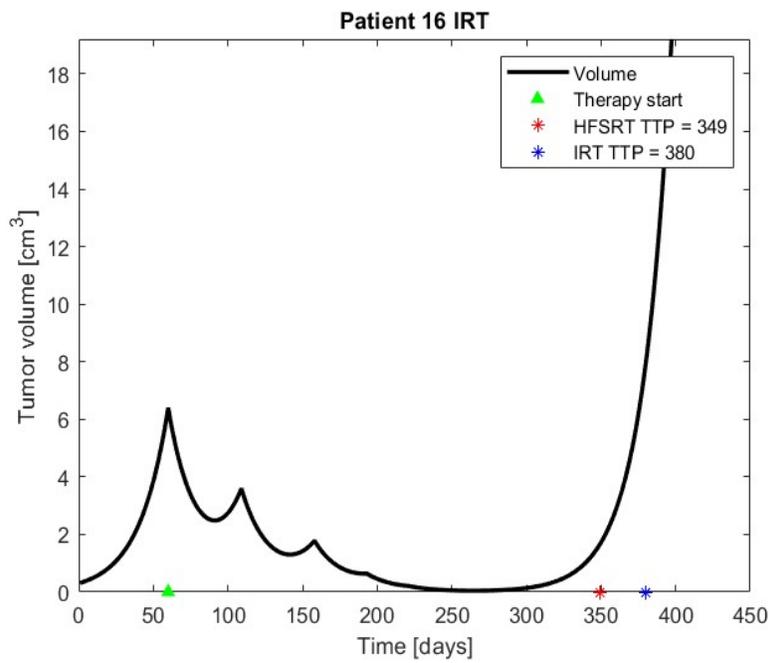
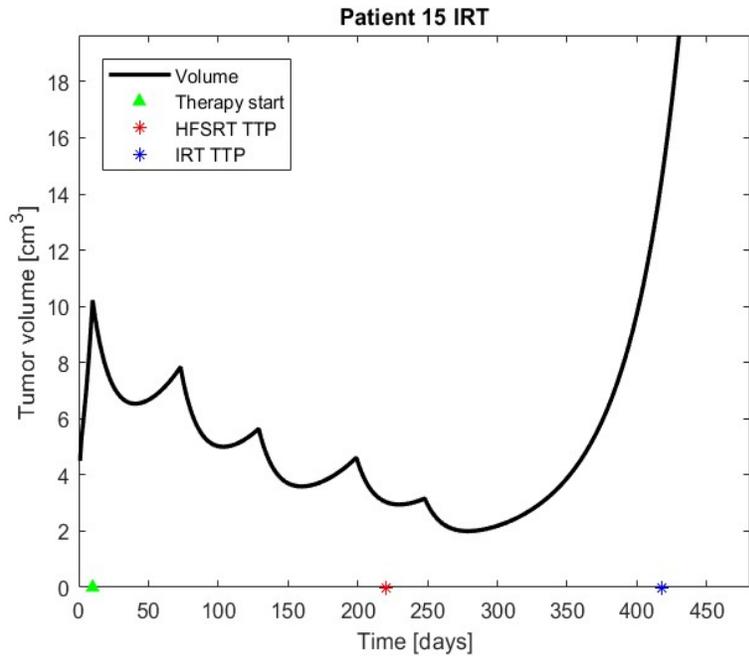












3.4.2 Data Analysis

A data analysis about the time distribution of each delivered fraction has been done, seeking for statistical information and patterns which may be exploited *a priori* when designing an intermittent protocol.

The normalized time distribution of each dose for the 16 patients and the related box plots are reported below. It is expected and straightforward to see from both the summary plot in Figure 3.6 and standard deviation values that subsequent doses have a greater time range of distribution. This tendency for a sharp-to-uniform spread of delivery times may provide practical useful information (e.g. how a group of patients treated simultaneously can be logistically distributed).

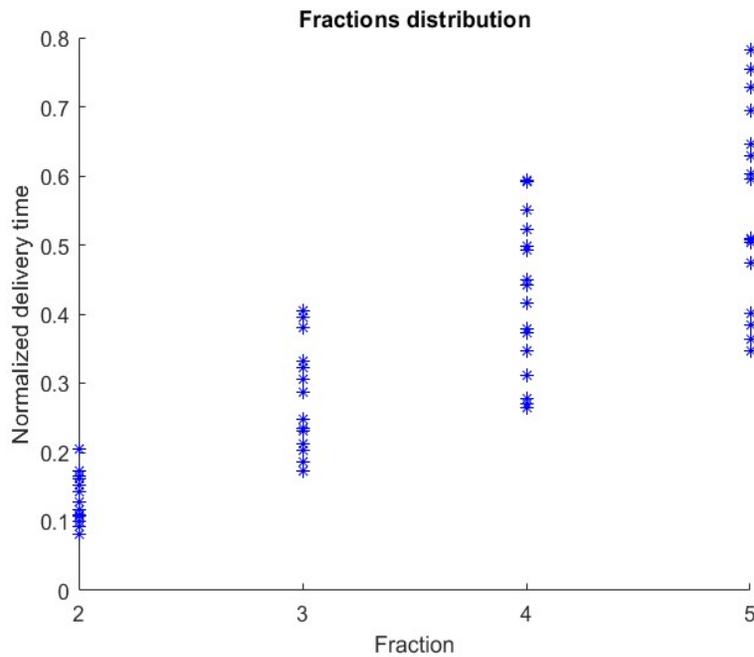
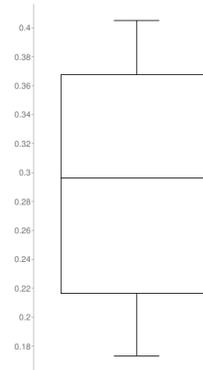
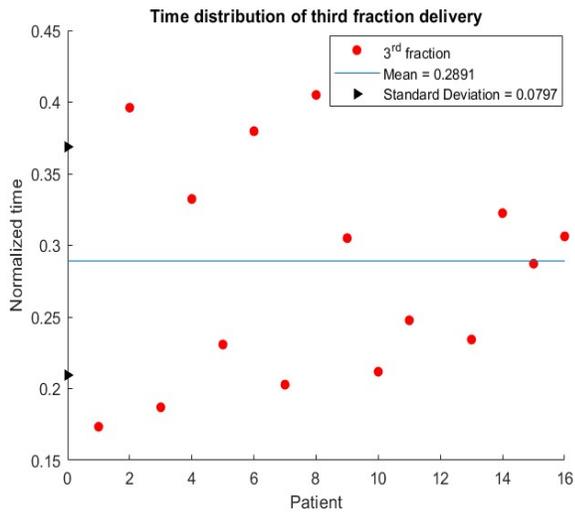
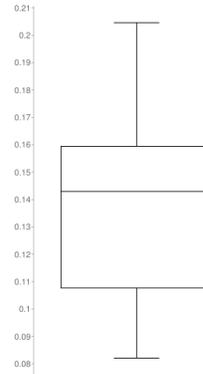
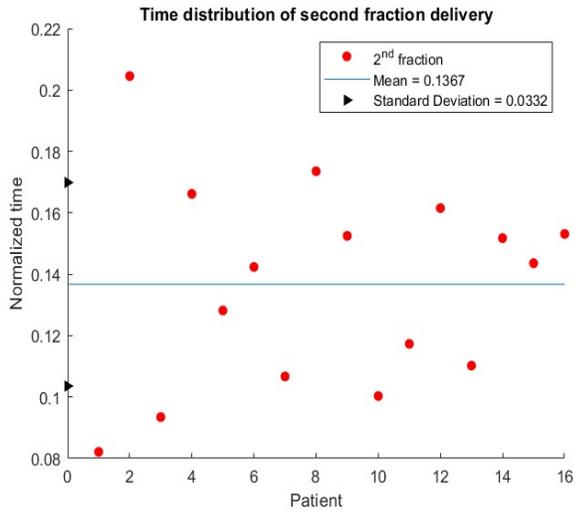
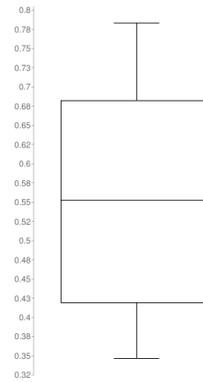
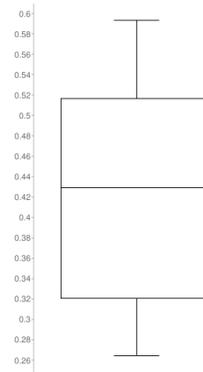
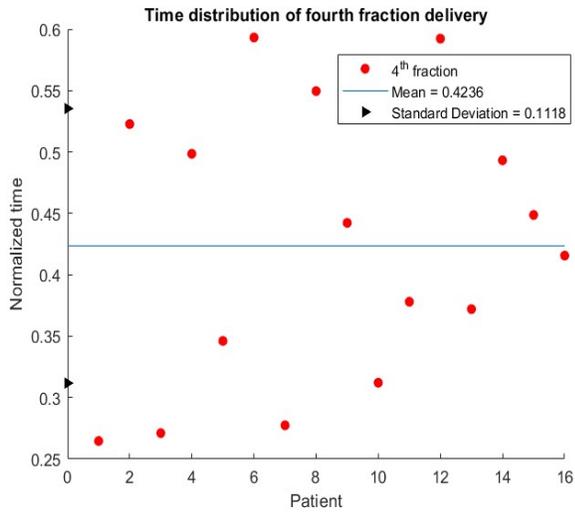


Figure 3.6: Summary plot of fractions distribution





3.5 Hyperfractionation

Normal and tumour tissue radiosensitivity may vary strongly between individual patients. The opportunity to either continue or interrupt treatment at each of the evaluation time-points holds great potential for treatment personalization.

The total number of delivered fractions can be adjusted to enable personalized dose escalation given the absence of acute normal tissue toxicity. In this model, potential tissue complications are not considered, although it is clinically confirmed that, in the intermittent setting, normal tissue may be capable to compensate for radiation-induced damage more effectively than the tumour. This would motivate an escalation of the total delivered dose in the IRT setting.

In [3], dose escalation does significantly enhance patients' survival probabilities. Here, instead of dose escalation, an IRT hyperfractionation schedule of the usual total dose amount is investigated, in order to see if it is (at least virtually) possible to improve the final outcome without necessarily incur in plausible radiation induced damage to normal tissue.

Since the number of fractions is increased but the therapeutic effect is left unchanged, the tumour surviving fraction varies as follows:

$$1 - S_{Hyper} = (1 - S_{Hypo}) \cdot \frac{5}{n} \implies S_{Hyper} = 1 - \left[\frac{5}{n} \cdot (1 - S_{Hypo}) \right]$$

Where S_{Hyper} is the new parameter value due to the introduction of hyperfractionation in the model. S_{Hypo} is the value found within the fitting stage and n is the total number of delivered fractions. The idea is that, when $n = 5$, the effect of a single radiation dose is left unchanged, while it decreases as n increases (i.e. when hyperfractionating).

3.5.1 Results

The plots below show the best predicted hyperfractionation schedules of each patient, obtained through the usual GA for IRT but with S_{Hyper} replacing S_{Hypo} . n is assumed to vary from 6 to 12, thus allowing up to 12 total fractions delivered. The optimization problem is the same as before, but this time the control vector, which again encodes the the allowed time intervals between doses, has n elements:

$$PST(\mathbf{c}) = \int_{t_0}^{t_{cut-off}(\mathbf{c})} d\tau \quad (3.15)$$

$\mathbf{c} = [c_1, c_2, \dots, c_n]$ where $c_i \in \{4, 5, 6, 7, 8, 9, 10 \text{ weeks}\}$

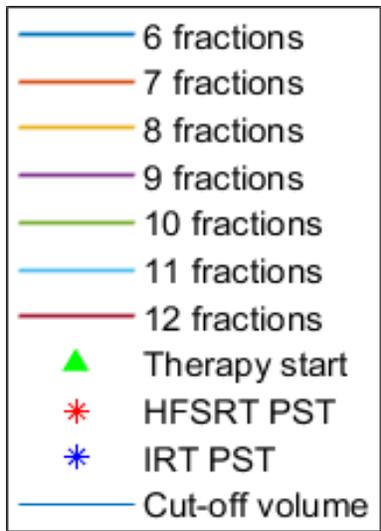
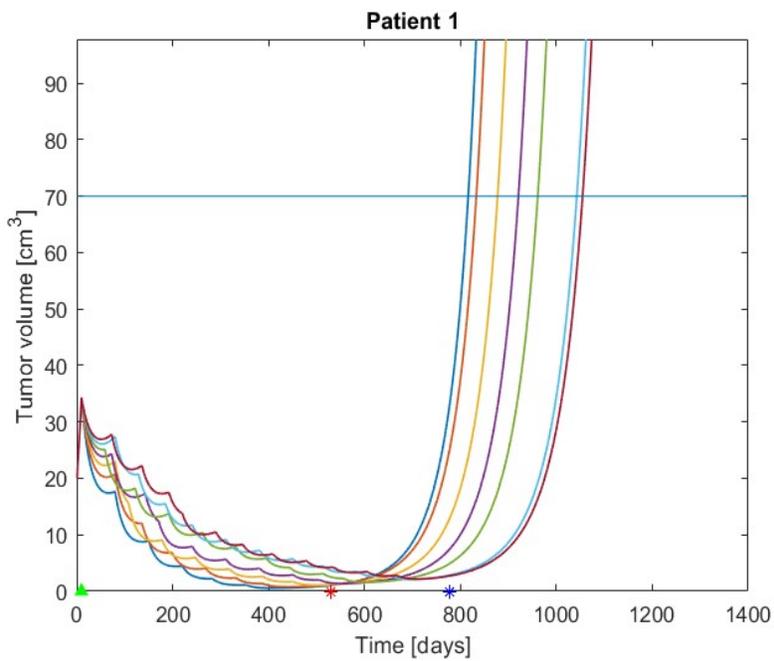
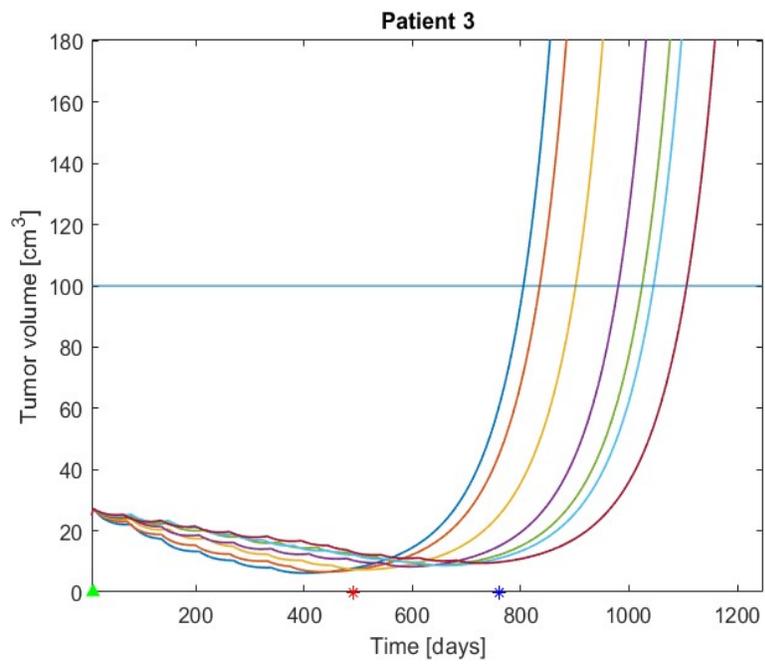
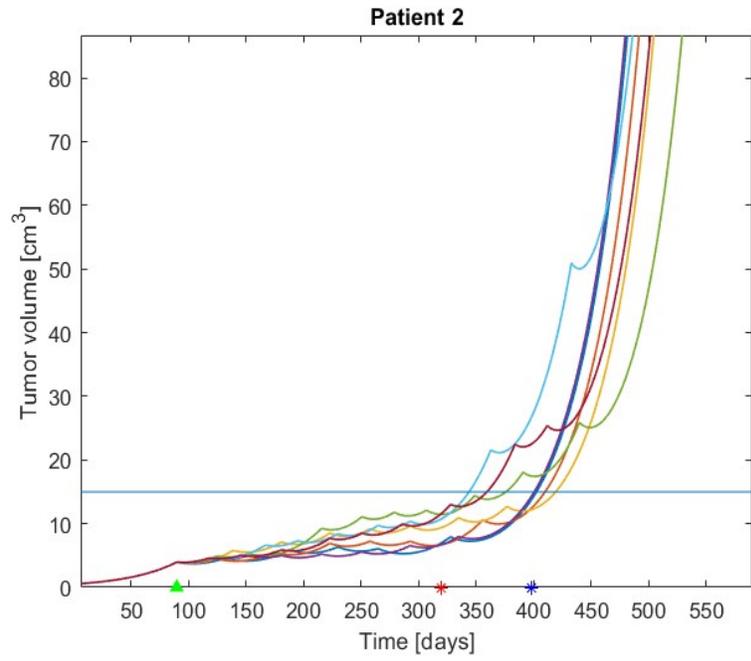
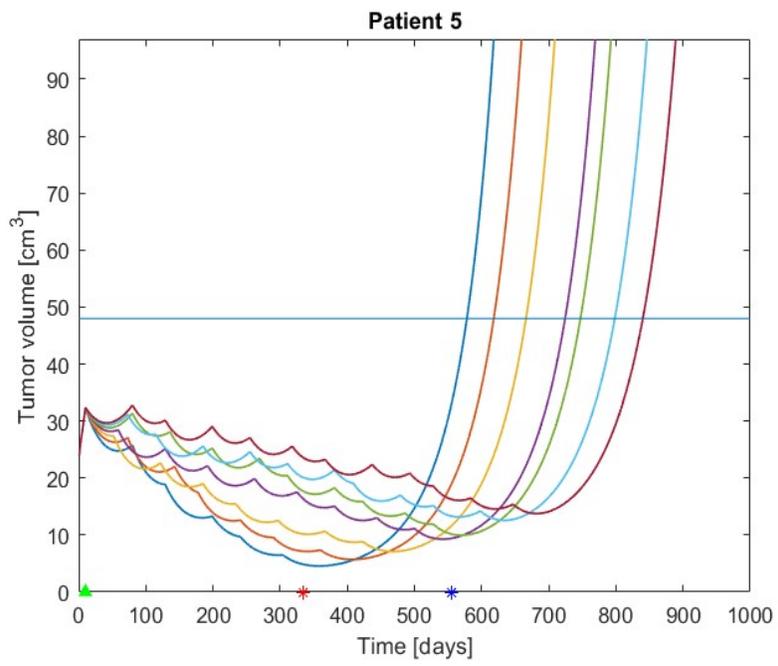
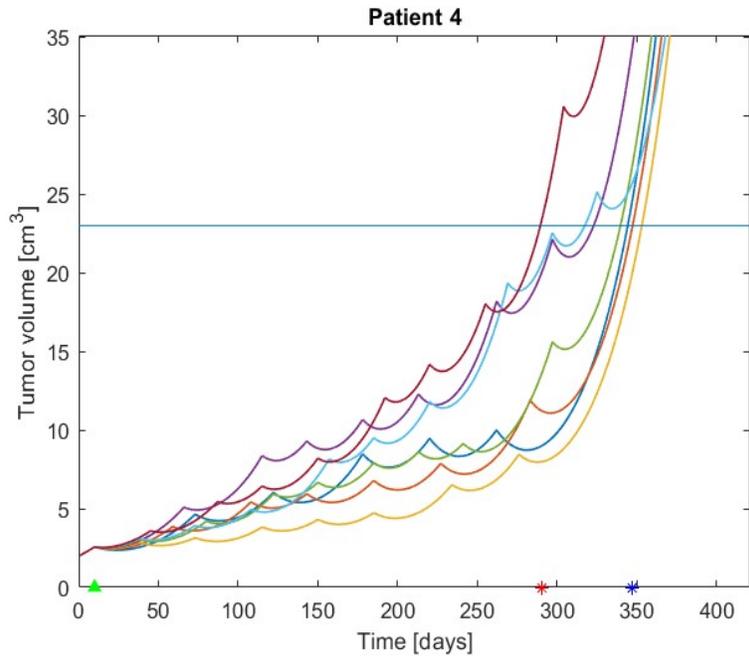
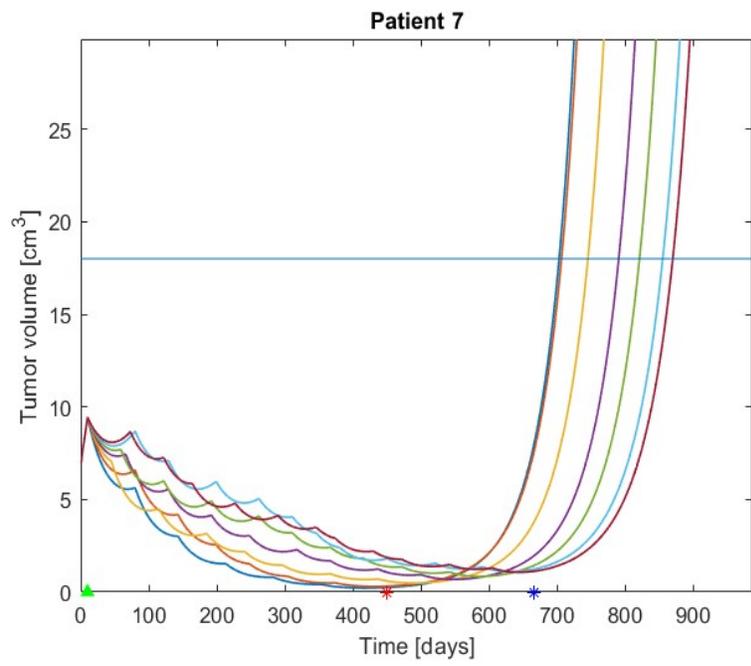
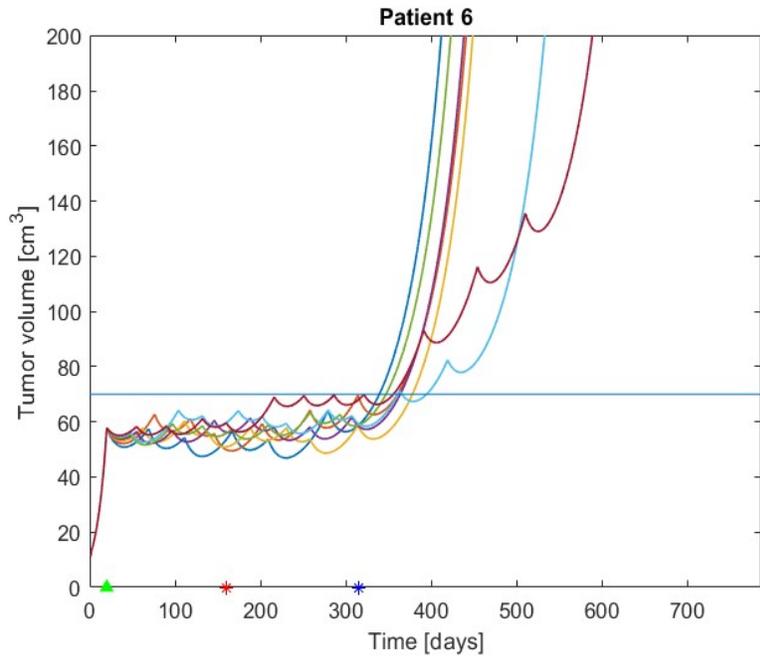


