UNIVERSITY OF ZAGREB

FACULTY OF SCIENCE

DEPARTMENT OF MATHEMATICS

Vanna Žnidar

MODELING AVASCULAR TUMOUR GROWTH

Diploma thesis

Zagreb, July, 2024

Ovaj diplomski rad obranjen je dana _____ pred ispitnim povjerenstvom
u sastavu:

1. _____ , predsjednik
2. _____ , član
3. _____ , član

Povjerenstvo je rad ocijenilo ocjenom _____ .

Potpisi članova povjerenstva:

1. _____
2. _____
3. _____

Table of Contents

Introduction

The study of avascular tumour growth has been a significant area of research in the field of mathematical biology and medicine. This is primarily due to the potential insights it can provide into the early stages of cancer development, which can be crucial for early detection and treatment. In this thesis, I focus on several classical phenomenological models of growth, describing, fitting, analyzing, and comparing them in the context of avascular tumour growth.

The data used in this study are time series of average sizes of avascular tumours grown *in vitro*, measured at discrete time points up to the lifetime of the tumour spheroid in culture. These measurements are part of the research under the Croatian Science Foundation project "Methods Developing in Mathematical Modelling in Biology and Medicine".

Through this work, I aim to contribute to the understanding of avascular tumour growth dynamics, potentially aiding in the development of more effective treatment strategies. I also hope to demonstrate the power and versatility of mathematical modeling as a tool in biological and medical research.

Chapter 1

Tumours

1.1 About Tumours

Tumours, also known as neoplasms, are the result of uncontrolled growth of cells that can appear in any part of the body. The study of tumours is known as oncology. There are two main types of tumours: benign and malignant. Benign tumours are non-cancerous and do not spread to other parts of the body, while malignant tumours, or cancers, have the ability to invade nearby tissues and spread through the bloodstream or lymphatic system which can cause serious health problems. In this thesis, I will study the growth of avascular tumours, i.e., tumours that do not have access to blood.

Tumour spheroids are three-dimensional cell culture models that closely mimic the *in vivo* tumour. They are used extensively in cancer research because they replicate many of the morphological and physiological characteristics of solid tumours, including cell-cell and cell-matrix interactions, nutrient and oxygen gradients, and drug penetration [1]. In this paper, I will also explore the growth dynamics of tumour spheroids, providing insights into the behaviour of avascular tumours.

Tumour discovery

The discovery and understanding of tumours have evolved significantly over the centuries, with advancements in medical technology and research playing a crucial role. The earliest known references to tumours date back to ancient civilizations. The Edwin Smith Papyrus, an Egyptian medical text from around 1600 BC, contains descriptions of tumours, indicating that ancient physicians had some understanding of these abnormal growths [2].

However, the modern understanding of tumours began to take shape in the 19th century with the advent of the microscope. This revolutionary tool allowed scientists to observe cells and tissues at an unprecedented level of detail. One of the key figures in this era

was Rudolf Virchow, a German physician and pathologist. Often referred to as the "father of modern pathology," Virchow was the first to propose that all cells, including cancerous ones, originate from other cells, a theory known as cellular pathology [3]. His work laid the foundation for the modern understanding of tumour formation and growth.

The late 19th and early 20th centuries saw further advancements in tumour detection and diagnosis. The discovery of X-rays by Wilhelm Conrad Roentgen in 1895 was particularly transformative. For the first time, doctors could non-invasively visualize the internal structures of the body, making it possible to detect tumours that were not palpable or visible from the outside [4].

In the mid-20th century, the development of molecular biology and genetics provided new insights into the mechanisms behind tumour growth and development. Scientists began to understand that tumours were the result of genetic mutations that lead to uncontrolled cell growth. This understanding has been refined and expanded over the decades, leading to the identification of numerous genes and pathways involved in tumour formation.

Today, tumours are classified into benign *(in situ)* and malignant (invasive) types. It is also classified into liquid, such as leukaemia, and solid tumours, such as carcinomas or sarcomas. Nowadays, the detection and diagnosis of tumours involve a combination of techniques, including physical examination, imaging tests like X-rays and MRIs, and molecular diagnostic tests. Despite these advancements, tumours remain a significant global health concern, highlighting the ongoing need for research and innovation in this field. The journey from the ancient descriptions in the Edwin Smith Papyrus to the sophisticated understanding we have today is a testament to the progress of medical science and the relentless pursuit of knowledge.