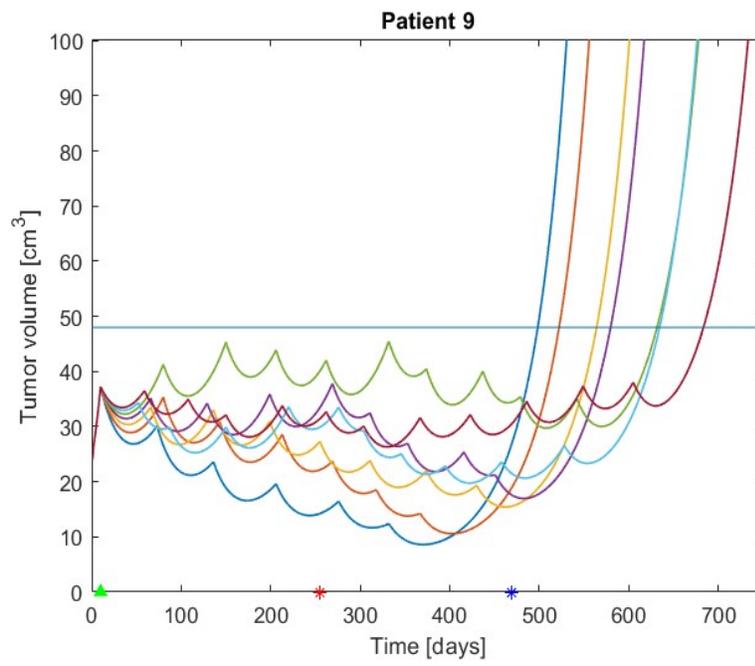
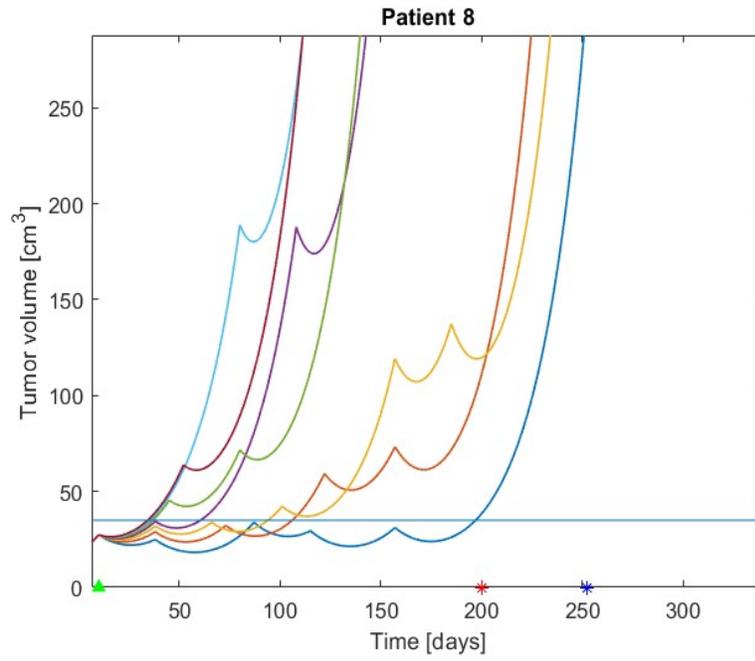
Figure 3.7: Legend of the hyperfractionation plots

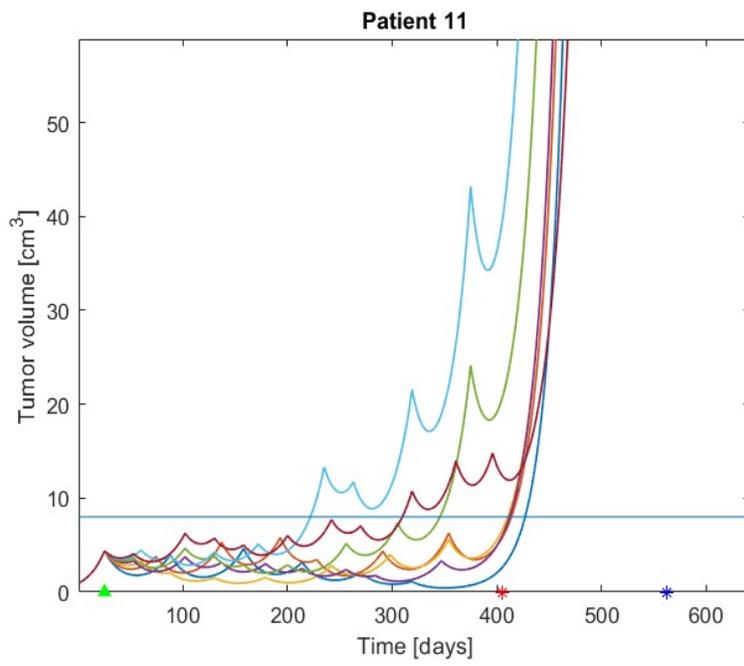
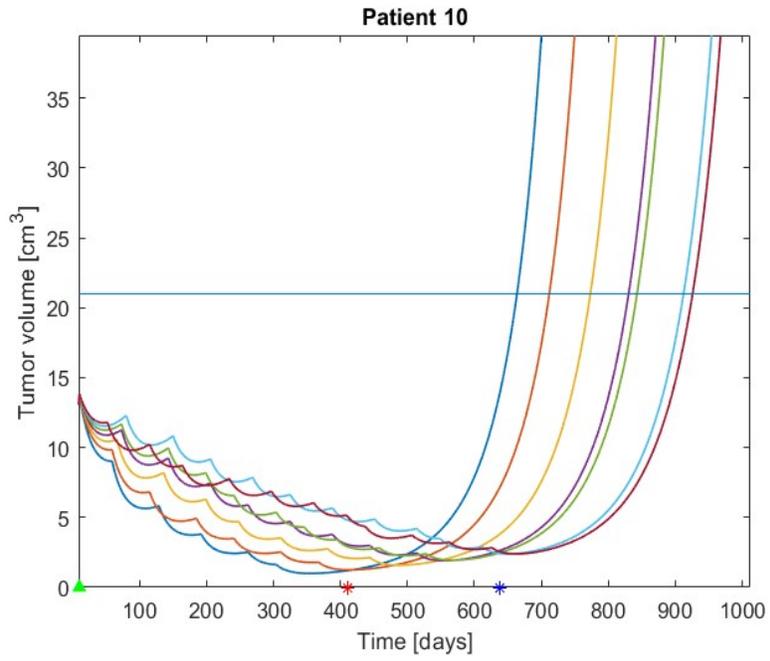


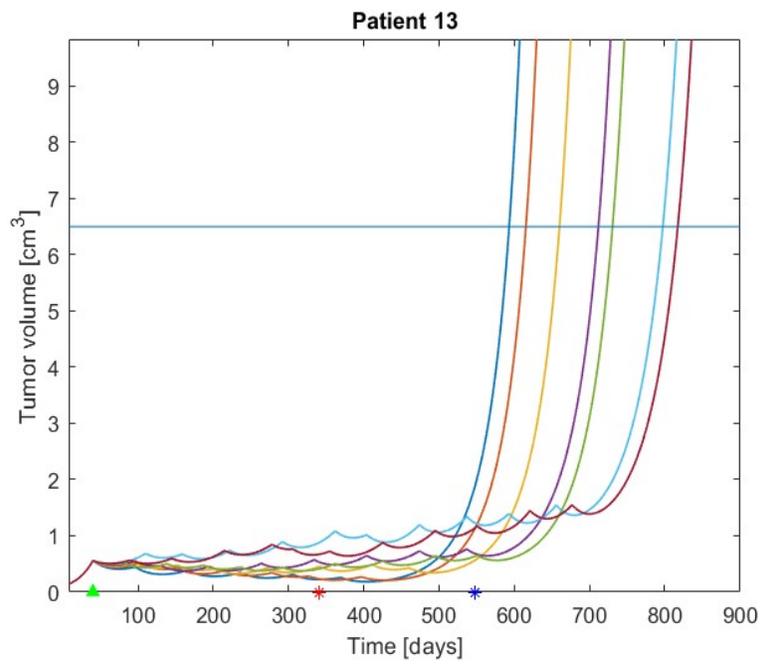
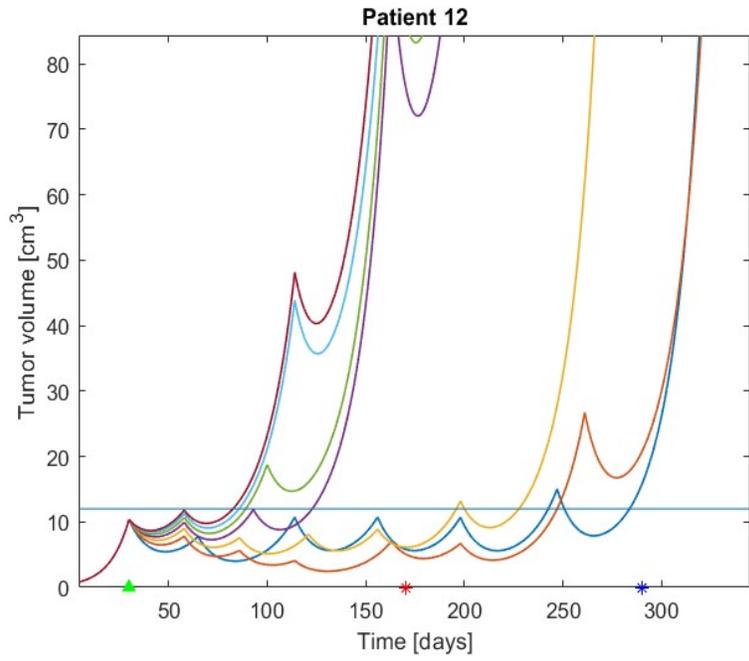


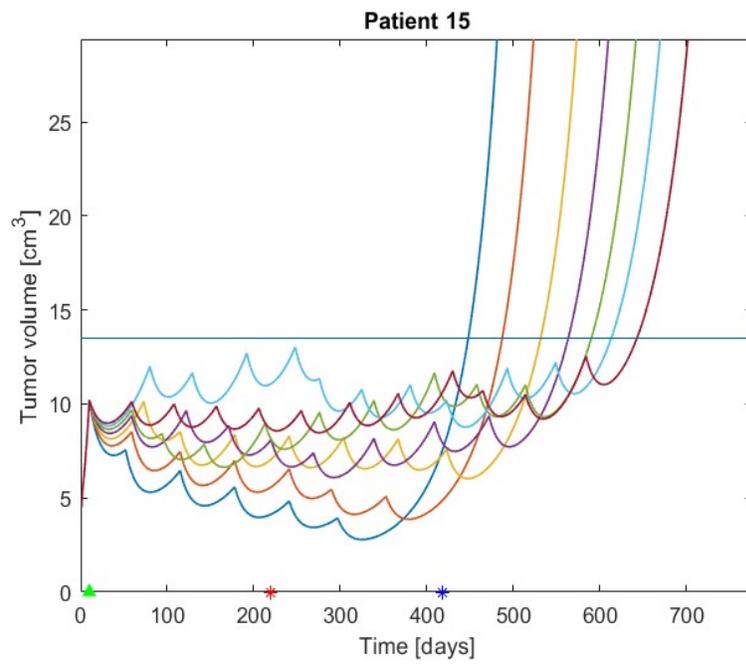
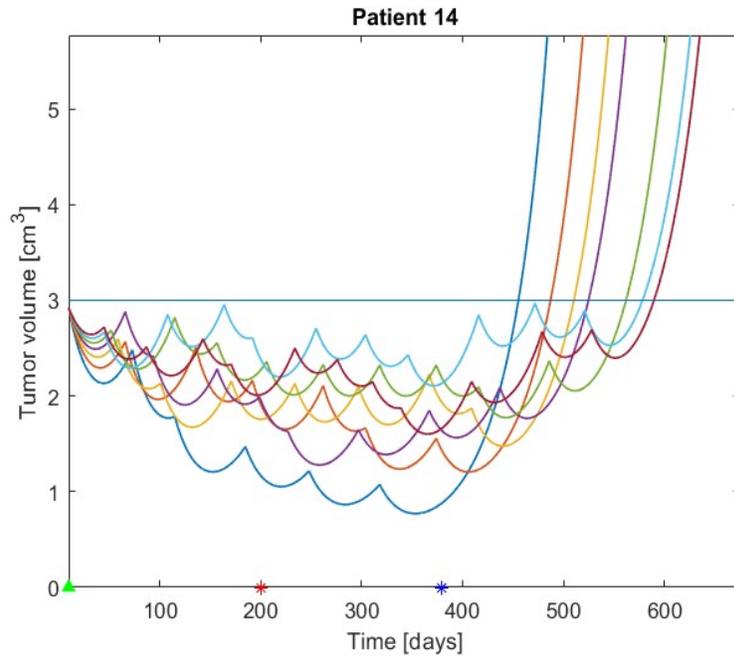


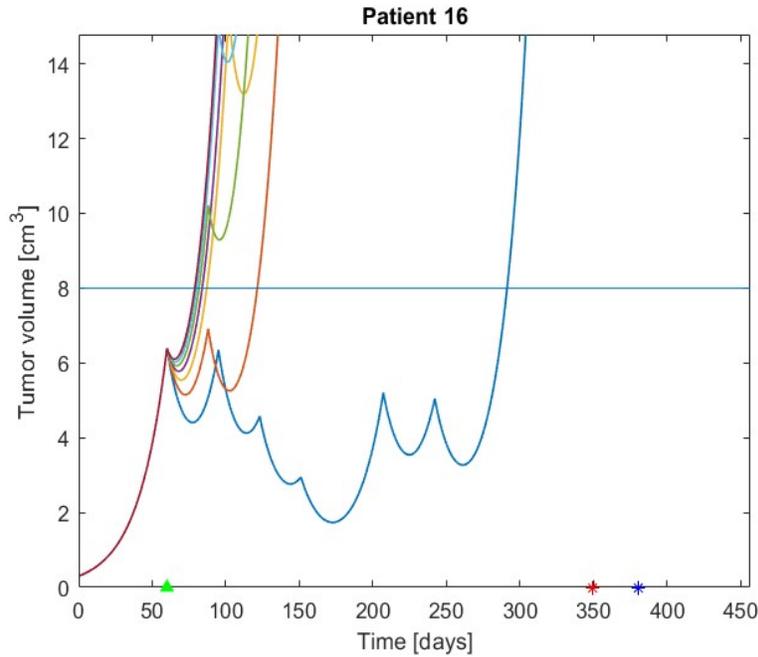












3.5.2 Data Analysis

Hyperfractionation protocol predictions exhibit broad heterogeneity. Every patient responds in its own way, with a highly-personalized therapy schedule for each fractionation.

Nevertheless, it is possible to identify three main groups of patients:

- **Group 1.** Hyperfractionation shows a profitable trend: the more the fractionation the more the PST is increased [Patients #1, 3, 5, 7, 9, 10, 13, 14, 15].
- **Group 2.** Patients of this group partially benefit from hyperfractionation. After a certain dose split, the tendency moves toward a smaller PST, thus making the IRT hypofractionation or even HFSTR more effective [Patients #2, 4, 6].
- **Group 3.** The hyperfractionation protocol has no advantage compared to IRT hypofractionation and sometimes even to HFSRT. Moreover, the trend is in general the opposite of that of group 1: the more the dose is split, the lower is the PST [Patients #8, 11, 12, 16].

In Figure 3.9, the distribution of patients' parameters is reported. Each group has been highlighted as in the legend (Figure 3.8). The growth rates λ show a uniform behaviour (it is however important to stress that the dataset is particularly limited). For what concerns resistance rates, it is clear that they are low for every patient of group 1. This means that hyperfractionation may be a good strategy when a good response to chemotherapy is observed. From the last picture is possible to see that patients of group 3 are characterized by a low surviving rate (S) of tumour after radiation is applied. This suggests to prefer a hypofractionated protocol whenever a good response to radiation therapy shows up.

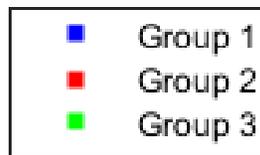


Figure 3.8: Plot legend of section 3.5.2

In Figure 3.10, each possible couple of parameters is reported. It seems that statistical correlation emerges in none of them, either in general and within a group.

In Figure 3.11 a complete scatter plot of the three parameters is reported. It is evident that patients of group 1 (which is the most numerous among the three) forms a cluster, thus corroborating that identifying such a group makes sense. This result may be clinically useful, since it allows to classify future patients with similar features as those who will probably benefit from a hyperfractionation more than a traditional hypofractionated protocol.

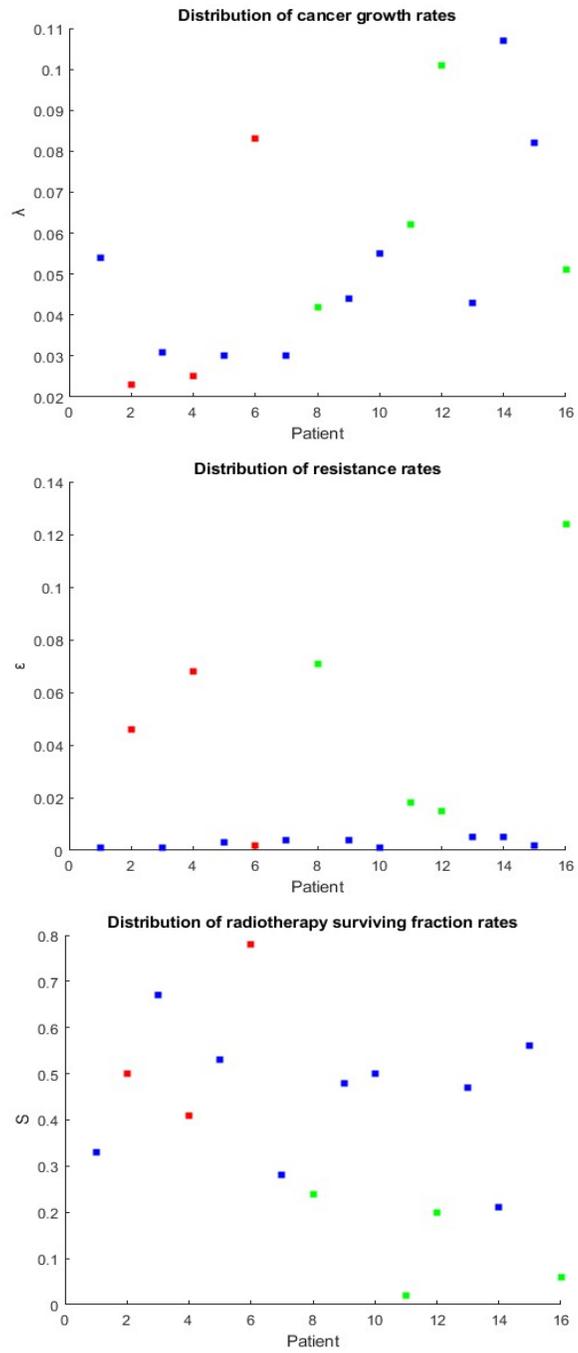


Figure 3.9: Single-parameter distributions

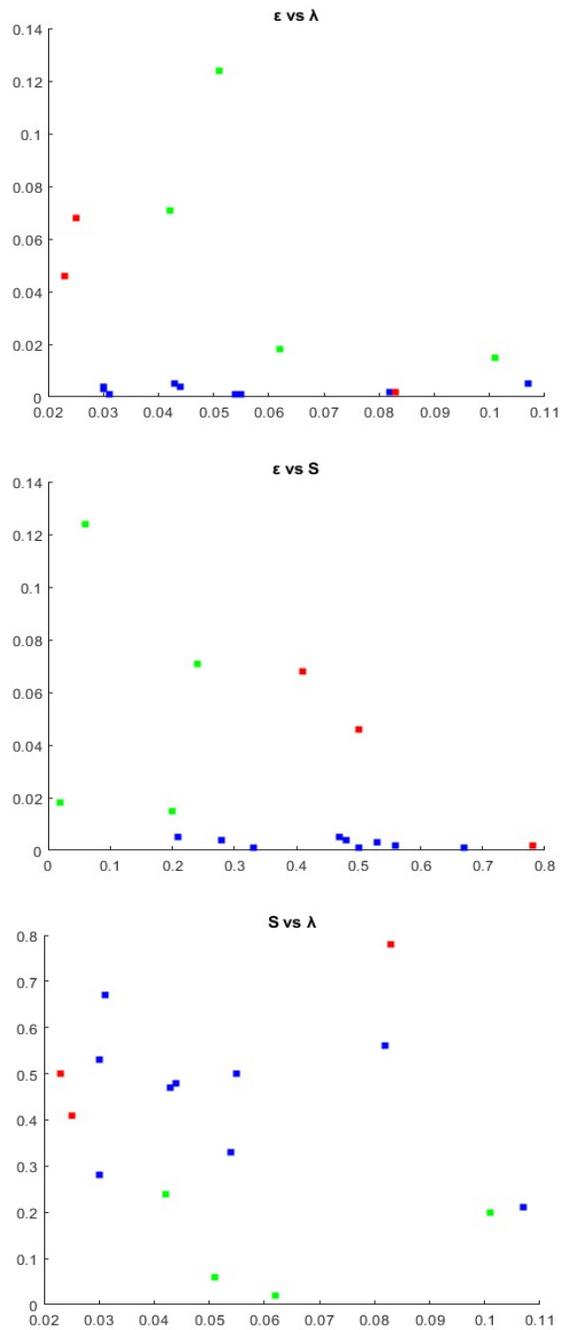


Figure 3.10: Parameters pairs. No evident statistical correlation comes off.

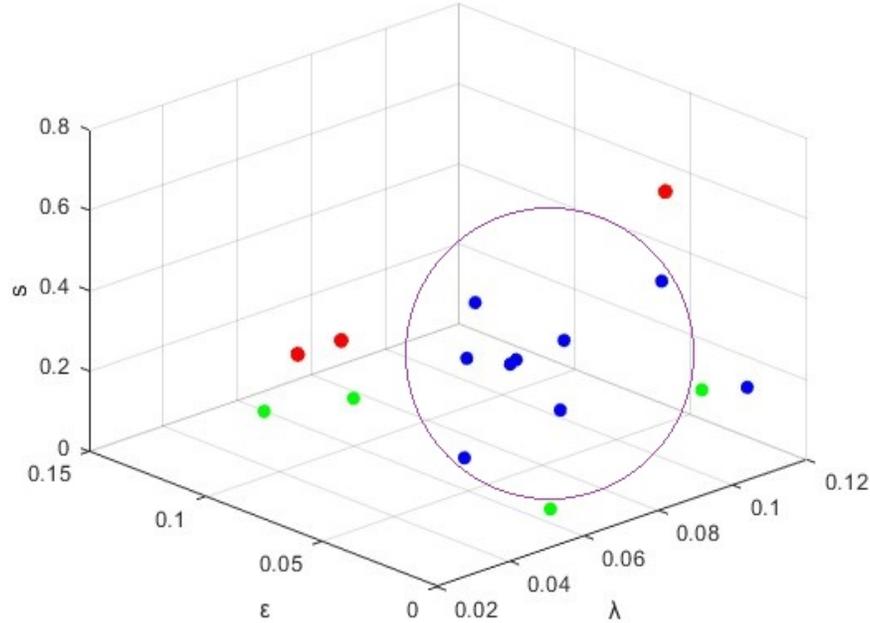


Figure 3.11: Scatter plot over all parameters. A purple circle highlights the presence of the aforementioned cluster.

3.6 Model reliability

Up to this moment, almost every result or consideration has been discussed *a posteriori*, namely knowing the course of the disease thanks to the available data. Actually, when designing a fractionation protocol, the amount of information is limited, since clinical decisions must be taken in a short time frame in order to prevent a irremediable cancer spread.

In this section a measure of model’s reliability is proposed. In particular, the posed question is:

On average, what is the amount of data required by the model in order to perform an arbitrarily good parameters’ estimation?

Mathematically, such a measure is performed through the *relative error* over parameters prediction as a function of the available amount of information:

$$\delta_P(m) = \left\| \frac{P_{partial}(m) - P_{complete}}{P_{complete}} \right\| \quad (3.16)$$

Where $P \in \{\lambda, \epsilon, S\}$ indicates the generic parameter, $P_{partial}$ is the parameter estimate obtained with m measures from the dataset and $P_{complete}$ is the original parameter estimate obtained with the full dataset as reported in section 3.3.

The *average* relative errors over the patient cohort together with the related standard deviations are reported in Figure 3.12. m varies from 2 to 6, since on average that's the maximum number of available measures. the first measure is the last took before starting the therapy, while $m = 2$ indicates the first measure under therapy effect.

The growth rate error diminishes as m increases, same for the resistance rate. However, the survival fraction rate doesn't show evident improvement with a larger m , so one may assume that this parameter can be safely inferred in a early therapy stage, even though it departs about 100-200% from the original value.

Notice that the resistance rate error assumes values definitely larger than those of the other two parameters. This means that the ϵ value predicted by the model is particularly sensitive with respect to perturbations of the original dynamics.

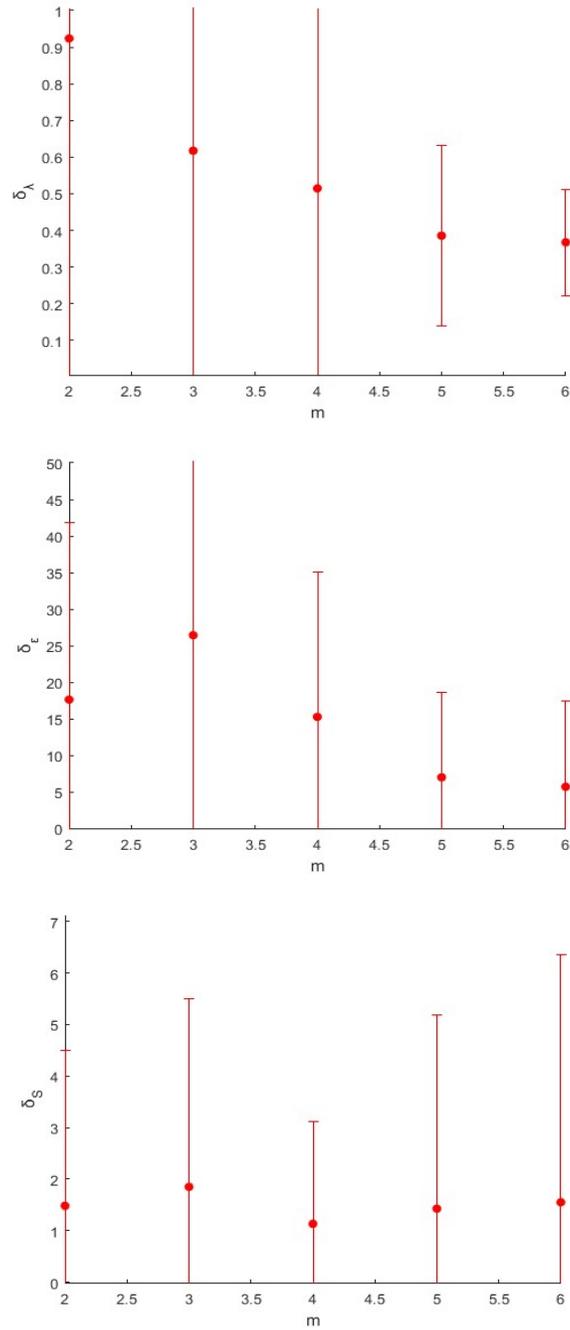


Figure 3.12: Relative errors for the three model parameters.

Chapter 4

Radiation resistance: polyclonal modelling of recurrent high-grade glioma

A neoplasm can be viewed from an evolutionary perspective as a large, genetically and epigenetically heterogeneous population of individual cells [11].

In the scientific community, malignant cancers are widely believed to be governed by Darwinian dynamics. The process of carcinogenesis includes genetic instability (i.e. the tendency for DNA mutations) and highly selective local microenvironments, the combination of which promotes somatic evolution. These microenvironmental forces, such as hypoxia and acidosis, are not only highly selective, but are also able to induce genetic instability. As a result, malignant cancers are dynamically evolving group of cells living in distinct microhabitats that almost certainly ensure the emergence of therapy-resistant populations [7].

Nevertheless, standard radiobiology theory of radiation response assumes a uniform innate radiosensitivity (i.e. the ability of the radiation to damage the tumor cell) of tumours. However, experimental data show that there is significant intratumoral heterogeneity of radiosensitivity [1].

In the presence of an intratumoral distribution of radiosensitivity, there is rapid selection of radiation-resistant cells over a course of fractionated radiation therapy. Standard treatment fractionation regimes result in the near complete replacement of the initial population of sensitive cells with a population of more resistant cells. Further, as treatment progresses, the tumour becomes more resistant to further radiation treatment, making each fractional dose less efficacious.

However, the emergence of a resistant phenotype is not in itself clinically sig-

nificant. That is, resistant cells affect patient outcomes only when they form a sufficiently large population to allow tumor progression and treatment failure: in destroying the entire population of sensitive cells, maximum dose therapy imposes intense selection for resistant phenotypes and, by eliminating all potential competitors, their proliferation is maximized, which is a well-known evolutionary phenomenon termed *competitive release* [5].

The time between radiation fractionations for IRT could allow for regrowth of both resistant and sensitive populations. Pre-clinical data and evolutionary convention suggest that resistant cells may display a fitness disadvantage relative to sensitive clones in the absence of the selective pressure, allowing for the sensitive subpopulation to preferentially repopulate the tumour [3].

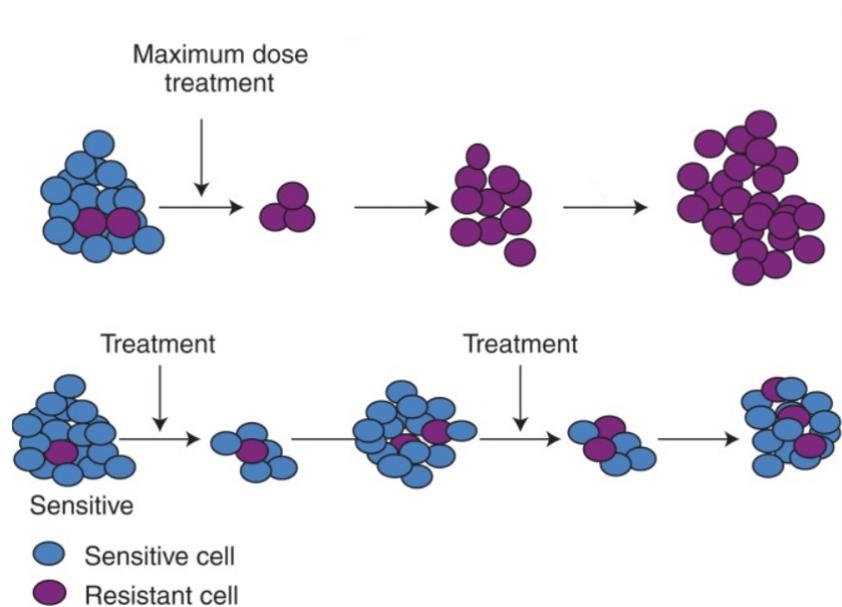


Figure 4.1: Standard HFSRT aims to kill the largest possible tumour mass by delivering the maximum radiation dose in a short time, leading to the undesired competitive release where resistant clones dominates the tumour (*top row*). With an intermittent approach, sensitive population is expected to exploit its fitness advantage to repopulate the environment more effectively than resistant population, allowing the patient to carry a bearable tumour burden which can be still managed due to the consistent presence of sensitive cells (*bottom row*).

It is expected that intermittent radiotherapy (IRT) treatments would provide an advantage over daily hypofractionated stereotactic radiation therapy (HFSRT) whenever sensitive subclones may repopulate more effectively between fractions.

In this chapter, the aim is to simulate highly-personalized IRT treatments that may soften the aforementioned unfavourable competitive release mechanism.

4.1 Multicompartmental model

The first proposed model is a slight modification of the monoclonal description presented in chapter 3. Raising resistance is modelled through a constant transfer rate from a *radiosensitive cells compartment* to a *radioresistant cells compartment*. Each of the two populations is characterized by its own growth rate:

$$\frac{dV_s(t)}{dt} = \lambda_s V_s(t) - p V_s(t) \quad (4.1)$$

$$\frac{dV_r(t)}{dt} = \lambda_r V_r(t) + p V_s(t) \quad (4.2)$$

Where p is the constant transfer rate [$time^{-1}$], λ_s and λ_r are the growth rates of the sensitive population and resistant population respectively.

The radiation therapy effect is again implemented through a survival fraction S of the tumour after undergoing a radiotherapy fraction at a certain time t_{RT} . The novelty here is that this effect applies not to the whole tumour mass, but only over the sensitive compartment.

$$V_s(t_{RT}^+) = S \cdot V_s(t_{RT}^-) \quad (4.3)$$

$$V_d(t_{RT}^+) = V_d(t_{RT}^-) + (1 - S) \cdot V_s(t_{RT}^-) \quad (4.4)$$

The dying compartment is the same as in the monoclonal model, with a decay rate equal to λ_s , since those are the cells suffering the subsequent mitotic catastrophe.

$$\frac{dV_d(t)}{dt} = -\lambda V_d(t) \quad (4.5)$$

The resistant population is thus immune to the whole adopted therapy, virtually proliferating without any limit even in case of a massive radiation dose escalation.

The system of ODEs (4.1), (4.2) and (4.3) is linear. Its solution is easily found:

$$V_s(t) = V_{s,0} \cdot e^{(\lambda_s - p)(t - t_0)} \quad (4.6)$$

$$V_r(t) = e^{\lambda_r(t - t_0)} \cdot \frac{(V_{r,0}\lambda_r - V_{r,0}\lambda_s + V_{s,0}p + V_{r,0}p)}{\lambda_r - \lambda_s + p} - \frac{V_{s,0} \cdot p \cdot e^{(\lambda_s - p)(t - t_0)}}{\lambda_r - \lambda_s + p} \quad (4.7)$$

$$V_d(t) = V_{d,0} \cdot e^{\lambda_s(t - t_0)} \quad (4.8)$$

Where $V_{s,0}$ and $V_{r,0}$ are the initial sensitive population and resistant population respectively. The total tumour volume is:

$$V(t) = V_s(t) + V_r(t) + V_d(t), \quad V(0) = V_{s,0} + V_{r,0}$$

As usual, pre-treatment growth of both populations is malthusian, namely $p = 0$ for $0 \leq t < t_0$.

Remark. In the following, the used dataset is the same as in chapter 3, obtained through the same therapy schedule, which is a combination of both HFSRT and chemotherapy. It is natural to wonder how here chemotherapy resistance is taken into account. Even though the model focuses on radioresistance rather than chemoresistance, it is essential to specify that here the transfer rate parameter p is assumed to describe both kind of resistance undistinguishably. The cost to pay for this simplification is that of working in a worst-case scenario: those cells developing chemoresistance automatically are assumed to gain radioresistance as well.

4.1.1 Linear Stability Analysis

Exactly as in the monoclonal model, (4.5) is decoupled from the rest, thus consisting in a one-dimensional dynamical system:

$$\dot{V}_d = f(V_d) = -\lambda_s \cdot V_d \quad (4.9)$$

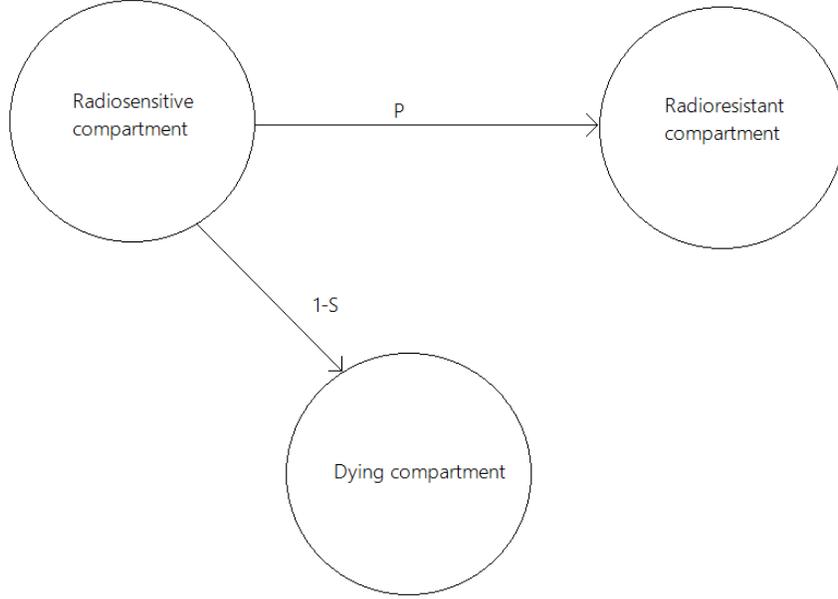


Figure 4.2: Schematic representation of the multicompartmental model

Again, by imposing the steady-state condition $\dot{V}_d = 0$ we find the unique fixed point $V_d^* = 0$, which is stable since $\lambda_s > 0$.

Equations (4.1) and (4.2) form a two-dimensional linear system of coupled ODEs, which can be rewritten in the compact form:

$$\dot{V}_s(t) = f(V_s, V_r) \quad (4.10)$$

$$\dot{V}_r(t) = g(V_s, V_r) \quad (4.11)$$

The unique fixed point is the origin of the phase space: $(V_s(t)^*, V_r(t)^*) = (0, 0)$.

Since the system is linear, its stability can be deduced from the coefficients matrix:

$$J(V_s, V_r) = \begin{bmatrix} \lambda_s - p & 0 \\ p & \lambda_r \end{bmatrix}$$

whose trace and determinant are:

$$\tau = \lambda_s + \lambda_r - p \equiv \lambda_{tot} - p$$

$$\Delta = \lambda_r(\lambda_s - p) \equiv \lambda_{eff}\lambda_q$$

Where λ_{tot} is the tumour *total growth rate* and λ_{eff} is the *effective growth rate* of the sensitive population, which decreases as the transfer rate p increases.

The stability changes based on parameters' values:

- If $\lambda_{eff} < 0$ the origin is a saddle point. Physically, this corresponds to a situation where the sensitive population becomes extinct, while the resistant population proliferates. This scenario never occurs in the following.
- If $\lambda_{eff} > 0$ and $\lambda_{tot} > p$ the FP is unstable. Both populations proliferate.
- If $\lambda_{eff} > 0$ and $\lambda_{tot} < p$ the FP is stable. However, this condition is fulfilled only if $\lambda_s + \lambda_r < p$ and $\lambda_s > p$ simultaneously, which would require $\lambda_r < 0$, making this configuration unfeasible.
- If $\lambda_{eff} > 0$ and $\lambda_{tot} = p$ the FP is a center. However, this condition never occurs in the following.

4.1.2 Model fitting

Once again, the parameters fitting stage has been performed by means of a Genetic Algorithm. The models' parameters are $\lambda_s, \lambda_r, S, V_{s,0}, V_{r,0}$ and p , encoded in bit-string chromosomes as the one shown in Figure 4.2. The length of a chromosome is 44 bits, so the search space has a size of $2^{44} \sim 10^{13}$ possible solutions.

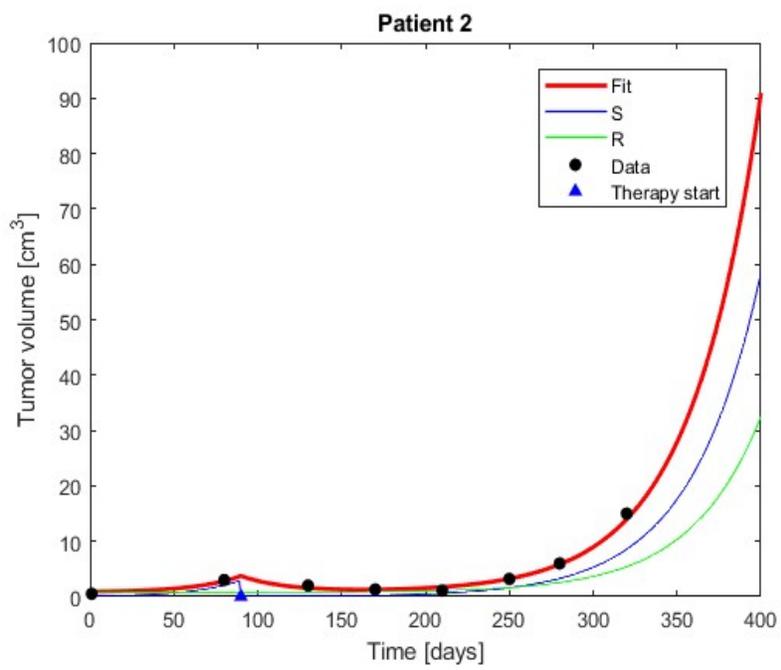
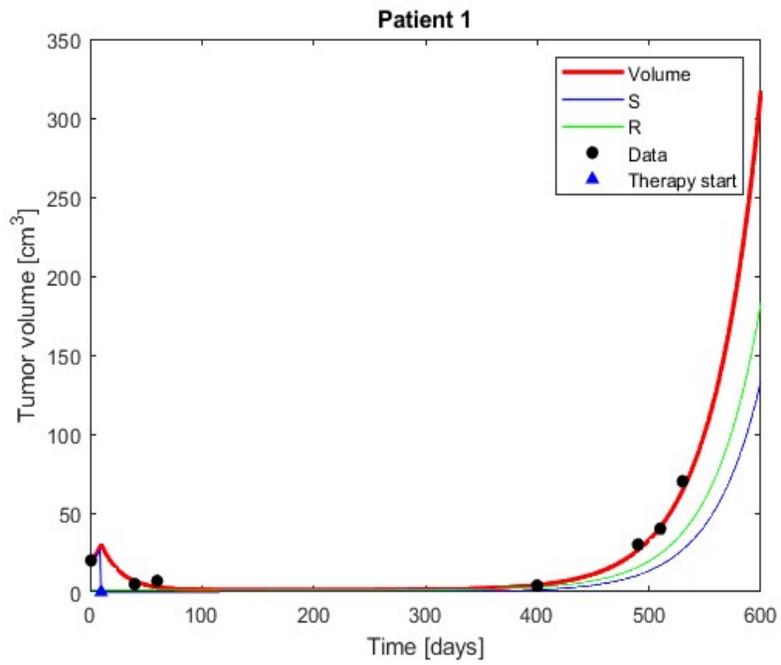
The used dataset is again the one presented in chapter 3 and same for the fitness function:

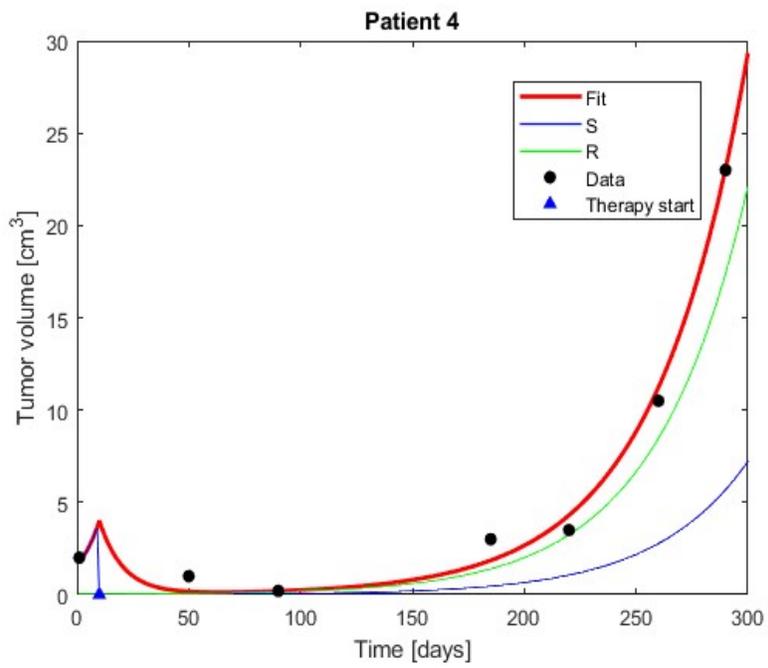
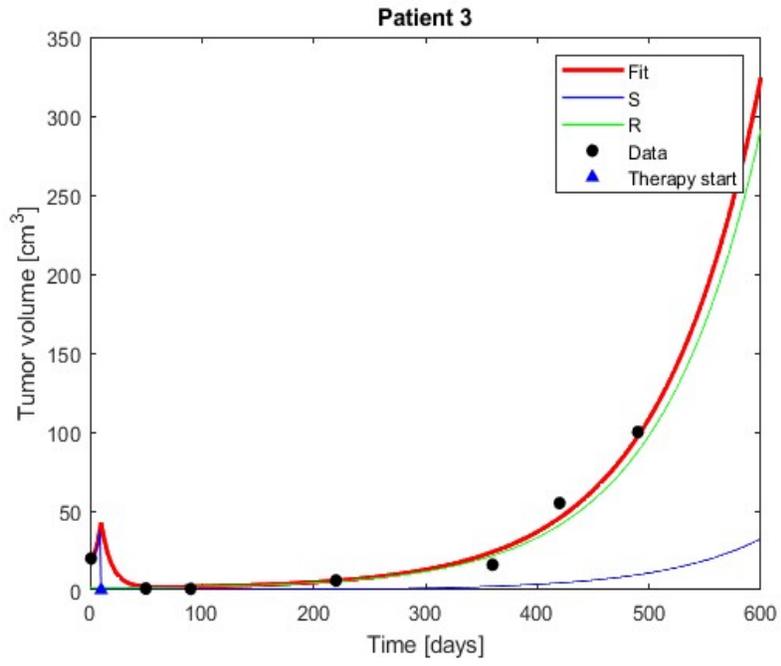
$$RMSE = \sqrt{\frac{\sum_{i=1}^N (V_{measured}(t_i) - V_{sim}(t_i))^2}{N}} \quad (4.12)$$

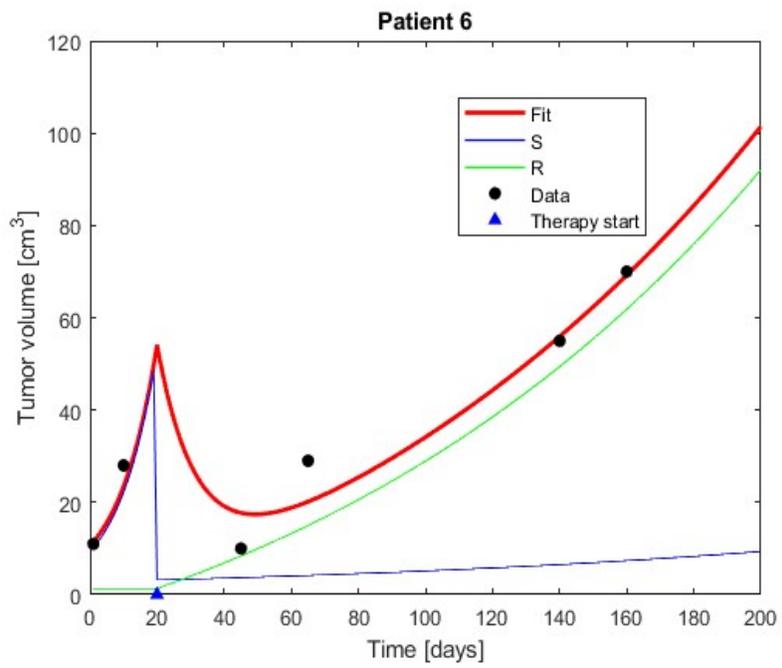
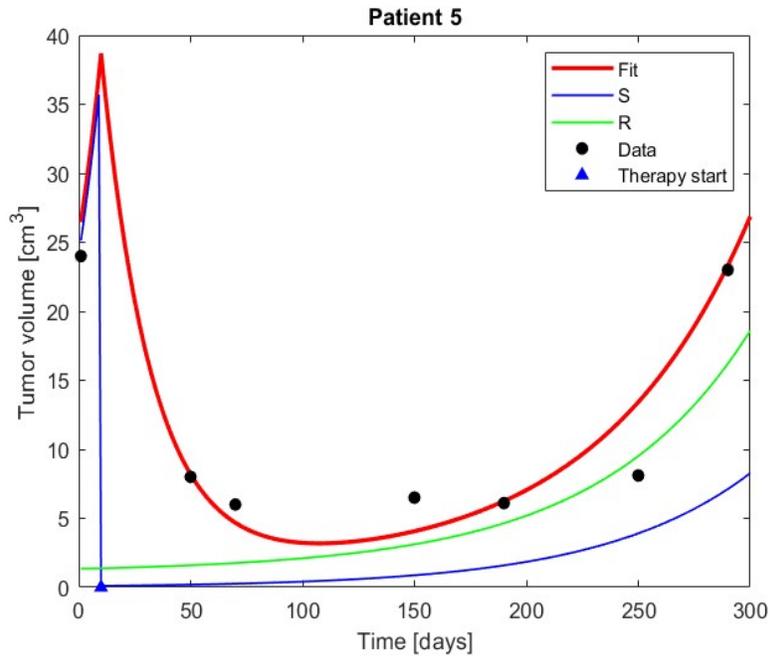
In the legend of the fitting plots, S and R indicate the volume of the predicted sensitive and resistant population respectively.