Tumour classification

One of the most fascinating findings in the field of tumours is that all tumour cells originate from a single cell that has undergone genetic mutations. These mutations can be caused by various factors, including exposure to carcinogens, genetic predisposition, and lifestyle choices. The genetic changes in tumour cells can lead to uncontrolled growth, invasion of surrounding tissues, and metastasis to distant organs. The differences between tumour types are based on molecular mechanisms, risk factors and spread patterns.
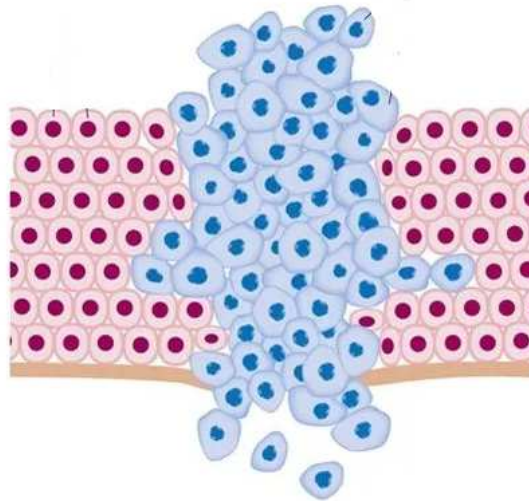
Figure 1.1: Cancer cells and normal cells

Ilustration from [5], here Figure 1.1, represents cancer cells colored in blue, while red-colored cells represent normal cells.

Tumours can be classified based on several factors, including histology, differentiation (grade), or expansion (TNM) [6].

There are five primary classifications of cancer, distinguished by their histological features: carcinoma, sarcoma, myeloma, leukaemia, and lymphoma. Additionally, there are certain hybrid types that exhibit characteristics of more than one category. Carcinomas are the most prevalent type of cancer [7].

Considering the differentiation of tumour cells, there are four grades: grade I (well-differentiated), grade II (moderately differentiated), grade III (poorly differentiated), and grade IV (undifferentiated) [6].

The TNM classification system is based on three key factors: the size and depth of the primary tumour (T), the presence of regional lymph node involvement (N), and the presence of distant metastasis (M). This system is used to stage tumours and guide treatment decisions [8].

Tumour properties

Tumours have several key properties that distinguish them from normal tissues. These properties are known as the hallmarks of cancer [9] and include:

- **Activating invasion and metastasis:** Tumour cells have the ability to invade surrounding tissues and spread to distant organs through the bloodstream or lymphatic system.

- **Inducing angiogenesis:** Tumours stimulate the growth of new blood vessels to supply nutrients and oxygen, allowing them to grow and survive.

- **Enabling replicative immortality:** Tumour cells have the ability to divide indefinitely, unlike normal cells that have a limited lifespan.

- **Resisting cell death (apoptosis):** Tumour cells can evade programmed cell death, allowing them to survive and proliferate.

- **Sustaining proliferative signaling:** Tumour cells can activate signaling pathways that promote cell growth and division. They have the ability to divide uncontrollably, leading to the formation of masses or lesions.

- **Evading growth suppressors:** Tumour cells can bypass mechanisms that normally inhibit cell growth, allowing them to proliferate unchecked.

- **Deregulating cellular energetics:** Tumour cells go through metabolic reprogramming to maintain their high proliferation rate and energy demands.

- **Avoiding immune destruction:** Tumour cells can evade the immune system's surveillance and destruction, allowing them to grow unchecked.

Tumour types

Tumour, a complicated and heterogeneous disease, includes a wide range of malignancies that can affect almost any organ of the body. These tumours are caused by abnormal cell growth and division, which frequently results in the production of masses or lesions that disrupt normal body activities. There are more than 100 different types of tumours that can occur in various parts of the body. Each type of tumour has its own characteristics, causes, and treatment methods.

Skin cancer is one of the most frequent types of cancer, consisting of basal cell carcinoma, squamous cell carcinoma, and melanoma. UV radiation exposure from sunshine and tanning beds are among the leading causes of skin cancer. While basal and squamous cell carcinomas are generally less malignant, melanoma can spread quickly and pose a serious hazard if not diagnosed early [10].