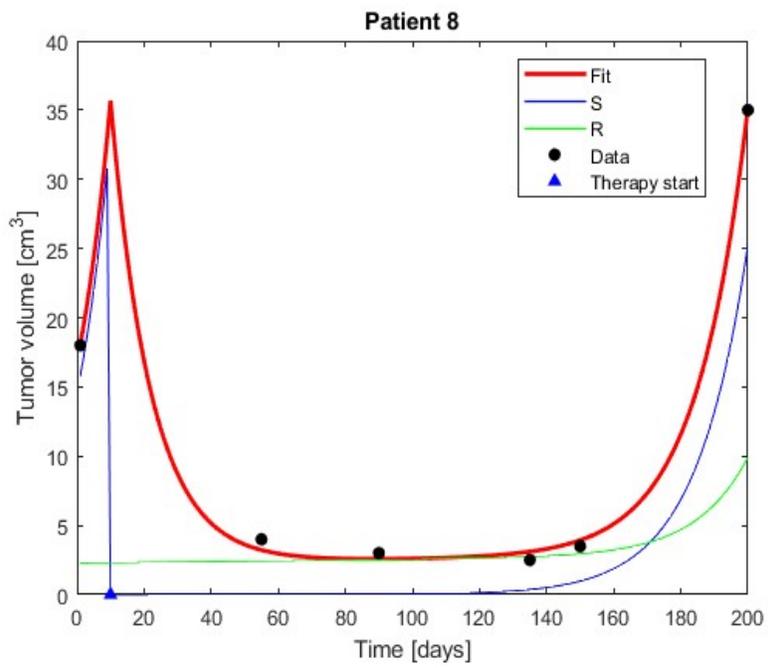
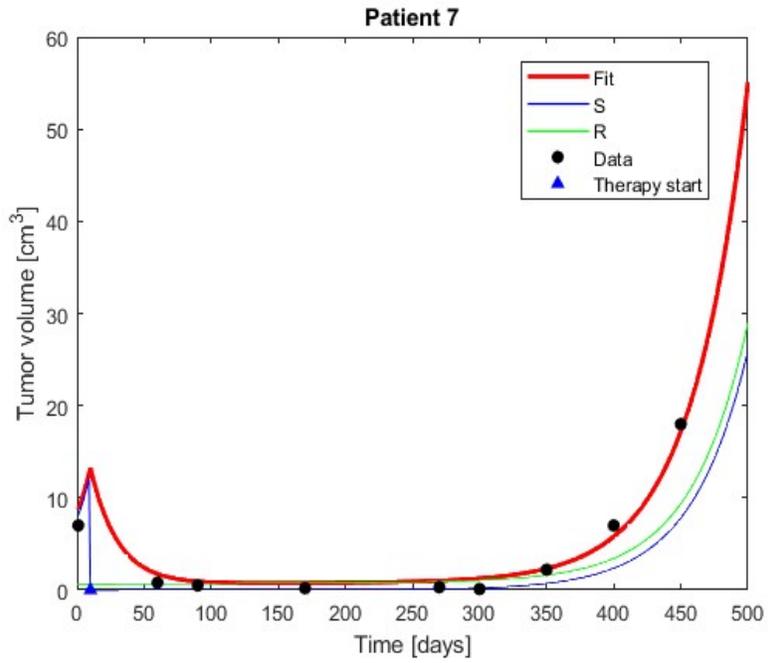
λ_s	λ_r	$V_{s,0}$	$V_{r,0}$	S	P
7 bits	7 bits	8 bits	8 bits	7 bits	7 bits

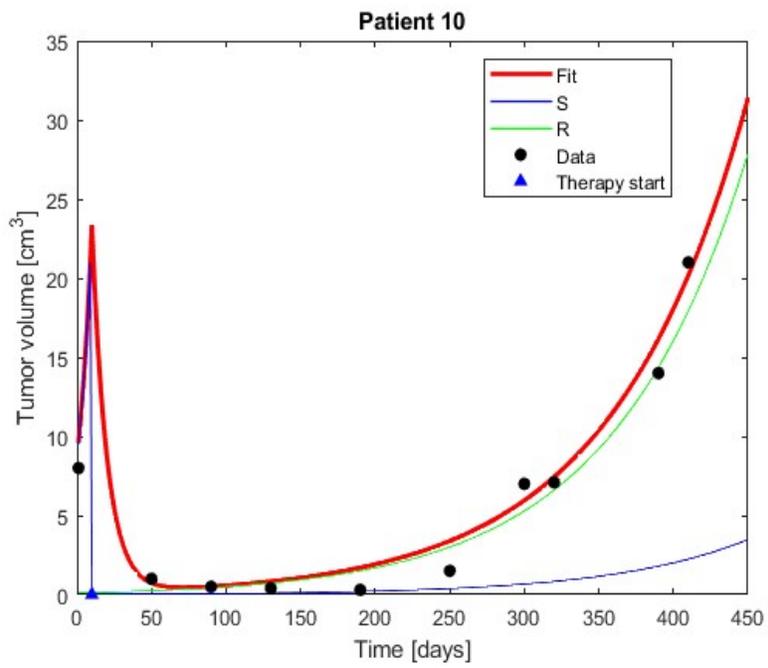
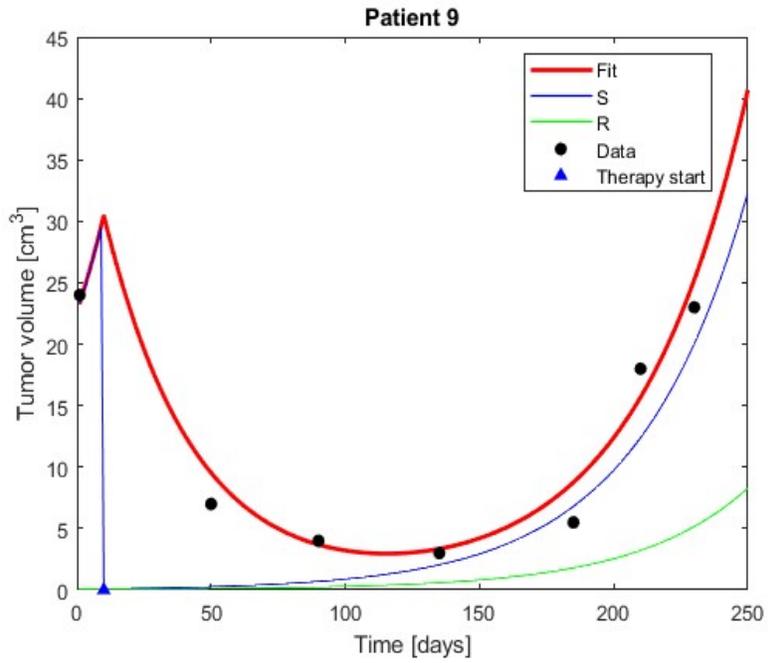
Figure 4.3: Partition of the bit-string chromosome

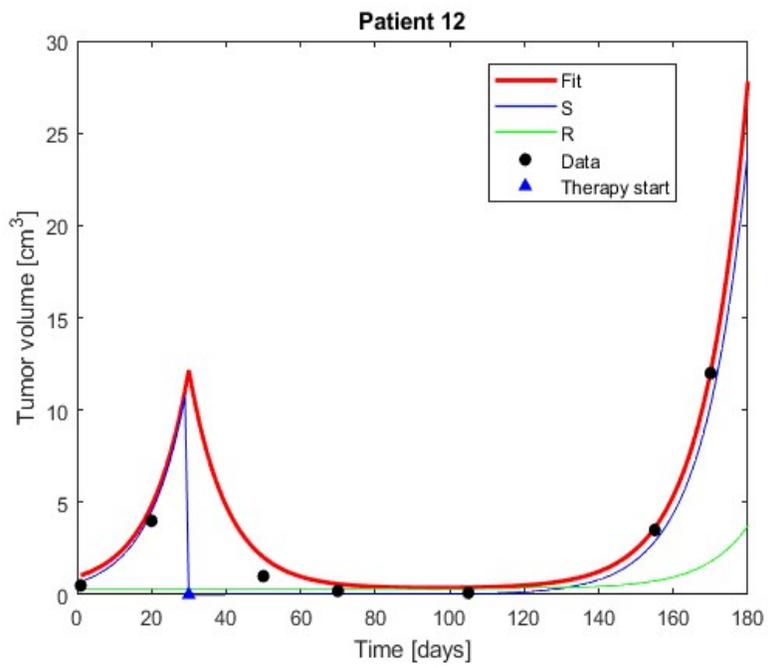
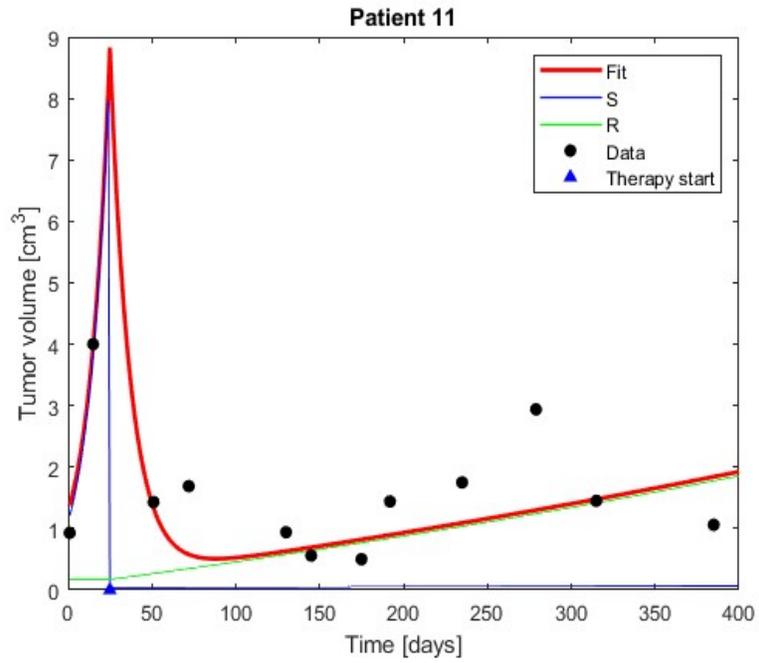


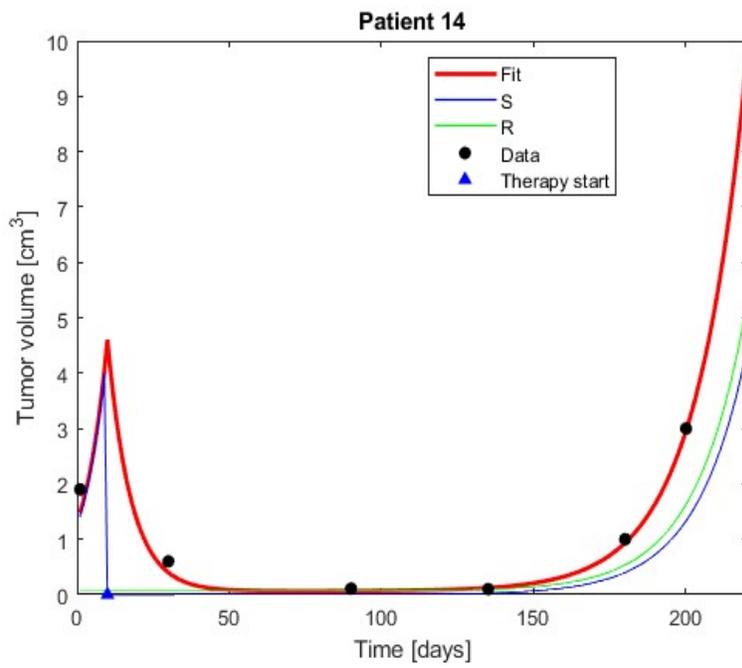
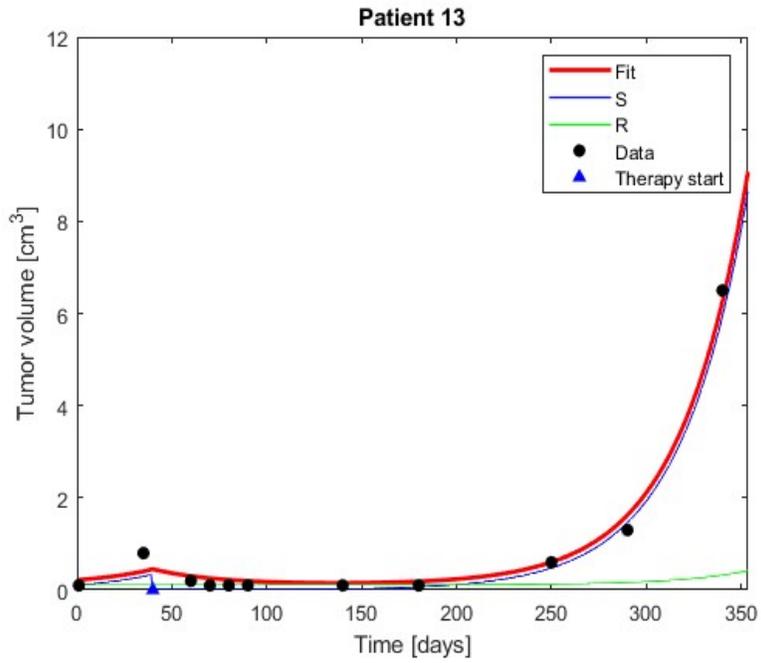


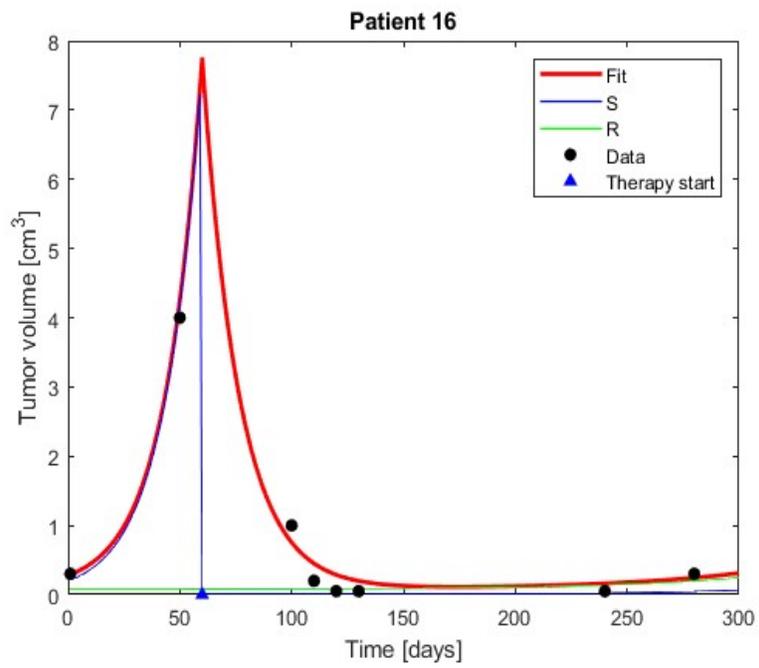
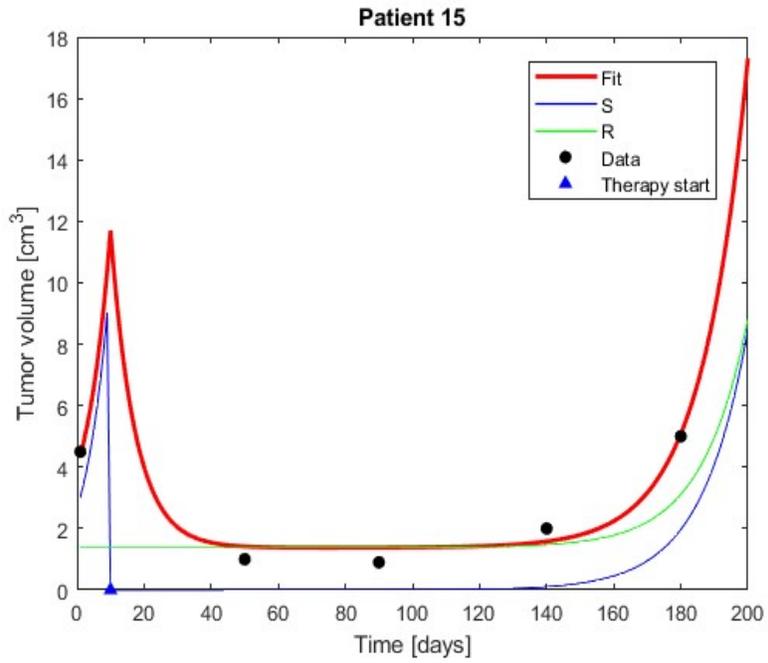












4.1.3 Data Analysis & Clustering

From the fitting plots it is possible to recognize and hypothesize the existence of two group of patients:

- **Group 1.** In late stage therapy, the predicted sensitive population dominates or it's at least comparable to the resistant one [Patients #1, 2, 5, 7, 8, 9, 12, 13, 15, 16].
- **Group 2.** In late stage therapy, the resistant population remarkably dominates over the sensitive one, making the therapy practically ineffective [Patients #3, 4, 6, 10, 11].

Based on this hypothesis, a data analysis has been carried out in order to corroborate this conjecture, also with the purpose of extrapolating *a priori* information which may be useful. However, model parameters in this case have to be handled carefully: cancer actually presents a very complex heterogeneity and the proposed two compartments (resistant and sensitive) are clearly an artifact of this scientific investigation which cannot be experimentally quantified (for example the initial resistant and sensitive population and their growth rates). Nevertheless, a total tumour growth rate could realistically be estimated:

$$\lambda_{tot} = \lambda_s + \lambda_r$$

along with hypothetical tumour survival rate S and transfer rate p .

In Figure 4.4, the distributions of the three investigated parameters (λ_{tot}, S, p) are reported. It can be clearly seen that patients assigned to group 2 are characterized by large p values: they are actually those whose cancer resistance develops abruptly.

Figure 4.5 shows each pair of parameters. A weak linear statistical correlation emerges between p and λ_{tot} . One may carefully infer from this model that a glioma with sharper growth could also present a tendency to evolve toward its resistant counterpart.

Figure 4.6, where all the parameters are gathered, shows clearly the existence of the two predicted clusters, making likely the initial conjecture.

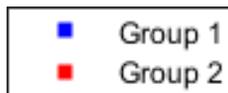


Figure 4.4: Plot legend of section 4.1.3

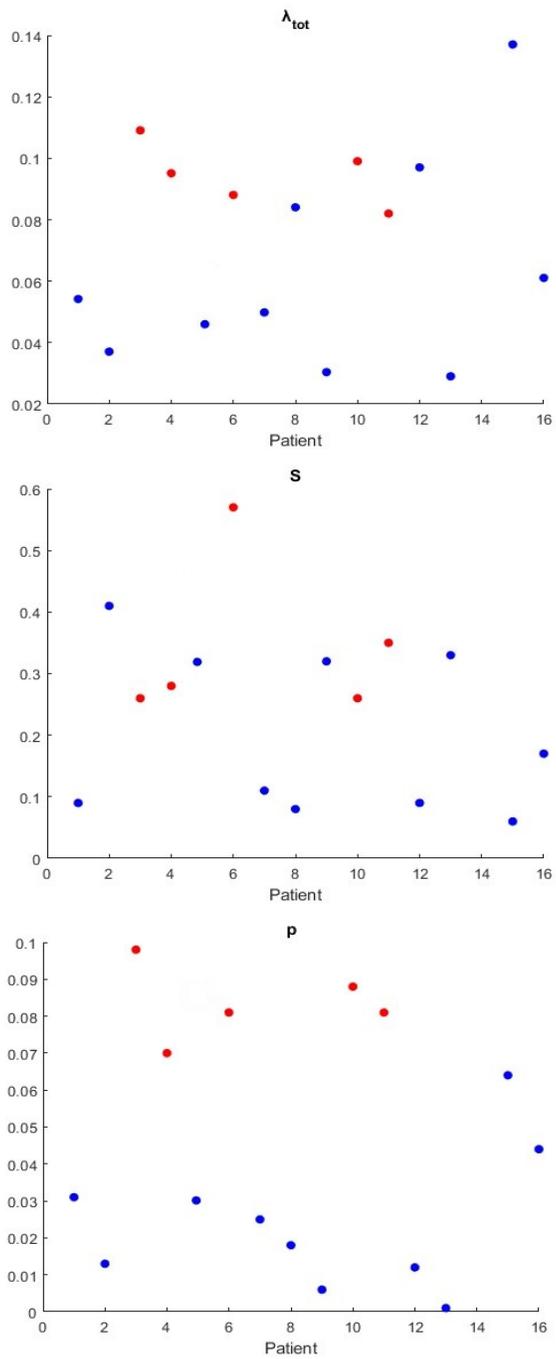


Figure 4.5: Single-parameter distributions. A distinction between the groups is evident for p .

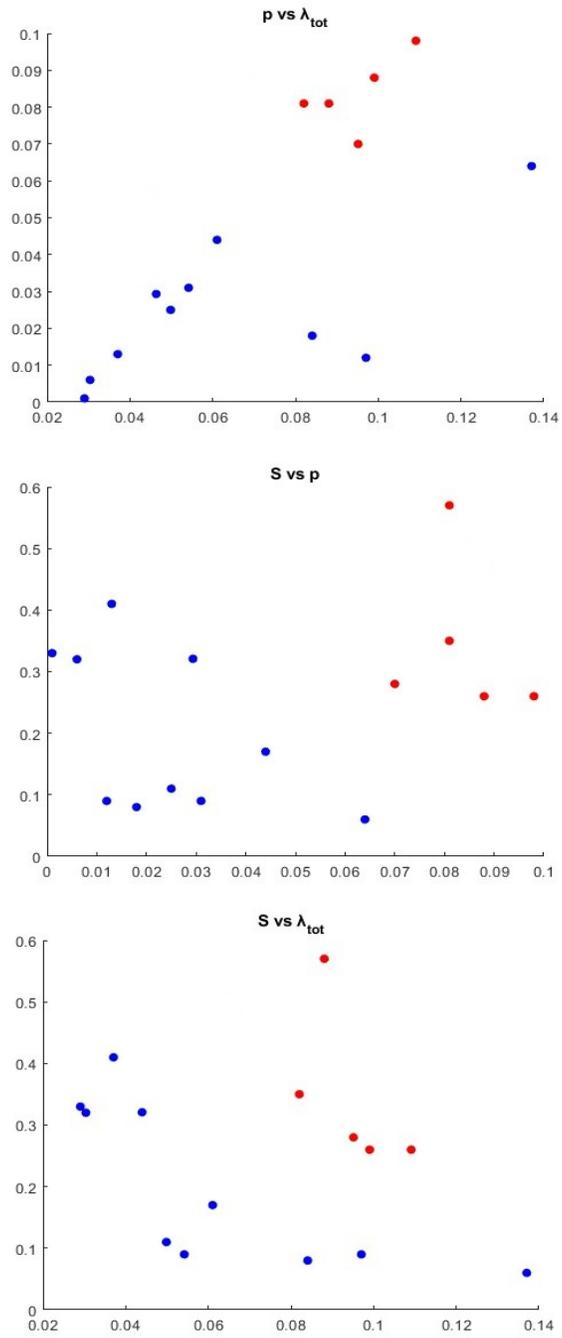


Figure 4.6: Parameter pairs. A slight linear correlation between p and λ_{tot} emerges.

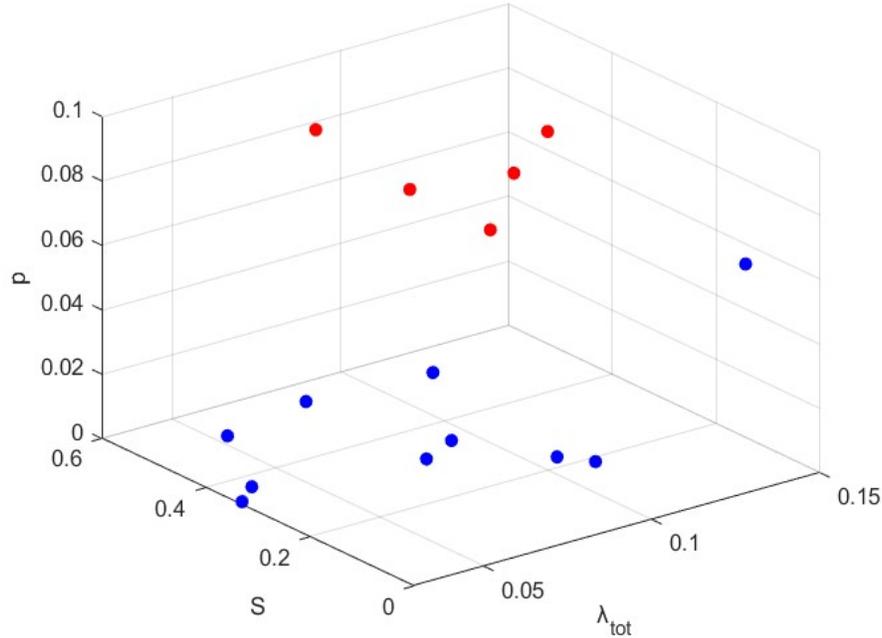


Figure 4.7: Scatter plot over all parameters. The two predicted clusters clearly show up. The rightmost point is that related to patient 15, which behaves as an outlier.

4.2 Maximizing Clonal Inversion Time

Model fitting predicts a common behavior: early sensitive population is readily outnumbered by the resistant one as a result of HFSRT (competitive release). In some cases, as time goes on, resistant cancer cells heavily dominate, making eventual further radiotherapy intervention almost worthless (as for patients belonging to group 2).

In this section, the usual alternative fractionation schedules are proposed, aiming to maximize again the patient survival time (PST) and to delay as much as possible the moment where early clonal inversion occurs, defined as *Clonal Inversion Time* (CIT). The hope of this strategy is to see if by increasing CIT it is possible to prevent (or at least to delay as much as possible) a long-term domination of the predicted resistant mass.

A Genetic Algorithm is employed to predict the best fractionation, where the fitness of each chromosome which has to be maximized is represented by the sum of the two following objective functionals:

$$PST(\mathbf{c}) = \int_{t_0}^{t_{cut-off}(\mathbf{c})} d\tau \quad (4.13)$$

$$CIT(\mathbf{c}) = \int_{t_0}^{t_{CIT}(\mathbf{c})} d\tau \quad (4.14)$$

So, the goal is to maximize $PST(\mathbf{c}) + CIT(\mathbf{c})$ by finding some control vector $\mathbf{c} = [c_1, c_2, \dots, c_n]$ (As in chapter 3, $n = 5$ for IRT, $n \in [6, 7, 8, 9, 10, 11, 12]$ when hyperfractionating) where $c_i \in \{4, 5, 6, 7, 8, 9, 10 \text{ weeks}\}$

4.2.1 IRT

The intermittent approach didn't lead to encouraging results. Although the early clonal inversion is in almost every case delayed, late domination of resistance population for patient of group 2 is not eradicated and, even worse, PST is often remarkably reduced.

For what concerns group 1, only patients 13, 16 noticeably benefit from the intermittent protocol. In all other cases, PSTs are often shortened with respect to the HFSRT case, without any improvement of late populations' behavior.

IRT simulations are reported below along with the legend.

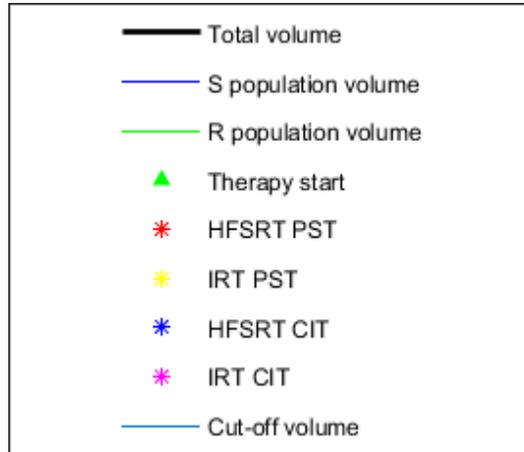
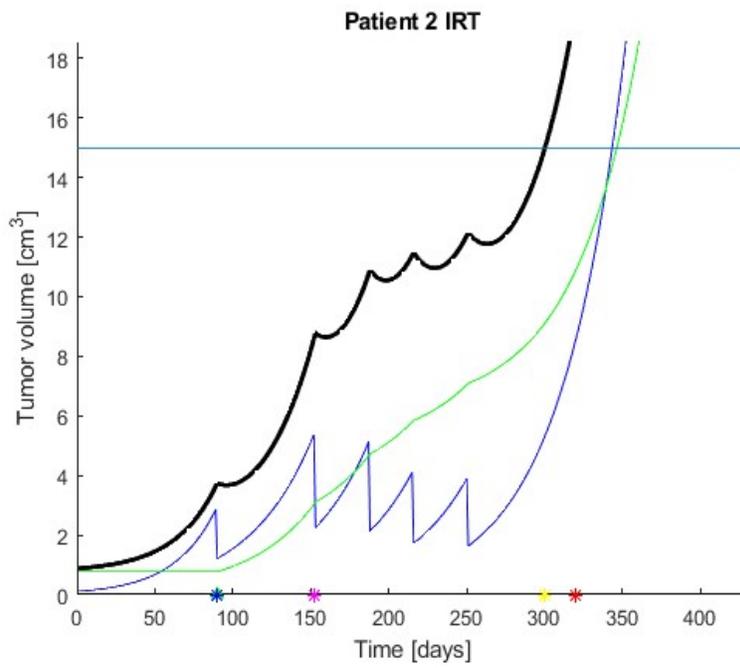
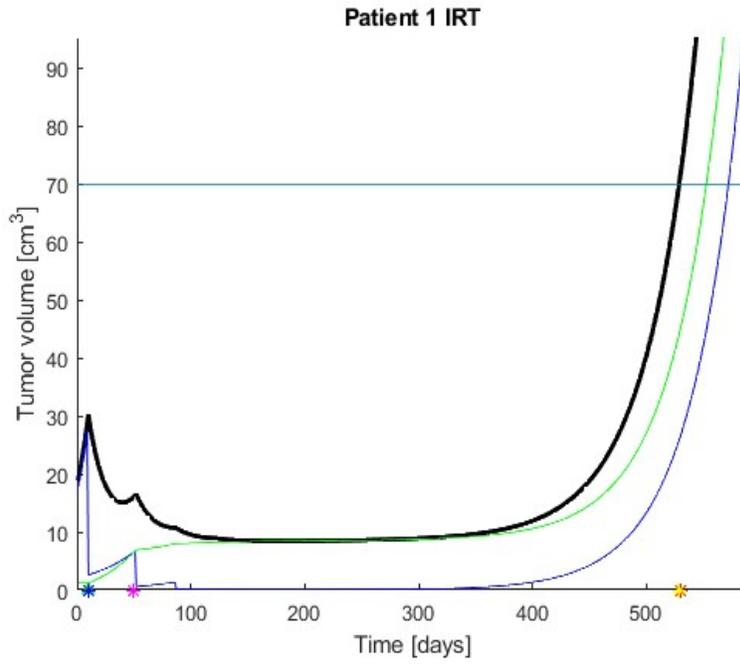
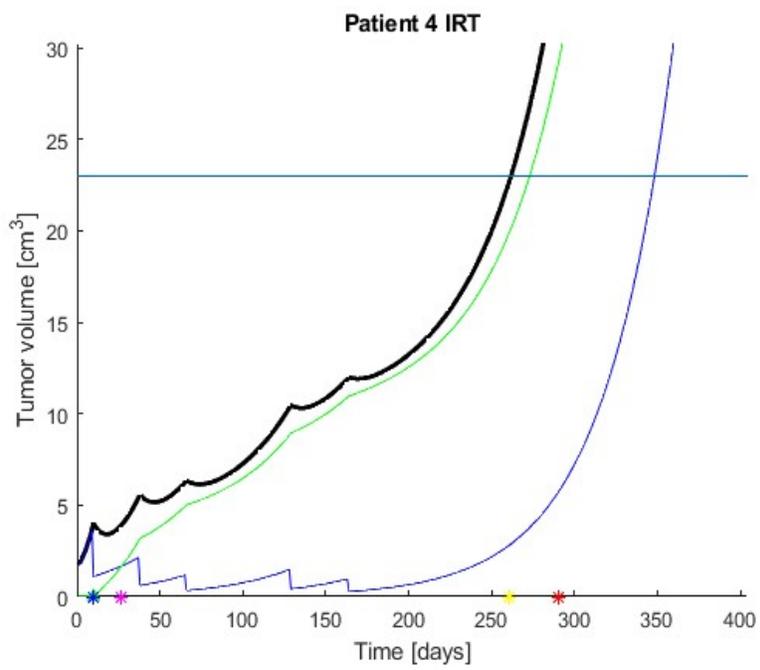
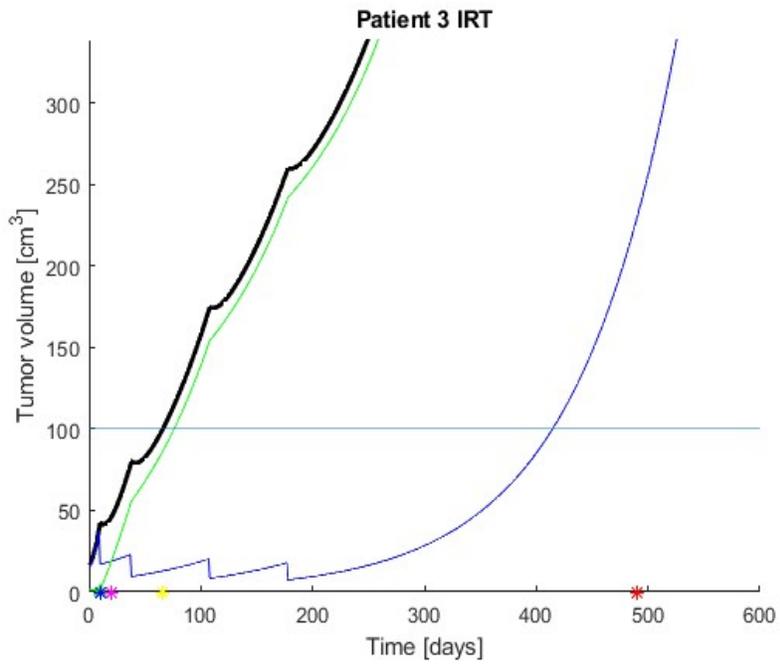
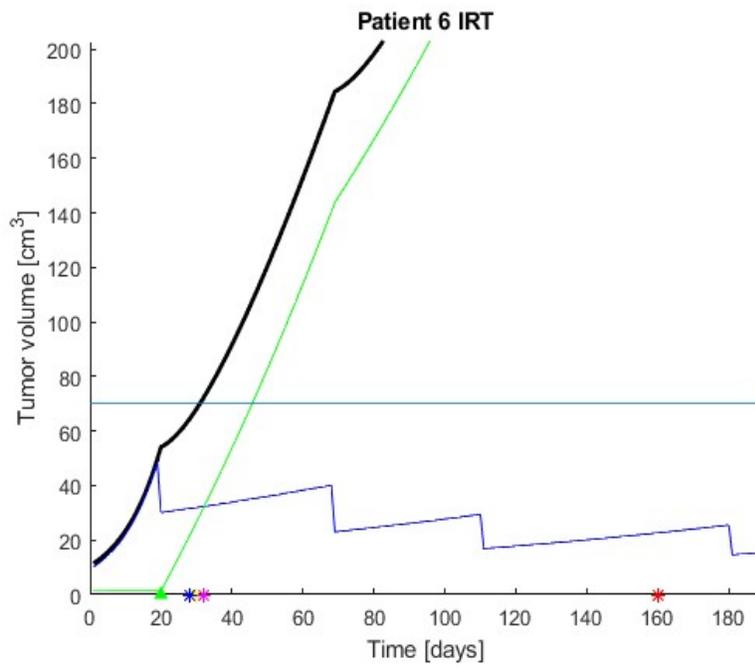
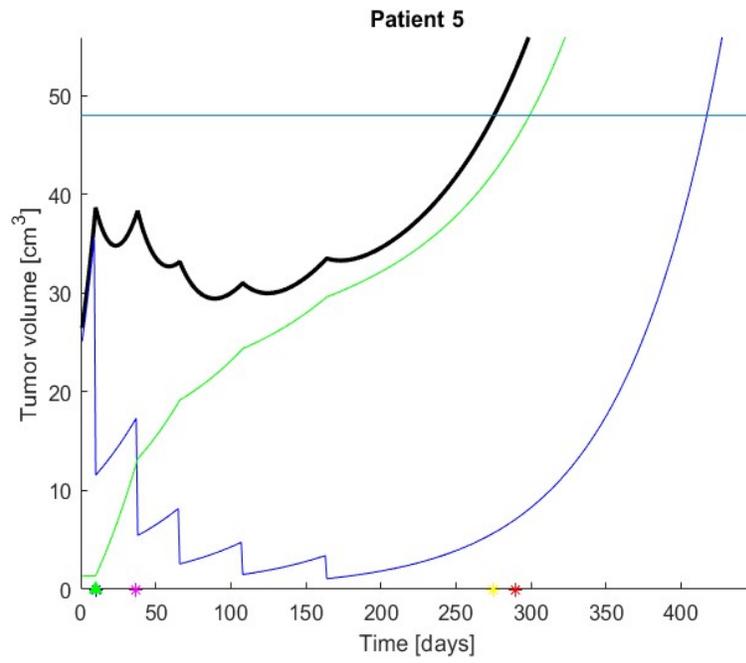
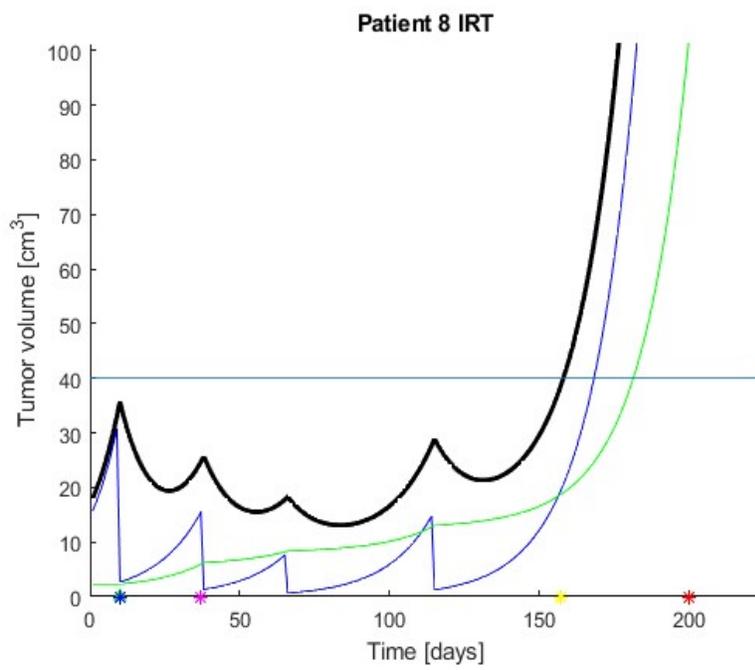
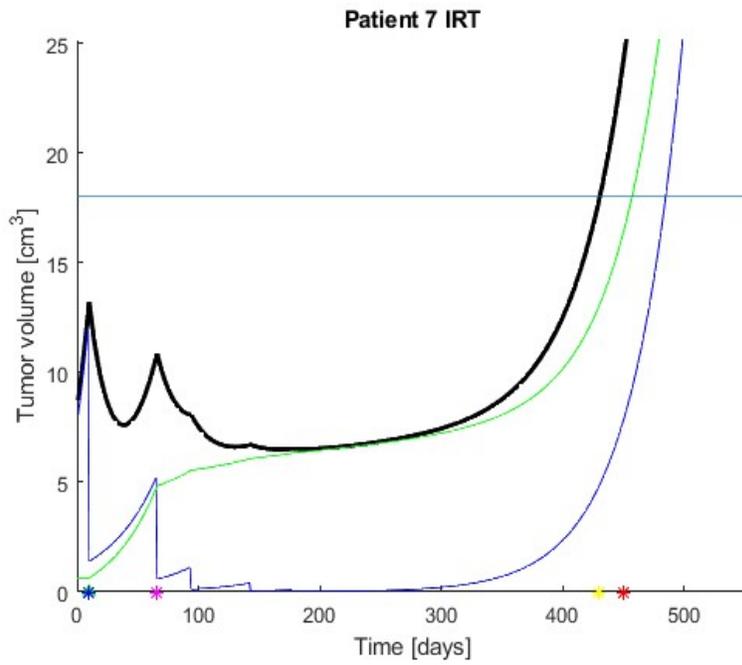


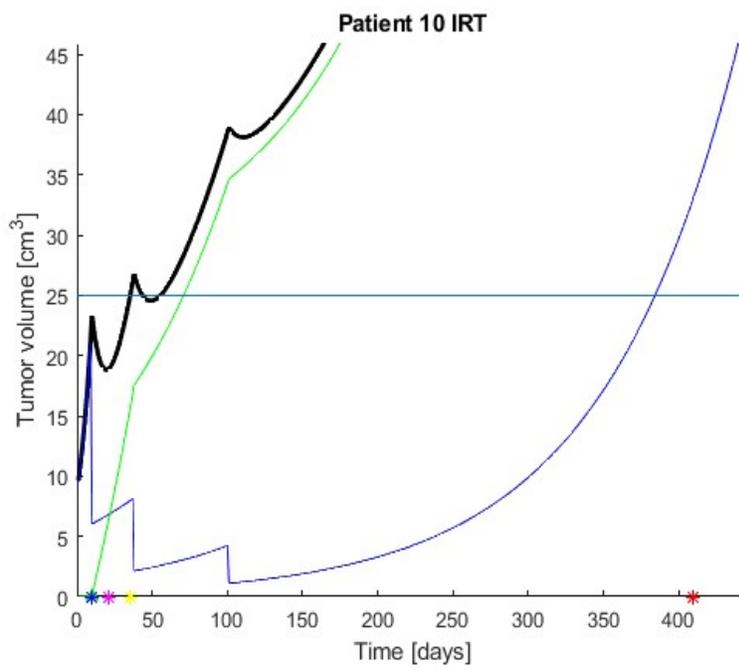
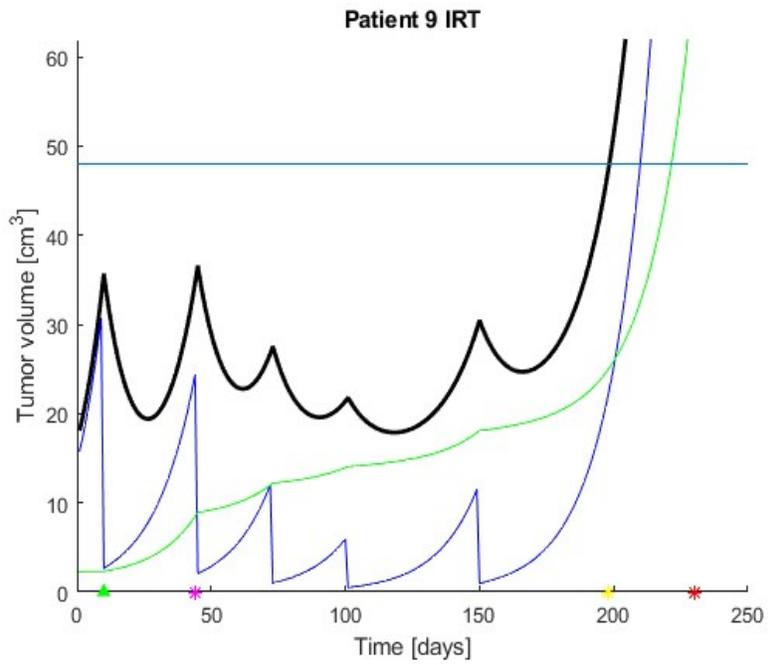
Figure 4.8: Legend of the fitting plots

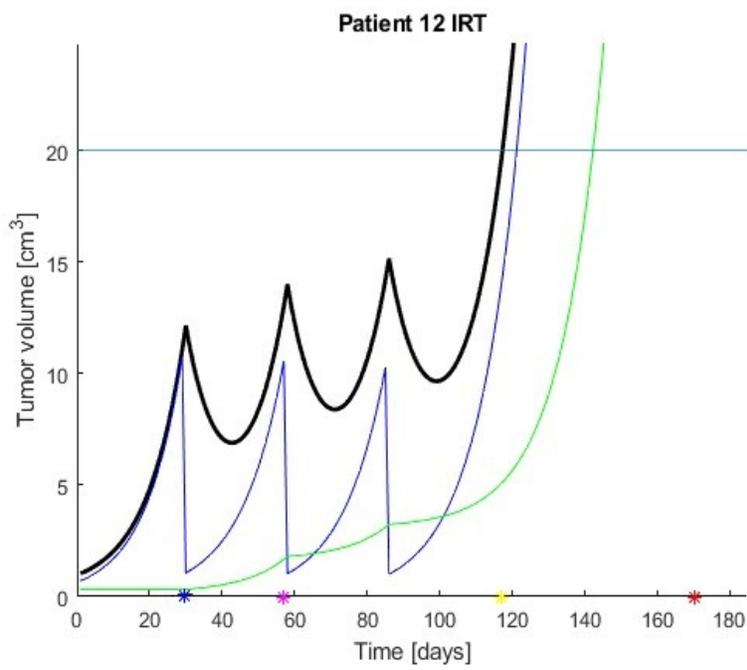
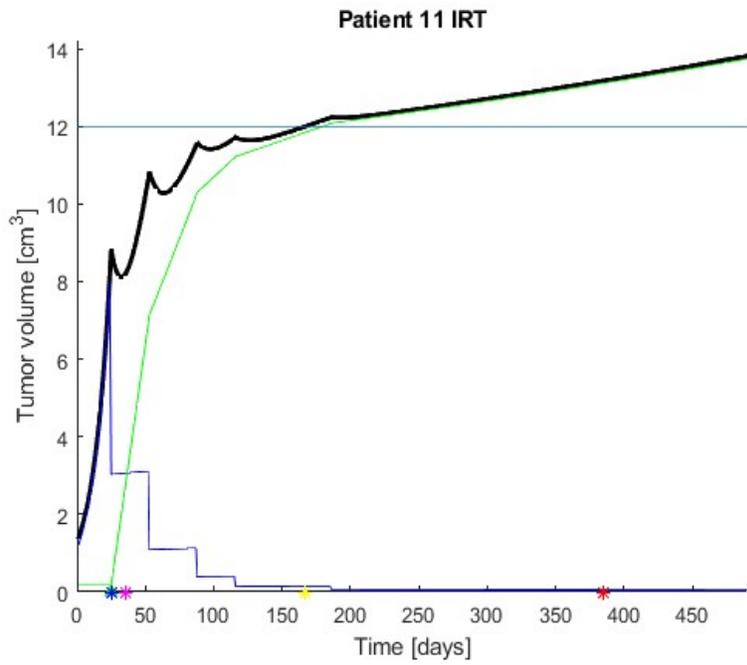


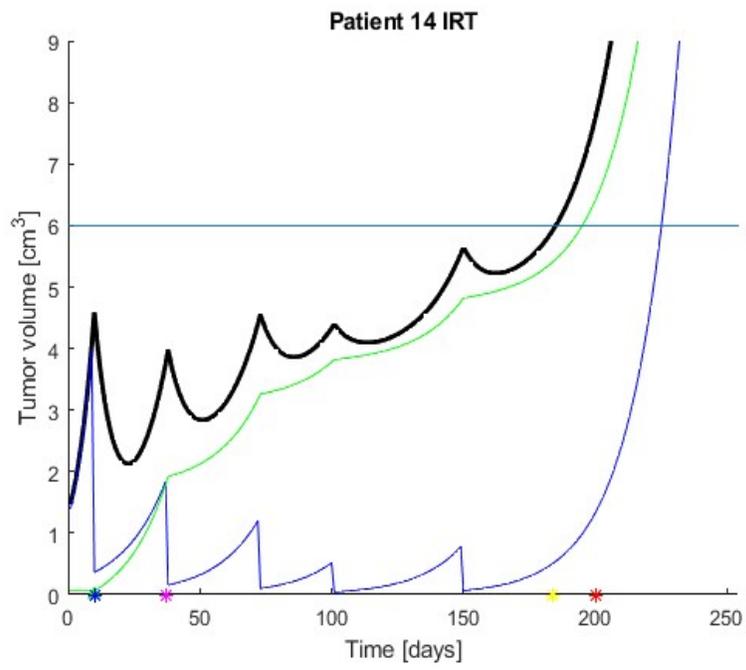
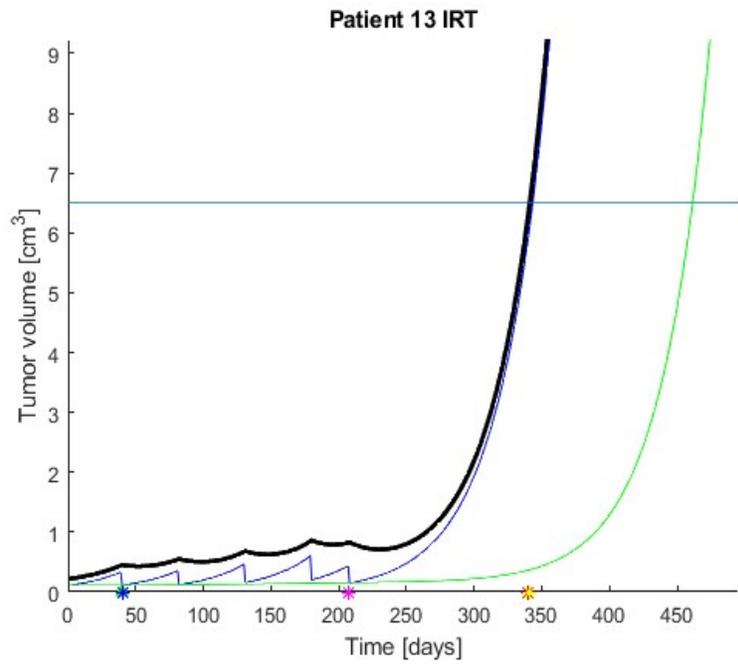