Lung cancer ranks among the deadliest cancers worldwide, owing mostly to tobacco use, while exposure to environmental pollutants such as asbestos and radon also plays a role. Lung cancer can be asymptomatic until advanced stages, resulting in a poor prognosis. Screening programs aimed at high-risk individuals, such as smokers, seek to enhance early detection rates and outcomes [11].

Breast cancer is another common malignancy that primarily affects women, but men can also develop it. Breast cancer can be classified into several subtypes, the most frequent of which are hormone receptor-positive and HER2-positive. Early identification using mammography and advancements in treatment modalities, such as targeted medicines and immunotherapy, have dramatically improved survival rates [12].

Prostate cancer mostly affects men and is one of the top causes of cancer-related death among men. Although many instances proceed slowly and do not require rapid treatment, aggressive types can spread to distant organs, posing a serious threat to health. Screening with prostate-specific antigen (PSA) tests and digital rectal exams can assist identify people who are at risk [13].

Colorectal cancer includes tumours of the colon and rectum and is one of the most common cancers worldwide. Risk factors include a high-processed-meat, low-fiber diet, sedentary lifestyle, and genetic predisposition. Screening techniques like colonoscopy can detect precancerous polyps or early-stage cancers, allowing for appropriate intervention [14].

Leukaemia, a malignancy of the blood and bone marrow, is caused by an abnormal production of white blood cells. This broad category of disorders includes acute lymphoblastic leukaemia (ALL), acute myeloid leukaemia (AML), chronic lymphocytic leukaemia (CLL), and chronic myeloid leukaemia (CML). While some types of leukaemia are more common in youngsters, others primarily afflict adults, emphasizing the age-related nature of cancer occurrence [15].

Osteosarcoma is a type of bone cancer that typically develops in osteoblast cells that form bone. It most commonly occurs in children and young adults, but also in older adults. Symptoms often include bone pain, swelling, and fractures [16]. The exact cause of osteosarcoma is unknown. Despite being a relatively rare, osteosarcoma is the most common type of bone cancer in children and teens.

This is just a small fraction of the possible types of tumours. Each of these tumour types can be further subdivided into subtypes with different characteristics and prognoses. Understanding the many forms of cancer, their prevalence, and related risks is critical for improving prevention and treatment techniques.
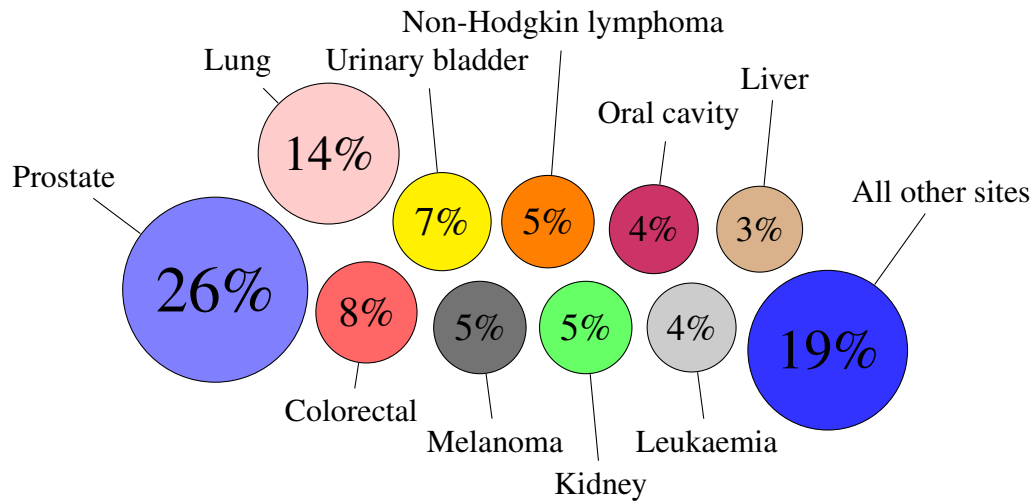
Figure 1.2: Cancer incidence for male patients in 2015 in the United States [17].
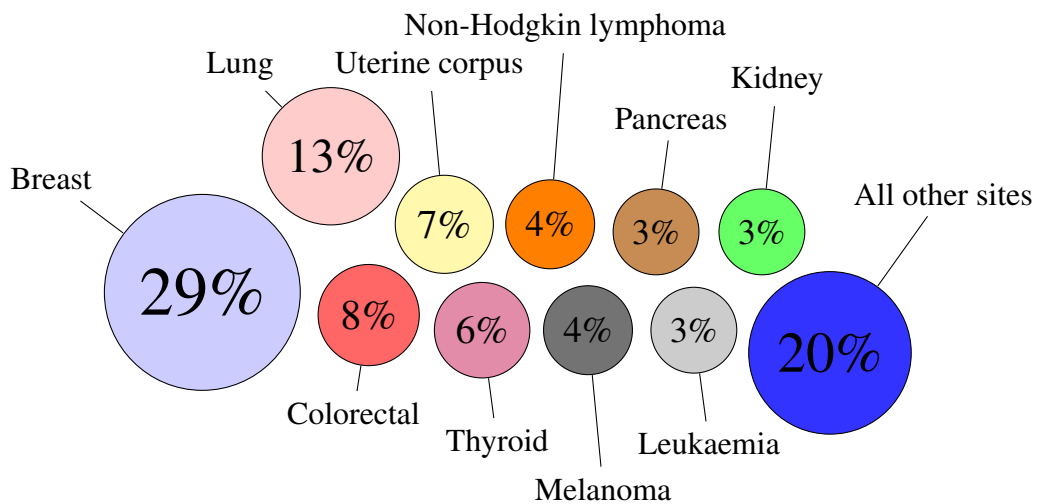
Figure 1.3: Cancer incidence for female patients in 2015 in the United States [17].
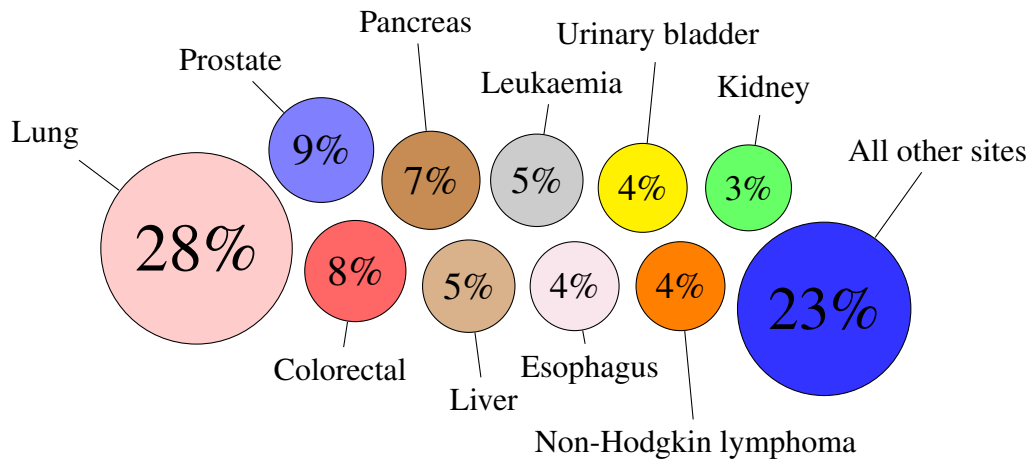
Figure 1.4: Cancer deaths for male patients in 2015 in the United States [17].
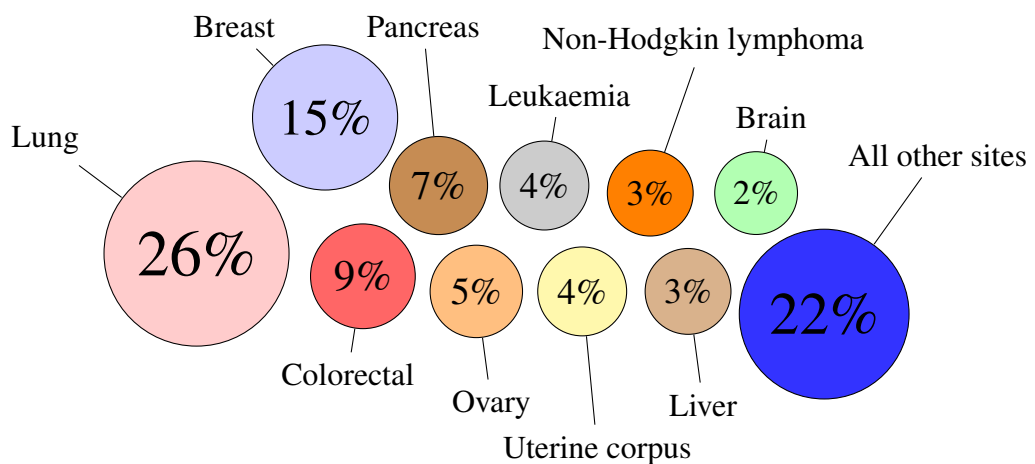
Figure 1.5: Cancer deaths for female patients in 2015 in the United States [17].