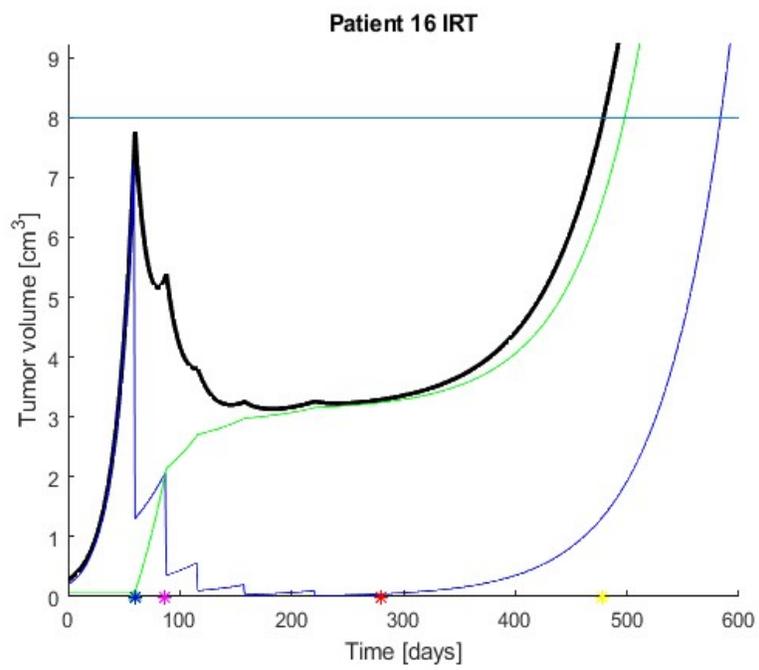
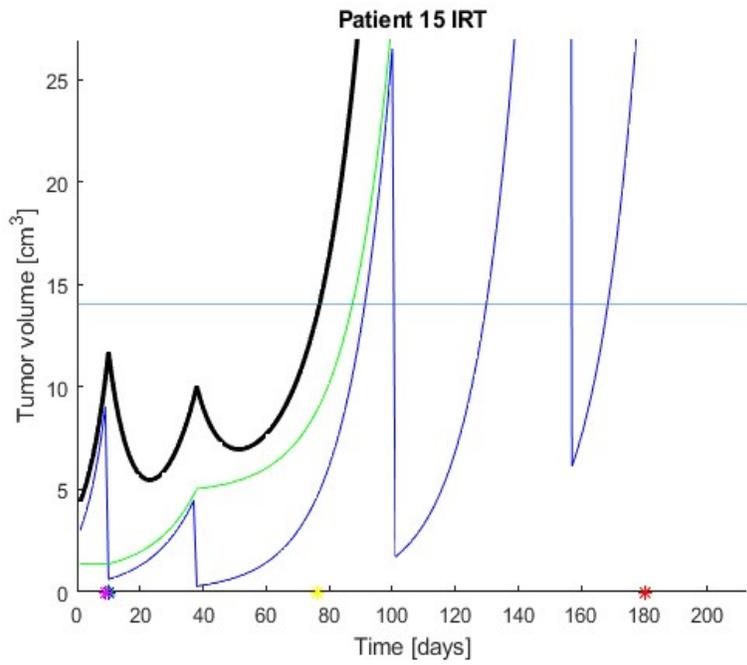












4.2.2 Hyperfractionation

The hyperfractionation strategy proposed in chapter 3 is once again exploited to check if improvements with respect to IRT are possible. Previous results further validated the hypothesis of dividing the patient cohort in two main groups: indeed, the response to IRT of the members was similar, in particular for those of group 2. So, for the sake of clarity, five representative patients have been selected to undergo the simulated hyperfractionation protocol:

- **Patient 1 (group 1)**, for which IRT increases CIT and with a PST similar to the HFSRT case. When reaching the cut-off volume, the resistant population slightly dominates.
- **Patient 6 (group 2)**, for which IRT fails completely.
- **Patient 9 (group 1)**, for which IRT leads to an increased CIT but with a visible, although reasonable, reduction of PST with respect to HFSRT. When reaching the cut-off volume, the sensitive population slightly dominates.
- **Patient 15 (group 1)**, which is clearly an outlier: it's the only case of group 1 for which IRT fails completely as for those patients of group 2. The anomalous feature is further confirmed from the clustering in Figure 4.6 (Patient 15 is the rightmost, isolated data point in the scatter plot)
- **Patient 16 (group 1)**, the one who benefits the most by IRT over the whole patient cohort.

The outcome is the same for each patient: hyperfractionation leads to a massive PST reduction. While IRT in some few cases could lead to a slight improvement, hyperfractionation seems to be, at least in this model, not effective. The related plots are reported below.

The answer to the initial question of section 4.2 about whether dealing early clonal inversion induces trend reversal of the final populations or not is clearly negative. One may wonder if investigating other aspects, for example by changing the objective functional and its weights, could lead to different schedule predictions and reduce as much as possible the raising resistant population. The answer may be positive; however, although this simple multicompartmental model allows for a first rough estimation of cancer heterogeneity and resistance phenomenon, it doesn't include a crucial aspect of cancer dynamics: cancer cells of different nature have different fitness responses to their shared environment and the fitness of a neoplastic cell is shaped by its interactions with cells and other factors in its microenvironment (its *ecology*) [11].

In the present model, a constant transfer rate p increases the resistant population proportionally to the size of the sensitive population, which is on average

larger when switching to an intermittent radiation protocol. It is a fact that the sensitive population is expected to have some fitness advantage over the resistant population [7]. This well-known biological feature will be taken into account in section 4.3, where competition between species is introduced.

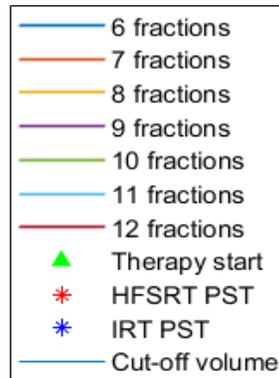
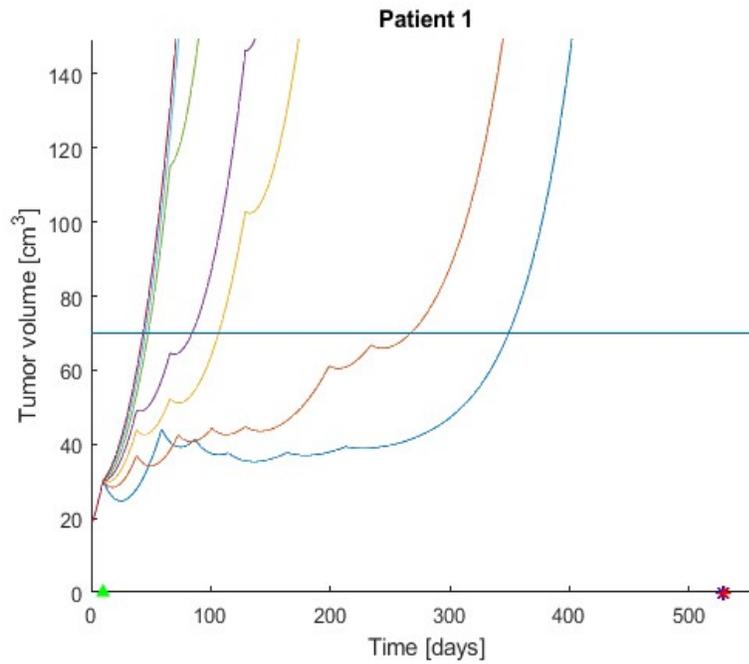
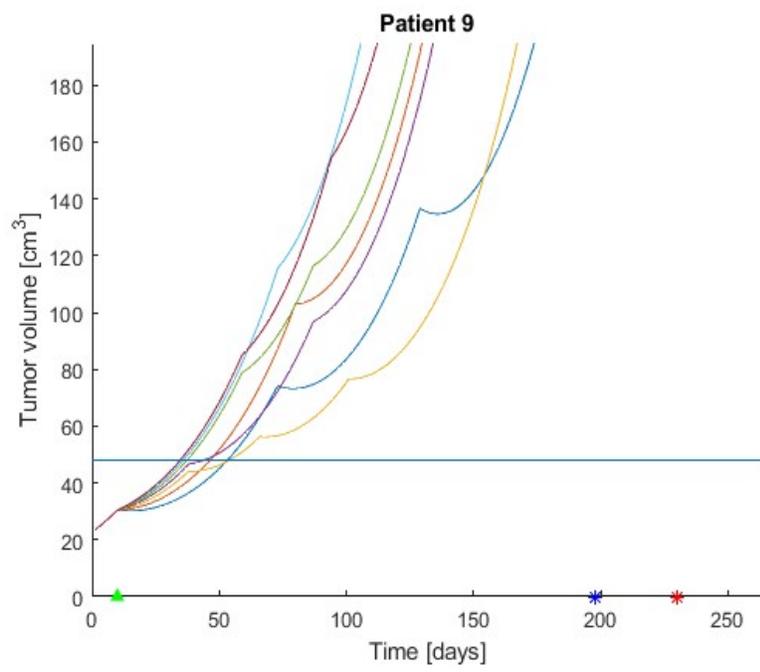
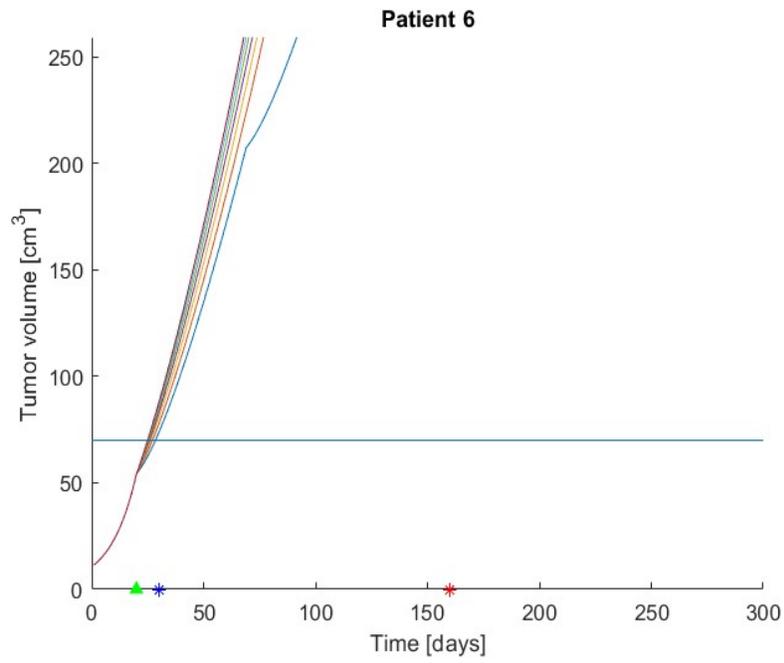
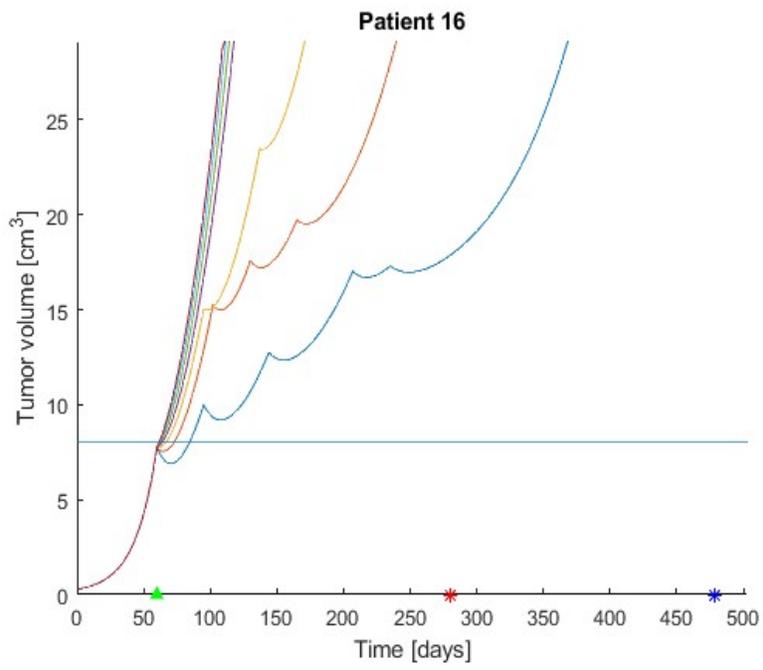
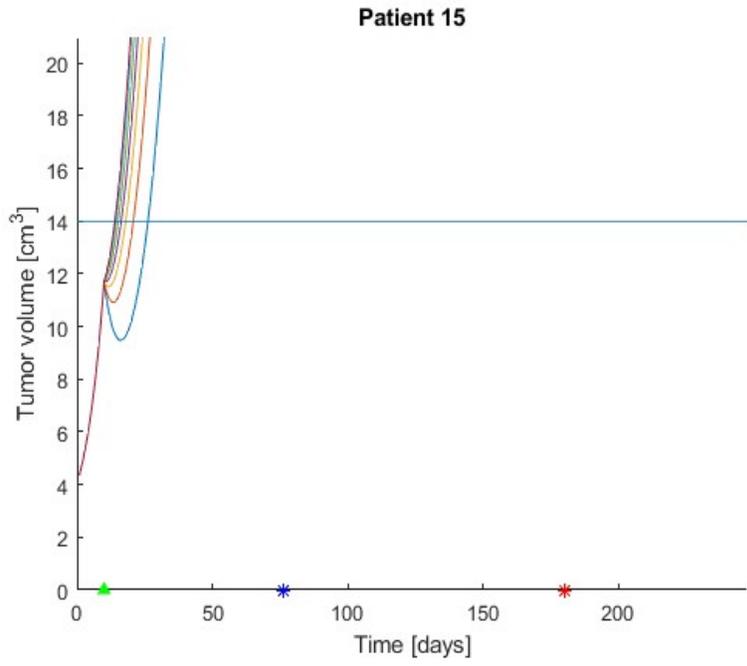


Figure 4.9: Legend of the hyperfractionation plots







4.3 Competitive models

The presence of clonal competition is an unavoidable fact of cancer biology. For neoplastic cells in a heterogeneous population, competition exists in the form of resource consumption (oxygen for example). However, neoplastic clones can also have direct negative effects on each other. For example, one clone can stimulate an immune response that clears other clones [11].

In the absence of treatment, one can infer that resistant cells are less fit than sensitive cells, as untreated cancers generally have a preponderance of cells that are sensitive to primary therapies. In controlled studies, it can be observed that some resistance mechanisms do indeed have a fitness penalty in which the resistant clones grow slower than the parental sensitive cells. This is probably related to resource allocation to resistance mechanisms which would reduce the energy available for proliferation [7].

4.3.1 Including intraspecific competition

At first, intraspecific competition has been introduced within the multicompartamental model presented in section 4.1 (equations (4.3), (4.4), (4.5) are left unchanged):

$$\frac{dV_s(t)}{dt} = \lambda_s V_s(t) - p V_s(t) - \lambda_s \frac{V_s(t)^2}{K_s} \quad (4.15)$$

$$\frac{dV_r(t)}{dt} = \lambda_r V_r(t) + p V_s(t) - \lambda_r \frac{V_r(t)^2}{K_r} \quad (4.16)$$

where K_s and K_r are the carrying capacities [*volume*] of the sensitive and resistant population respectively, they are both positive constants.

Even though it makes sense to expect this internal competition mechanism, it is straightforward to note roughly with a quick glance from the data set that the characteristic S-shaped curve of this logistic growth doesn't fit the available data. Indeed, when the genetic algorithm has been used to estimate model's parameters, it predicted very large values for K_s, K_r , making the newly introduced terms in equations (4.15),(4.16) definitely negligible.

Moreover, since the model is non-linear, the computational cost of the fitting stage increased noticeably (the genetic algorithm makes use of the MATLAB R2022a function `ode45()` solver in the numerical integration of the ODE system). For both reasons, including intraspecific competition has been deemed worthless in terms of mathematical modelling and therefore in the following this aspect, which anyhow is physically present, has been neglected.

4.3.2 Including interspecific competition

On the basis of the discussion at the very beginning of this chapter, interspecific competition between the two species in the multicompartamental model introduced in section 4.1 has been included (equations (4.3), (4.4) and (4.5) are unaltered):

$$\frac{dV_s(t)}{dt} = \lambda_s V_s(t) - p V_s(t) - \lambda_s b_{sr} V_s(t) V_r(t) \quad (4.17)$$

$$\frac{dV_r(t)}{dt} = \lambda_r V_r(t) + p V_s(t) - \lambda_r b_{rs} V_s(t) V_r(t) \quad (4.18)$$

Where the new positive constants b_{sr} and b_{rs} describes respectively the competitive effect over the sensitive population due to the resistant one and vice versa. They are dimensionally [$volume^{-1}$].

4.3.3 Model fitting

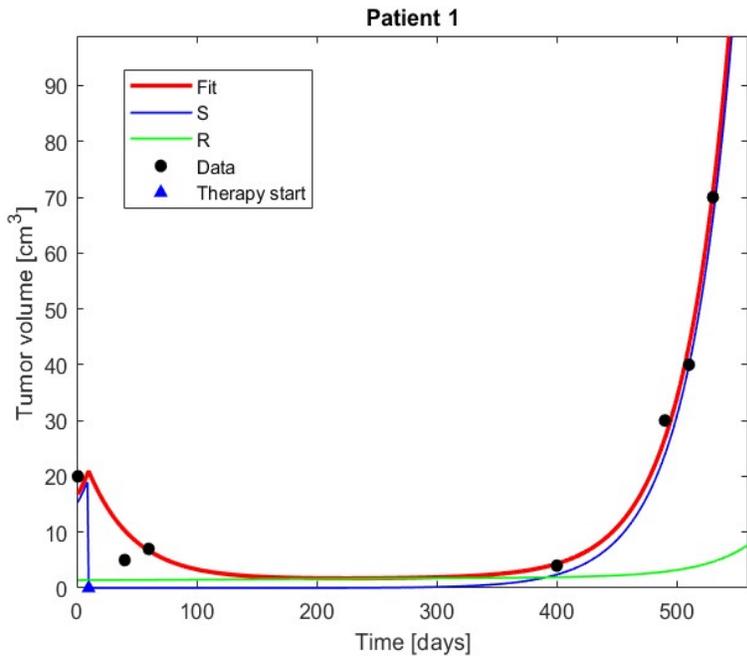
Following the usual work path, parameters fitting has been performed by means of a Genetic Algorithm (supported by the MATLAB R2022a function *ode15s()* to integrate the non-linear system of ODEs). The models' parameters are $\lambda_s, \lambda_r, S, V_{s,0}, V_{r,0}, b_{sr}, b_{rs}$ and p , encoded in a bit-string chromosome as shown in Figure 4.10. The length of the chromosome is 99 bits, so the search space has a size of $2^{99} \sim 10^{29}$ possible solutions. Fitness function is again the RMSE in (4.12). In the fitting plots, S and R indicate the volume of the predicted sensitive and resistant population respectively.

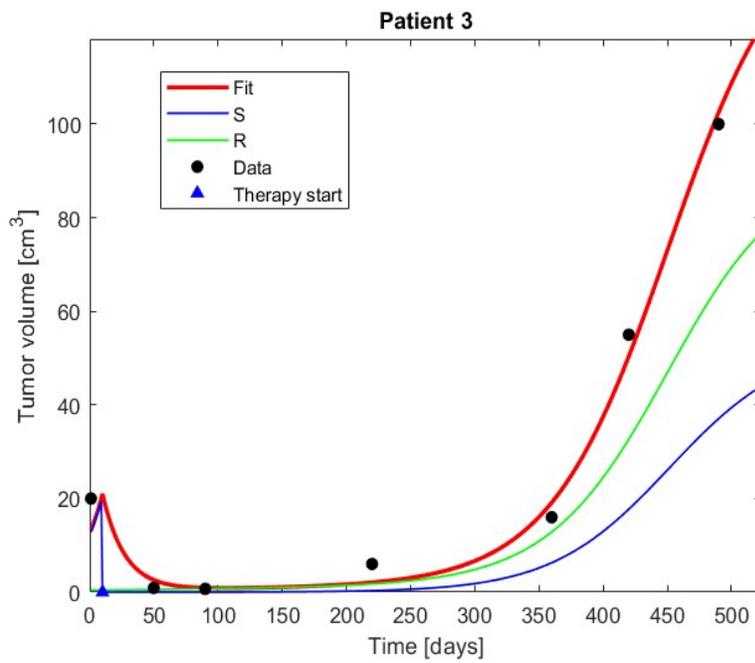
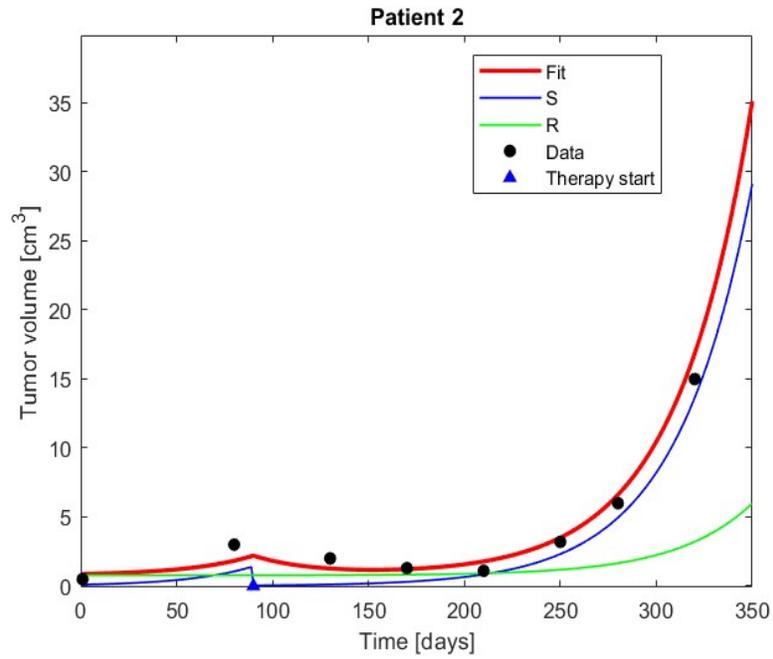
λ_s	λ_r	$V_{s,0}$	$V_{r,0}$	S	p	b_{sr}	b_{rs}
10 bits	10 bits	11 bits	11 bits	7 bits	10 bits	20 bits	20 bits

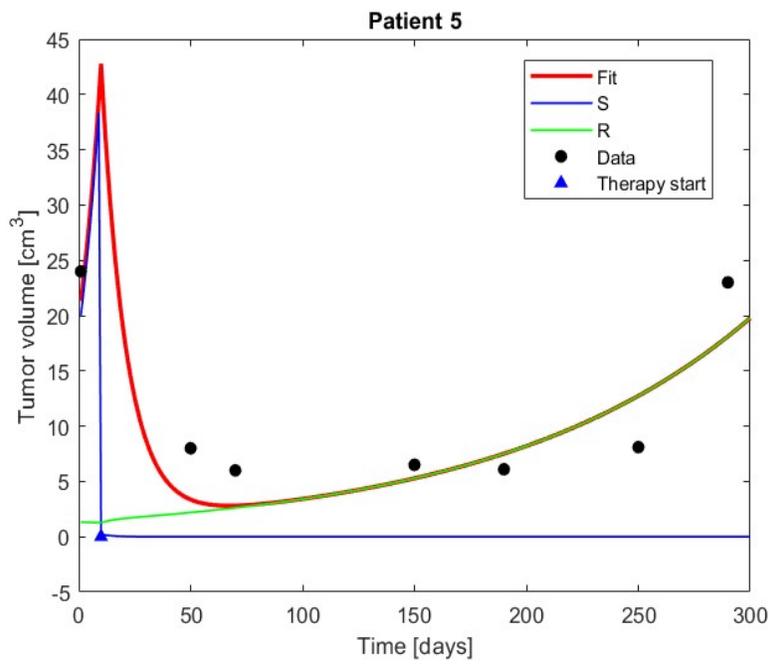
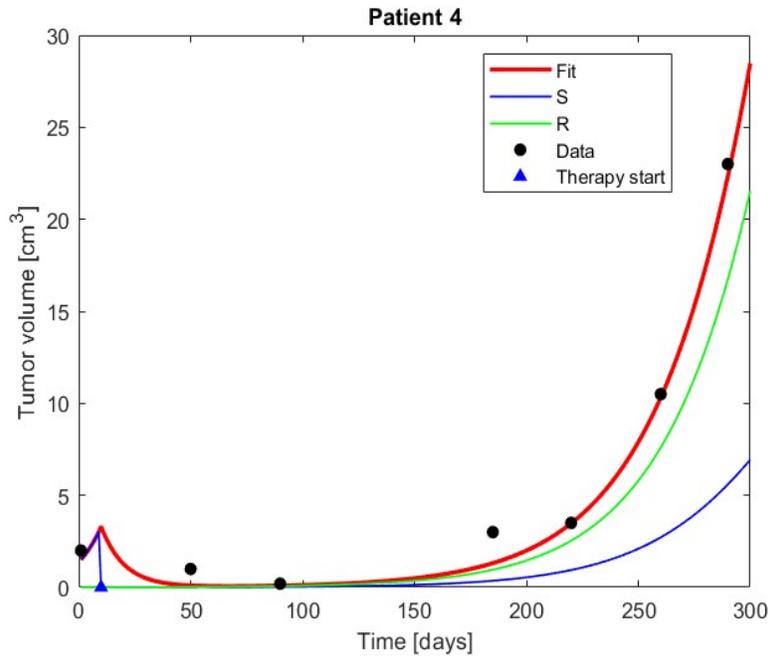
Figure 4.10: Partition of the bit-string chromosome

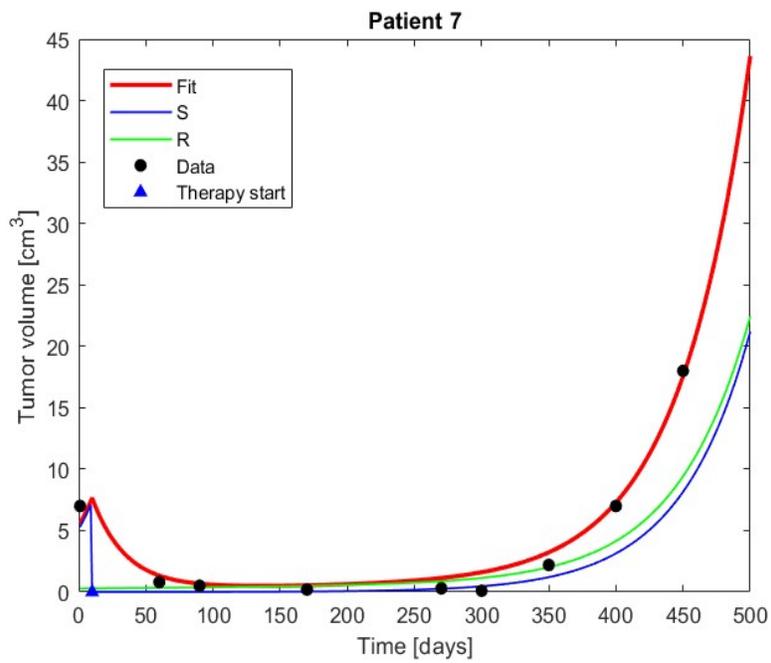
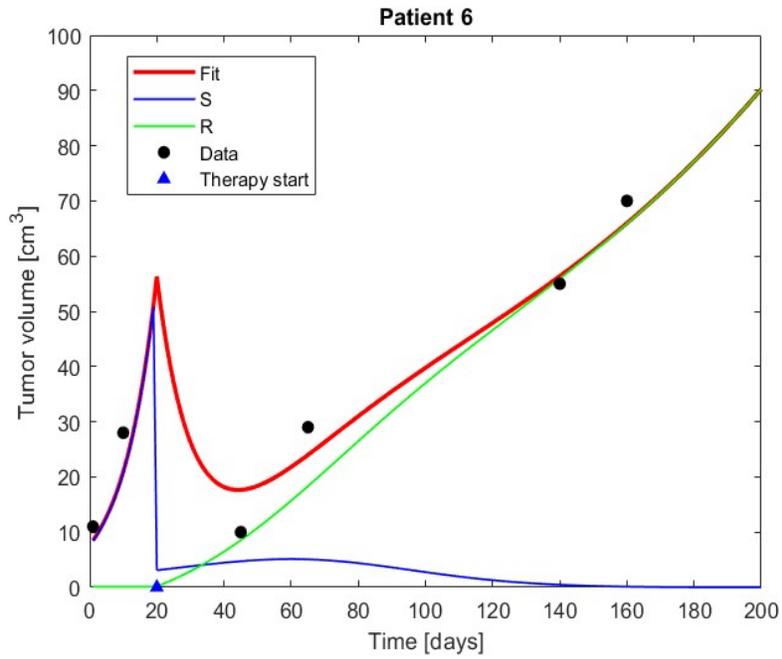
Remarkably, the GA predicted that sensitive cells are almost in every case more competitive than their resistant counterpart (see table below). This is in line with the idea discussed previously: sensitive cells are usually expected to have a fitness advantage and therefore to repopulate the environment in a more effective way.

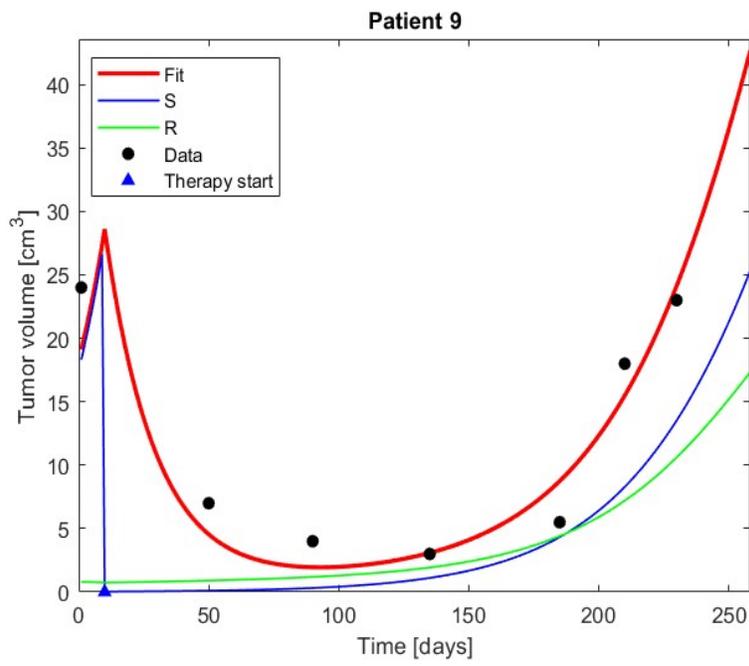
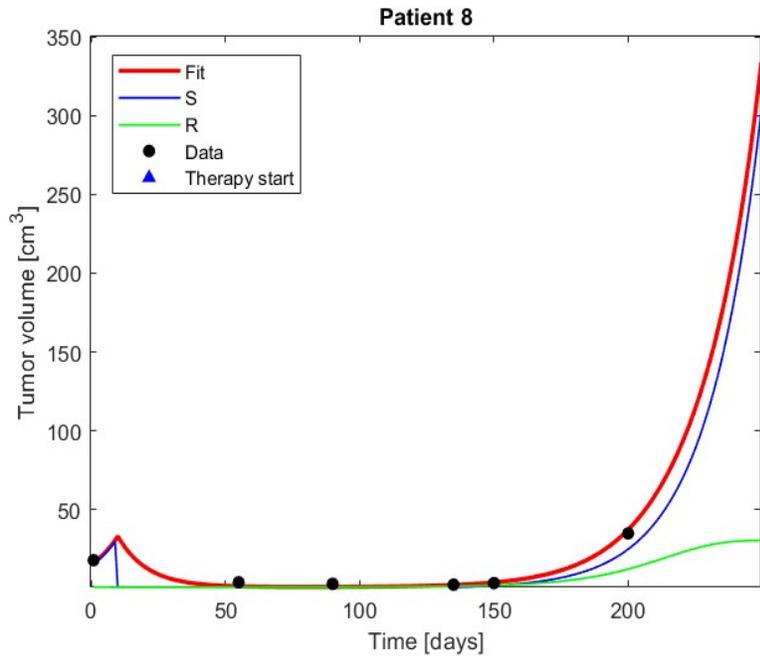
Patient	$b_{sr}[cm^3]$	$b_{rs}[cm^3]$
1	0.000023	0.000094
2	0.000018	0.000031
3	0.0046	0.088
4	0.0039	0.00038
5	0.00011	0.053
6	0.077	0.014
7	0.0000013	0.027
8	0.00092	0.85
9	0.0098	0.15
10	0.0012	1.6
11	0.34	5.2
12	0.056	2.2
13	0.079	0.21
14	0.026	1.4
15	0.016	3.3
16	0.0059	0.55

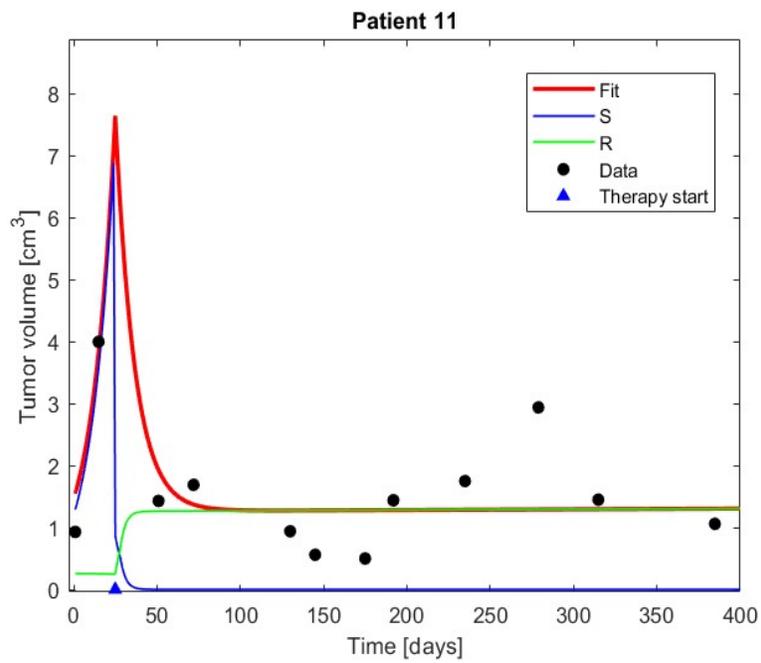
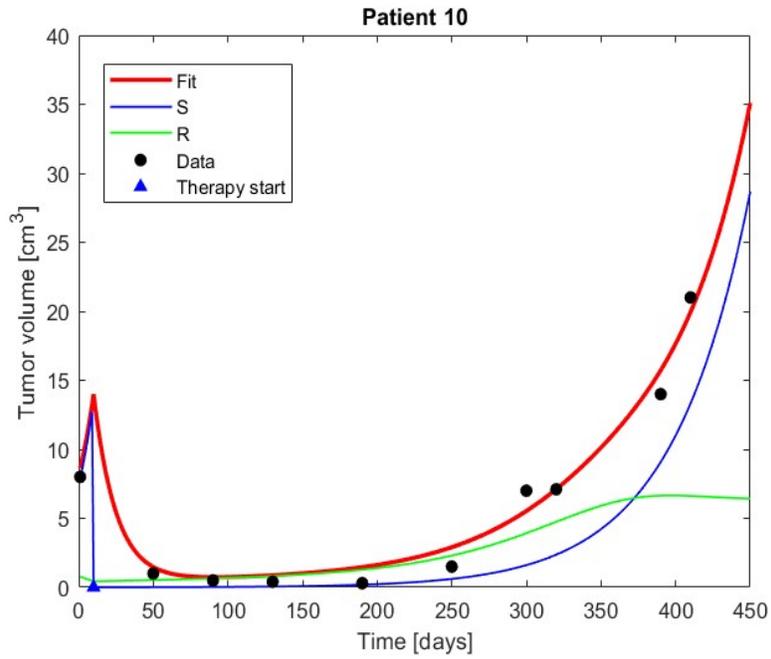


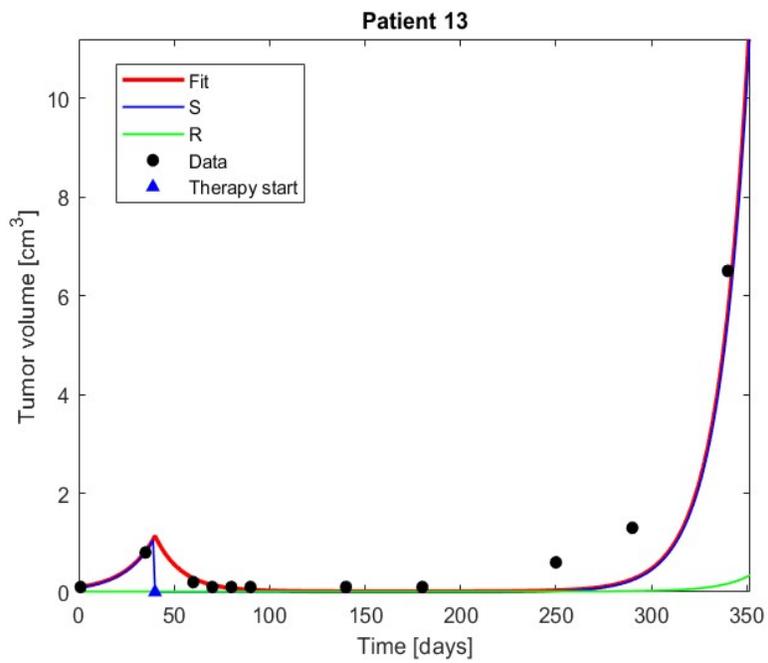
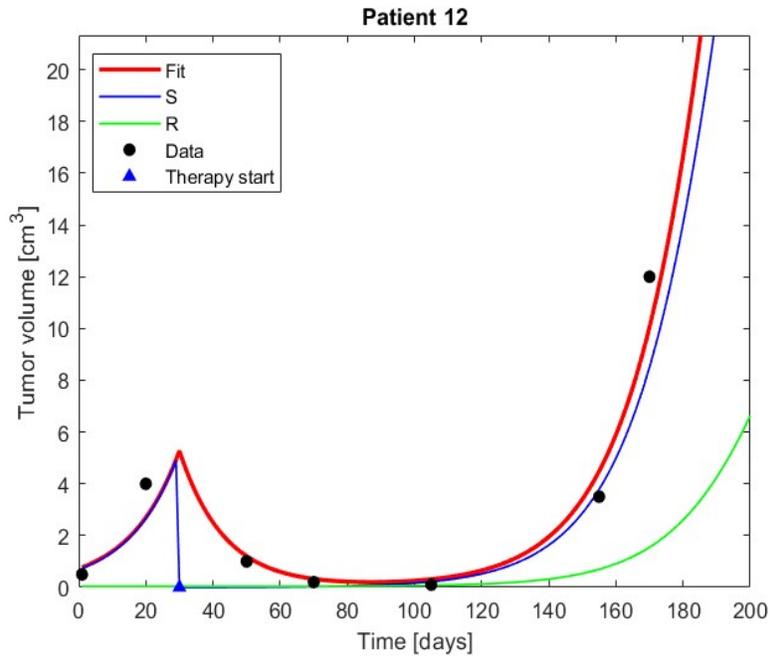


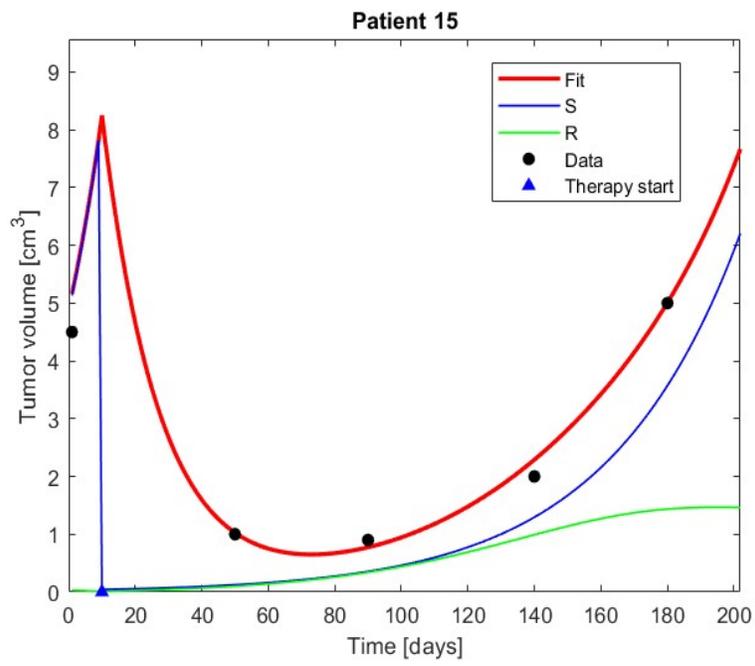
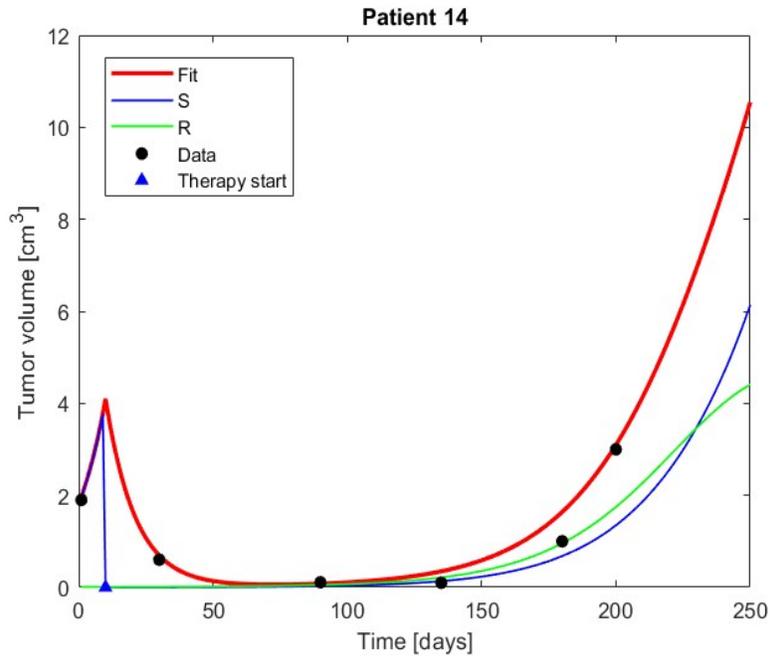


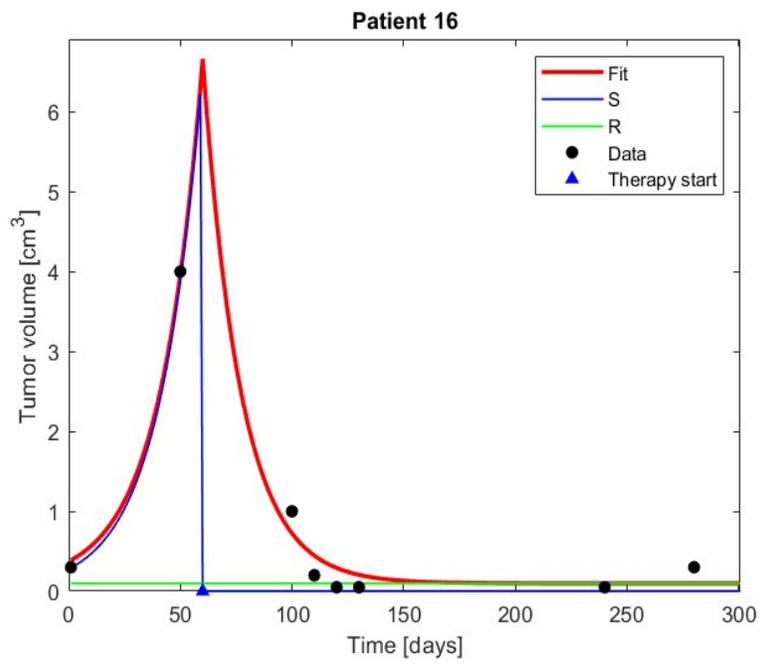












4.3.4 IRT

From the GA fitting, it is possible also in this case to see that a consistent group of patients is predicted to undergo the aforementioned competitive release mechanism. Following what's been done previously, a GA is used to predict the optimal intermittent radiotherapy schedule aiming to increase each patient survival time (PST) and, hopefully, to delay resistance development as much as possible.

Patients #1, 3, 4 have been selected to compose a sample for this study. IRT prediction plots are reported below.

- Patient 1 didn't particularly show the competitive release phenomenon as a consequence of HFSRT. The intermittent protocol didn't alter significantly this behaviour, but increased its PST.
- For patient 3 the HFSRT GA predicted an evident competitive release. However, IRT allows sensitive cells to exploit their larger competitiveness to overcome resistance raising and, noticeably, to induce a relapse of the resistant population, with an increase of patient's PST.
- Patient 4, just as patient 3, showed up competitive release within the HFSRT frame. However, in this case the GA predicted that resistant cells have a competition parameter value higher than that of sensitive ones. As a consequence, IRT led to an even more accentuated competitive release with respect to the hypofractionated case.

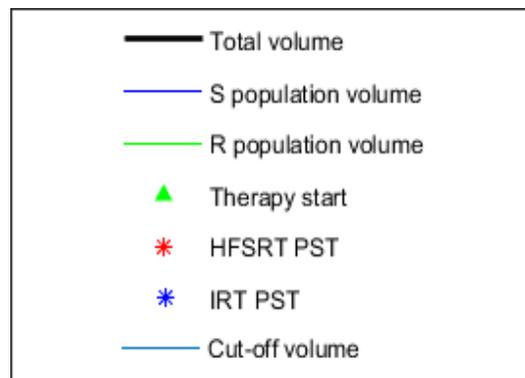
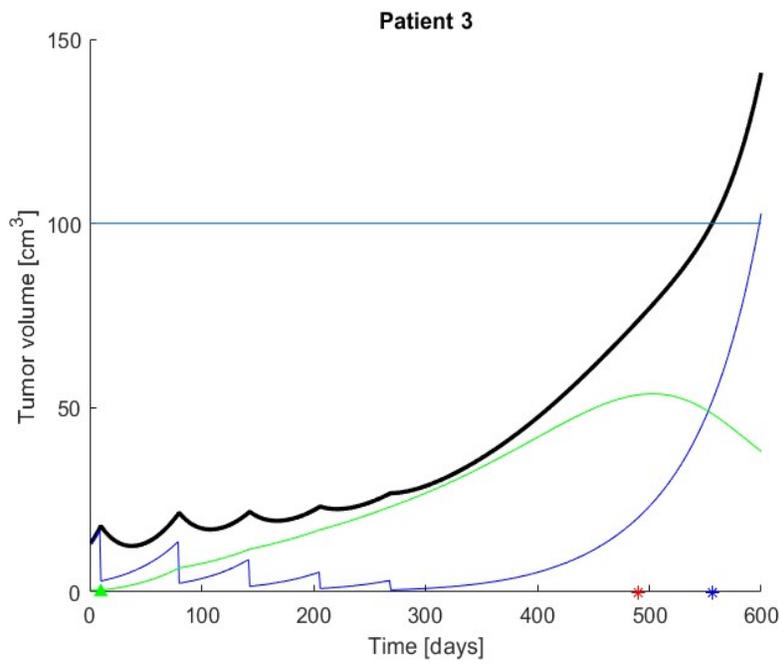
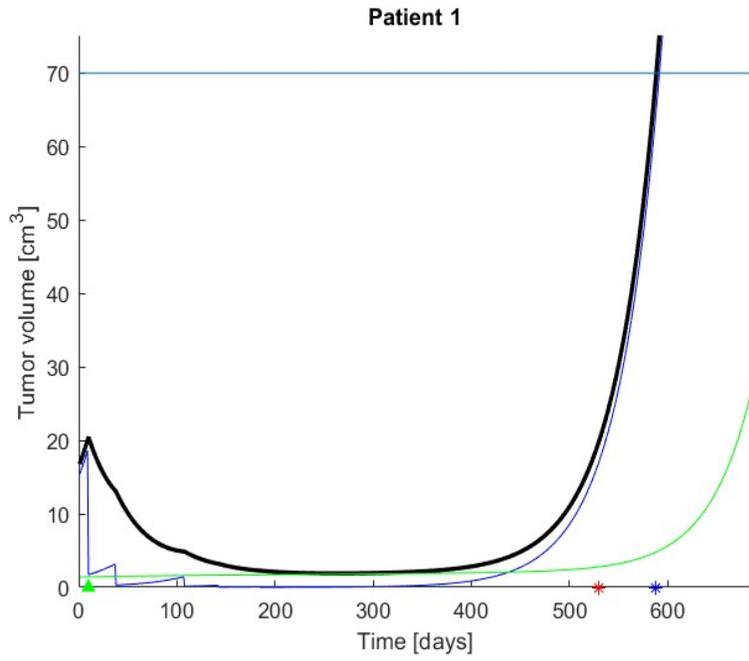
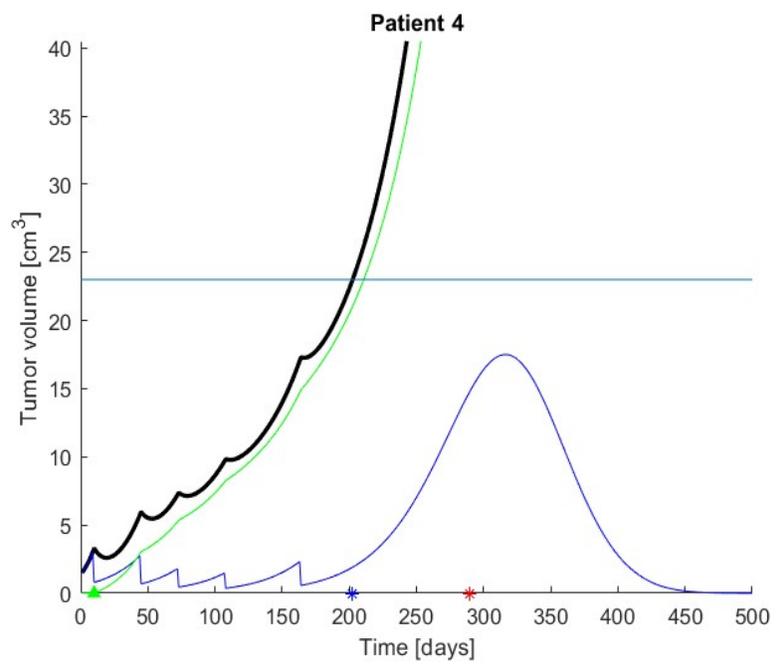


Figure 4.11: Plot legend of section 4.3.3





Conclusion

The heterogeneous physics of recurrent high-grade glioma, just like many other tumours in general, is endowed by several complex features (articulate spatial distribution, biological nature of cells and surrounding tissues, stochasticity, environmental noise, etc. . .). Mathematical models of great complexity are usually exploited to include many of those aspects.

In this thesis, modelling complexity has been reduced to the minimum. Nevertheless, crucial features like therapy resistance and clonal competition were included effectively, providing a rough but reasonable description.

A possible immediate extension of the proposed models may provide a more sophisticated description of the resistance mechanism, for example by including a non-constant resistance development, which realistically is expected to emerge slowly, afterwards reaching some asymptotic value.

Moreover, genetic algorithms turned out to be very effective in parameter estimation of this kind of models. Not only, they allowed within a certain extent to automate the seek of optimal radiotherapy treatments, allowing a high degree of patient's therapy customization.

Results of this study suggest that in some cases an intermittent radiotherapy approach may help clinicians in their struggle against a definitely hard-to-deal disease such as rHGG, for example by delaying as much as possible the well known competitive release phenomenon by exploiting, whenever possible, the fitness advantage of radiation sensitive cells, which typically dominate within the highly heterogeneous scenario as that of cancer.

The choice of a certain personalized therapy can be further improved by gathering more data and not only MRI scans just like in the proposed database. For example, patient-specific biological and medical information combined with the tools of mathematical modelling would lead to a higher treatment quality and, possibly, to a better final outcome.

Appendix A

GA template for model parameters' fitting

```
tic
%prepare run
clear
clc
close all
%select operation
fitting=true;
IRT=false;

if fitting==true
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    % GA parameters
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    n_generations=1000;
    M=100; % # of population members
    M_mating= M/2;

    %crossover parameters
    p_c=0.8;
    n_crossover= 1000;

    %mutation parameters
    p_m=0.01;
    n_mutation=1000;

    %Quality threshold
    RMSE_target= 1;

    ga_used_parameters=[M M_mating n_generations p_c n_crossover...
        p_m n_mutation ];
end
```

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% fitting through HFSRT
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
treatment= 'HFSRT';
patient= 11;
t_measured=xlsread('dataset.xlsx', patient, 'A1:A12');
vol_measured=xlsread('dataset.xlsx', patient, 'G1:G12');
t_rt_init=xlsread('dataset.xlsx', patient,'B1:B1');

%time range
T_fin=420;
time=1:1:T_fin;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%generating randomly initial population
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
nbit = 21;
population = zeros(M,nbit);
for i = 1:M
    for j=1:nbit
        population(i,j)=randi([0 1], 1);
    end
end

%start GA
while n_generations>0
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    %Selection
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    mb_indices=ones(M_mating, 1);
    mb=zeros(M_mating, nbit);
    fitness=zeros(M,1);
    for i=1:M
        epsilon=((bit2int(population(i,1:7)') ,7))*10^-3)+0.001;
        lambda=((bit2int(population(i,8:14)') ,7))+17)*10^-3;
        S=((bit2int(population(i,15:21)') ,7))*10^-2)+0.01;
        [Vl,Vd,vol]=tumor_volume(lambda,epsilon,S,time,...
            vol_measured(1),treatment,t_rt_init);

        index=0;
        for j = 1:T_fin
            if ismember(j,t_measured)
                index=index+1;
                fitness(i)=fitness(i)+...
                    (((vol_measured(index))-(vol(j))))^2);
            end
        end
    end
    n_generations=n_generations+1;
end

```

```

        end
    end
    fitness(i)=(1/sqrt(length(vol_measured)))*sqrt(fitness(i));
end

if min(fitness)<RMSE_target
    break
end

fitness_tot=sum(fitness);
p=zeros(M,1);
for i=1:M
    p(i)= fitness(i)/fitness_tot;
end

t=zeros(M,1);

for i=1:M
    for j=1:i
        t(i)=t(i)+p(j);
    end
end

already_used=zeros(M,1);
for i=1:M_mating
    r=rand();
    for j=1:M
        if r<t(j) && already_used(j)==0
            mb_indices(i)=j;
            already_used(j)=1;
            for k=1:nbit
                mb(i,k)=population(j,k);
            end
            break
        end
    end
end

end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Recombination
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Crossover
for i=1:n_crossover
    if rand()<p_c
        c1= randi([1 nbit], 1);
    end
end

```

```

        c2= randi([c1 nbit], 1);
        crossover1=randsample(mb_indices,1);
        crossover2=randsample(mb_indices,1);
        for j=c1:c2
            swap=population(crossover1,j);
            population(crossover1,j)=population(crossover2,j);
            population(crossover2,j)=swap;
        end
    end
end

%Mutation
for i=1:n_mutation
    if rand()<p_m
        mutation=randsample(mb_indices,1);
        bit_mutation=randsample(1:nbit,1);
        population(mutation,bit_mutation)~=population(mutation,...
            bit_mutation);
    end
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% calculating mating buffer members' fitness
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
fitness_mb=zeros(M_mating,1);

for i=1:M_mating
    epsilon=((bit2int(population(mb_indices(i),1:7)',7))*10^-3)+0.001;
    lambda=((bit2int(population(mb_indices(i),8:14)',7))+17)*10^-3;
    S=((bit2int(population(mb_indices(i),15:21)',7))*10^-2)+0.01;
    [Vl,Vd,vol]=tumor_volume(lambda,epsilon,S,time,...
        vol_measured(1),treatment,t_rt_init);

    index=0;
    for j = 1:T_fin
        if ismember(j,t_measured)
            index=index+1;
            fitness_mb(i)=fitness_mb(i)+((vol_measured(index)-vol(j))^2);
        end
    end
    fitness_mb(i)=(1/sqrt(length(vol_measured)))*sqrt(fitness_mb(i));
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% new population
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

```

```

for i=1:M_mating
    candidate=randsample(1:M,1);
    if fitness_mb(i)>fitness(candidate)
        for j=1:nbit
            population(candidate,j)=mb(i,j);
        end
    end
end
disp(n_generations)
n_generations=n_generations-1;

end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% best fitting and statistical analysis
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%final population fitness
fitness=zeros(M,1);
for i=1:M
    epsilon=((bit2int(population(i,1:7)',7))*10^-3)+0.001;
    lambda=((bit2int(population(i,8:14)',7))+17)*10^-3;
    S=((bit2int(population(i,15:21)',7))*10^-2)+0.01;
    [Vl,Vd,vol]=tumor_volume(lambda,epsilon,S,time,...
        vol_measured(1),treatment,t_rt_init);

    index=0;
    for j = 1:T_fin
        if ismember(j,t_measured)
            index=index+1;
            fitness(i)=fitness(i)+...
                (((vol_measured(index))-(vol(j))))^2);
        end
    end
    fitness(i)=(1/sqrt(length(vol_measured)))*sqrt(fitness(i));
end

best_fitting_index=1;
for i=2:M
    if fitness(i)<fitness(best_fitting_index)
        best_fitting_index=i;
    end
end

best_RMSE= fitness(best_fitting_index);
epsilon=((bit2int(population(best_fitting_index,1:7)',7))*10^-3)+0.001;

```

```

lambda=((bit2int(population(best_fitting_index,8:14)',7))+17)*10^-3;
S=((bit2int(population(best_fitting_index,15:21)',7))*10^-2)+0.01;
[Vl,Vd,vol]=tumor_volume(lambda,epsilon,S,time,vol_measured(1),treatment,t_rt_init);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Saving results & plots
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
mkdir(sprintf('patient %i RMSE %.3f', patient,best_RMSE));
folder_path=strcat('C:\Users\Francesco\Desktop\my code\', ...
    sprintf('patient %i RMSE %.3f', patient,best_RMSE),' ');
cd (folder_path);

plot(time,vol, 'red', 'LineWidth', 2)
xlabel('Time')
ylabel('Tumor volume')
title(sprintf('patient %i',patient))
hold ;
scatter(t_measured, vol_measured, 'o','black', 'filled')
scatter(t_rt_init,0, '^','blue','filled')
legend(sprintf('Fit, RMSE = %.3f', best_RMSE), 'Data', 'Therapy start')
savefig('fitting.fig')

fitted_parameters=[lambda epsilon S];
writematrix(fitted_parameters,'fitted_parameters.xlsx');
writematrix(ga_used_parameters,'GA_parameters.xlsx');
cd 'C:\Users\Francesco\Desktop\my code\'

toc

end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% IRT
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

if IRT== true
    treatment='IRT';
    fractions=[11];
    t_bt看_fractions=[28 42 56 70];
    max_fractions=max(fractions);
    max_t_bt看_fractions=max(t_bt看_fractions);

    patient=1;
    T_fin=1500;
    time=1:1:T_fin;

```

```

t_rt_init=xlsread('dataset.xlsx',patient,'B1:B1');
t_rt(1)= t_rt_init;
V0=xlsread('dataset.xlsx', patient,'G1:G1');
folder_path=strcat('C:\Users\Francesco\Desktop\my code\fitting\', ...
    sprintf('patient %i', patient),'\ ');
cd (folder_path);

lambda=xlsread('fitted_parameters.xlsx',1,'A1');
epsilon=xlsread('fitted_parameters.xlsx',1,'B1');
S=xlsread('fitted_parameters.xlsx',1,'C1');
cd 'C:\Users\Francesco\Desktop\my code\';

for i=1:max_fractions
    if ismember(i, fractions)
        for j=1:max_t_btw_fractions
            if ismember(j,t_btw_fractions)
                for k=2:i
                    t_rt(k)=j+t_rt(k-1);
                end
                for l=1:T_fin
                    [Vl,Vd,vol]=tumor_volume(lambda,epsilon,S,time,...
                        V0,treatment,t_rt);
                end
                plot(time,vol, 'blue', 'LineWidth', 2)
                xlabel('Time')
                ylabel('Tumor volume')
                title(sprintf('patient %d iRT fractions=%d Tbtwf=%d',...
                    patient,i,j))
                savefig(sprintf('p%d f%d tbf%d',patient,i,j));
            end
        end
    end
end
toc
end

function [Vl,Vd,vol] = tumor_volume(lambda,epsilon,S,t,V0, treatment,t_rt)
% calculate the tumor volume under RT effect
Vl=zeros(length(t),1);
Vd=zeros(length(t),1);
Vl(1)=V0;
V0_d=0;
t0=0;

```

```

if strcmp('HFSRT',treatment)
    t_rt_init=t_rt(1);
    for i=2:t_rt_init
        V1(i)=V0*exp(lambda*(t(i)-t0));
    end

    for i=t_rt_init:1:length(t)
        if i==t_rt_init
            V1_old= V1(i);
            Vd_old= Vd(i);
            for j=1:5
                V1_new=S*V1_old;
                Vd_new=Vd_old+(1-S)*V1_old;
                V1_old=V1_new;
                Vd_old=Vd_new;
            end
            V1(i)= V1_new;
            Vd(i)=Vd_new;
            V0=V1(i);
            V0_d=Vd(i);
            t0=t(i);
        else
            V1(i)=V0*exp(lambda*(t(i)-t0)+(lambda/epsilon)*(exp(-epsilon*(t(i)-t0))-1));
            if i~=1
                Vd(i)=V0_d*exp(-lambda*(t(i)-t0));
            end
        end
    end
end

if strcmp('IRT', treatment)
    t_rt_init=t_rt(1);

    for i=2:t_rt_init
        V1(i)=V0*exp(lambda*(t(i)-t0));
    end

    for i=t_rt_init:1:length(t)
        if ismember(i,t_rt)
            if i~=t_rt_init
                V1(i)=V0*exp(lambda*(t(i)-t0)+(lambda/epsilon)*(exp(-epsilon*(t(i)-t0))-1));
                if i~=1
                    Vd(i)=V0_d*exp(-lambda*(t(i)-t0));
                end
            end
        end
    end
end

```

```

        Vd(i)=Vd(i)+(1-S)*Vl(i);
        Vl(i)=S*Vl(i);

        V0=Vl(i);
        V0_d=Vd(i);
        t0=t(i);
    else
        Vl(i)=V0*exp(lambda*(t(i)-t0)+(lambda/epsilon)*(exp(-epsilon*(t(i)-t0))-1));
        if i~=1
            Vd(i)=V0_d*exp(-lambda*(t(i)-t0));
        end
    end
end
end
vol=Vl+Vd;

```

Appendix B

GA template for the optimal IRT protocol search

```
tic
%prepare run
clear
clc
close all

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% GA parameters
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
n_generations=100;
M=100; % # of population members
M_mating= M/2;

% crossover parameters
p_c=0.8;
n_crossover= 1000;

% mutation parameters
p_m=0.01;
n_mutation=1000;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% fitting through HFSRT
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
treatment= 'IRT';
patient=15;
lambda=0.082;
epsilon=0.002;
S_5=0.56;
```

```

max_nfrac=12;

PST_HFSRT=220;
PST_IRT=418;
vol_init=xlsread('dataset.xlsx', patient, 'G1:G1');
vol_PST=13.5;
t_rt_init=xlsread('dataset.xlsx', patient, 'B1:B1');

%time range
T_fin=1000;
time=1:1:T_fin;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%generating randomly initial population
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%start GA
for q=6:max_nfrac %use this in case of hyperfractionation
    nbit = 3*q;
    S=1-((5/q)*(1-S_5));
    population = zeros(M,nbit);
    for i = 1:M
        for j=1:nbit
            population(i,j)=randi([0 1], 1);
        end
    end

    n_generations=100;
    while n_generations>0

        %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
        %Selection
        %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
        mb_indices=ones(M_mating, 1);
        mb=zeros(M_mating, nbit);

        fitness=zeros(M,1);
        t_rt=zeros(M,max_nfrac);
        nfrac=max_nfrac*ones(M,1);
        for i=1:M
            t_rt(i,1)=t_rt_init;
            index=4;
            index_rt=2;
            while index<nbit
                if bit2int(population(i,index:index+2)',3)==0
                    nfrac(i)=nfrac(i)-1;
                end
            end
        end
    end
end

```

```

else
    t_rt(i,index_rt)=t_rt(i,index_rt-1)+...
        ((bit2int(population(i,index:index+2)',3))+...
        3)*7;
    index_rt=index_rt+1;
end
index=index+3;

end

[Vl,Vd,vol]=tumor_volume(lambda,epsilon,S,time,...
    vol_init,treatment,t_rt(i,1:nfrac(i)));
for j=t_rt_init+20:T_fin
    if vol(j)>vol_PST
        fitness(i)= j-1;
        break
    end
end
end

fitness_tot=sum(fitness);
p=zeros(M,1);
for i=1:M
    p(i)= fitness(i)/fitness_tot;
end

t=zeros(M,1);

for i=1:M
    for j=1:i
        t(i)=t(i)+p(j);
    end
end

already_used=zeros(M,1);
for i=1:M_mating
    r=rand();
    for j=1:M
        if r<t(j) && already_used(j)==0
            mb_indices(i)=j;
            already_used(j)=1;
            for k=1:nbit
                mb(i,k)=population(j,k);
            end
            break
        end
    end
end

```

```

        end
    end
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Recombination
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Crossover
for i=1:n_crossover
    if rand()<p_c
        c1= randi([1 nbit], 1);
        c2= randi([c1 nbit], 1);
        crossover1=randsample(mb_indices,1);
        crossover2=randsample(mb_indices,1);
        for j=c1:c2
            swap=population(crossover1,j);
            population(crossover1,j)=population(crossover2,j);
            population(crossover2,j)=swap;
        end
    end
end

%Mutation
for i=1:n_mutation
    if rand()<p_m
        mutation=randsample(mb_indices,1);
        bit_mutation=randsample(1:nbit,1);
        population(mutation,bit_mutation)=~population(mutation,...
            bit_mutation);
    end
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% calculating mating buffer members' fitness
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
fitness_mb=zeros(M_mating,1);
t_rt=zeros(M_mating,max_nfrac);
nfrac=max_nfrac*ones(M_mating,1);
for i=1:M_mating
    t_rt(i,1)=t_rt_init;
    index=4;
    index_rt=2;
    while index<nbit
        if bit2int(population(mb_indices(i),index:index+2)',3)==0
            nfrac(i)=nfrac(i)-1;
        end
    end
end

```

```

else
    t_rt(i,index_rt)=t_rt(i,index_rt-1)+...
    ((bit2int(population(mb_indices(i),index:index+2)',3))+3)*7;
    index_rt=index_rt+1;
end
index=index+3;

end

[Vl,Vd,vol]=tumor_volume(lambda,epsilon,S,time,...
    vol_init,treatment,t_rt(i,1:nfrac(i)));
for j=t_rt_init:T_fin
    if vol(j)+20>vol_PST
        fitness_mb(i)= j-1;
        break
    end
end
end
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% new population
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
for i=1:M_mating
    candidate=randsample(1:M,1);
    if fitness_mb(i)>fitness(candidate)
        for j=1:nbit
            population(candidate,j)=mb(i,j);
        end
    end
end
end
%disp(n_generations)
n_generations=n_generations-1;

end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% best fitting and statistical analisys
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%final population fitness
fitness=zeros(M,1);
t_rt=zeros(M,max_nfrac);
nfrac=max_nfrac*ones(M,1);
for i=1:M
    t_rt(i,1)=t_rt_init;

```

```

index=4;
index_rt=2;
while index<nbit
    if bit2int(population(i,index:index+2)',3)==0
        nfrac(i)=nfrac(i)-1;
    else
        t_rt(i,index_rt)=t_rt(i,index_rt-1)+...
            ((bit2int(population(i,index:index+2)',3))+3)*7;
        index_rt=index_rt+1;
    end
    index=index+3;
end

[Vl,Vd,vol]=tumor_volume(lambda,epsilon,S,time,...
    vol_init,treatment,t_rt(i,1:nfrac(i)));
for j=t_rt_init+20:T_fin
    if vol(j)>vol_PST
        fitness(i)= j-1;
        break
    end
end
end

best_fitting_index=1;
for i=2:M
    if fitness(i)>fitness(best_fitting_index)
        best_fitting_index=i;
    end
end

best_fitness= fitness(best_fitting_index);
[Vl,Vd,vol]=tumor_volume(lambda,epsilon,S,time,vol_init(1),treatment,...
    t_rt(best_fitting_index,1:nfrac(best_fitting_index)));

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Saving results & plots
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
plot(time(1:T_fin),vol(1:T_fin),'LineWidth',1)
xlabel('Time [days]')
ylabel('Tumor volume [cm^3]')
title(sprintf('Patient %i',patient))
%legend(sprintf('%i fractions',q))%, 'Therapy start',...
sprintf('HFSRT PST = %i',PST_HFSRT),sprintf('IRT PST (no hyperfractionation)...

```

```

    = %i',PST_IRT));
    hold on;
    %writematrix(t_rt(best_fitting_index),...
    1:nfrac(best_fitting_index),...
    sprintf('Hyperfractionation_%i.xlsx',q));
    disp(S)
end
scatter(t_rt_init,0, '^','green','filled');
scatter(PST_HFSRT, 0, '*','red');
scatter(PST_IRT , 0, '*','blue');
line([0 T_fin], [vol_PST vol_PST])
legend show
savefig('Hyperfractionation.fig')
toc

```

Bibliography

- [1] JCL Alfonso and L Berk. Modeling the effect of intratumoral heterogeneity of radiosensitivity on tumor response over the course of fractionated radiation therapy. *Radiation Oncology*, 14(1):1–12, 2019.
- [2] Antonella Belfatto, Barbara Alicja Jereczek-Fossa, Guido Baroni, and Pietro Cerveri. Model-supported radiotherapy personalization: in silico test of hyper-and hypo-fractionation effects. *Frontiers in Physiology*, 9:1445, 2018.
- [3] Sarah C Brüningk, Jeffrey Peacock, Christopher J Whelan, Renee Brady-Nicholls, Hsiang-Hsuan M Yu, Solmaz Sahebjam, and Heiko Enderling. Intermittent radiotherapy as alternative treatment for recurrent high grade glioma: A modeling study based on longitudinal tumor measurements. *Scientific reports*, 11(1):1–13, 2021.
- [4] Michael R Chicoine and Daniel L Silbergeld. Mitogens as motogens. *Journal of neuro-oncology*, 35(3):249–257, 1997.
- [5] Robert Gatenby and Joel Brown. The evolution and ecology of resistance in cancer therapy. *Cold Spring Harbor perspectives in medicine*, 8(3):a033415, 2018.
- [6] Robert A Gatenby, Ariosto S Silva, Robert J Gillies, and B Roy Frieden. Adaptive therapy. *Cancer research*, 69(11):4894–4903, 2009.
- [7] Robert J Gillies, Daniel Verduzco, and Robert A Gatenby. Evolutionary dynamics of carcinogenesis and why targeted therapy does not work. *Nature Reviews Cancer*, 12(7):487–493, 2012.
- [8] Daniel J Glazar, G Daniel Grass, John A Arrington, Peter A Forsyth, Natarajan Raghunand, Hsiang-Hsuan Michael Yu, Solmaz Sahebjam, and Heiko Enderling. Tumor volume dynamics as an early biomarker for patient-specific evolution of resistance and progression in recurrent high-grade glioma. *Journal of clinical medicine*, 9(7):2019, 2020.
- [9] Ahmad Hassanat, Khalid Almohammadi, Esra’a Alkafaween, Eman Abunawas, Awni Hammouri, and VB Surya Prasath. Choosing mutation

and crossover ratios for genetic algorithms—a review with a new dynamic approach. *Information*, 10(12):390, 2019.

- [10] John McCall. Genetic algorithms for modelling and optimisation. *Journal of computational and Applied Mathematics*, 184(1):205–222, 2005.
- [11] Lauren MF Merlo, John W Pepper, Brian J Reid, and Carlo C Maley. Cancer as an evolutionary and ecological process. *Nature reviews cancer*, 6(12):924–935, 2006.
- [12] James D Murray. *Mathematical biology II: spatial models and biomedical applications*, volume 3. Springer New York, 2001.
- [13] James Dickson Murray. *Mathematical biology: I. An introduction*. Springer, 2002.
- [14] Andrei Petrovski. *An application of genetic algorithms to chemotherapy treatment*. PhD thesis, 1998.
- [15] Steven H Strogatz. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. CRC press, 2018.
- [16] Shabana Tabassum, Norhayati Binti Rosli, and Mazma Sayahidatul Ayuni Binti Mazalan. Mathematical modeling of cancer growth process: a review. In *Journal of Physics: Conference Series*, volume 1366, page 012018. IOP Publishing, 2019.
- [17] Darrell Whitley. A genetic algorithm tutorial. *Statistics and computing*, 4(2):65–85, 1994.
- [18] Lalita Yadav, Naveen Puri, Varun Rastogi, Pranali Satpute, and Vandana Sharma. Tumour angiogenesis and angiogenic inhibitors: a review. *Journal of clinical and diagnostic research: JCDR*, 9(6):XE01, 2015.

Acknowledgements

Un ringraziamento speciale al Prof. Delitala, mio relatore e mentore, che mi ha fatto scoprire ed amare il meraviglioso mondo della matematica applicata e che mi ha insegnato come essa possa essere impiegata per quello che trovo sia lo scopo più nobile della scienza: migliorare la vita di qualcuno.

Un caro ringraziamento al Prof. Mario Ferraro, che con la sua infinita gentilezza mi ha guidato in questo percorso di tesi con i suoi preziosi consigli.

Per ultimi, ma non per importanza, ringrazio di cuore la mia famiglia e i miei amici, che mi hanno dato tanto sostegno e affetto nei momenti più bui e che hanno festeggiato con me tutti i piccoli traguardi che sono riuscito a tagliare nel corso di questo meraviglioso viaggio.