In Figure 1.2 and Figure 1.3, we can see the incidence of cancer for male and female patients in the United States in 2015, respectively. In Figure 1.4, we can see deaths caused by cancer in the United States in 2015 for male, while Figure 1.5 represents the data for female. Same tumour types are colored exactly the same in all graphs.

Figure 1.6 shows the distribution of cancer cases by age groups in the United States in 2022. As the population ages, the incidence of cancer is expected to rise. That is why tumours are also known as "the disease of the elderly" because the risk of developing cancer increases with age but can affect individuals of all ages. However, certain age groups may be more susceptible to specific types of malignancies. For instance, childhood tumours such as leukaemia and brain tumours are more prevalent in younger age groups, while prostate and breast cancers are commonly diagnosed in older individuals. However, advancements in research, early detection, and personalized treatment approaches offer hope in mitigating the burden of tumour across diverse populations.
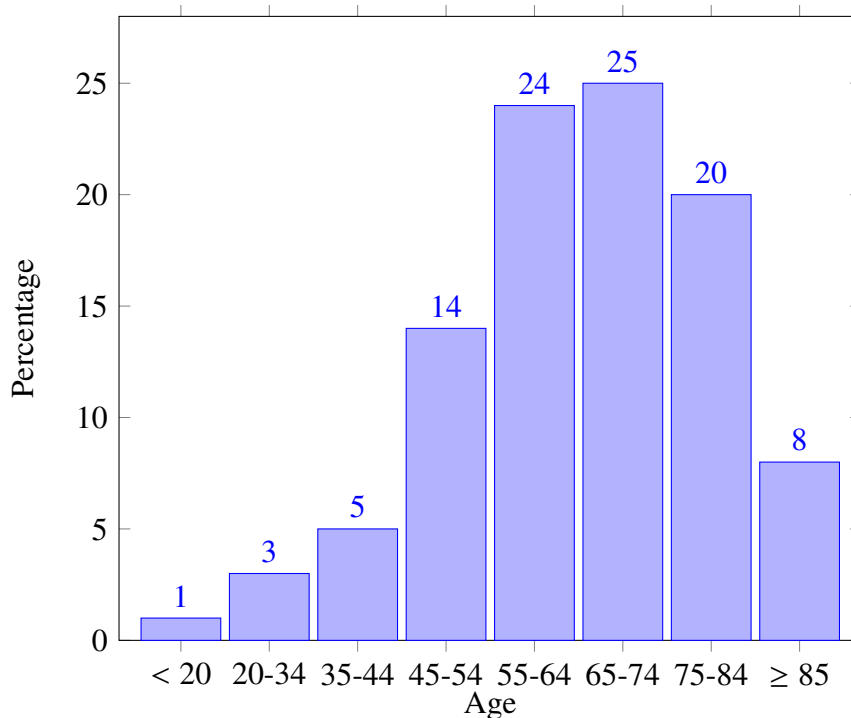
Figure 1.6: Cancer cases by age groups in the United States in 2022 [18].

1.2   Osteosarcoma

Osteosarcoma (OS) is a primary malignant bone tumour with worldwide incidence of 3.4 cases per million people per year. The mail feature of OS is osteoid production in association with malignant mesenchymal cells [19]. OS incidence is bimodal, with peaks of incidence at age 18 and ages over 60 [20]. It is a rare cancer, accounting for less than 1% of all cancers diagnosed each year. It develops particularly in the long bones of the arms and

legs [16]. The tumour is solid, hard, irregular, and often painful. If the tumour is located near a joint, it may cause limited range of motion and difficulty moving the affected limb. In advanced cases, the tumour may spread to other parts of the body, such as the lungs, causing additional symptoms.

The exact cause of OS is unknown, but it is thought to be related to rapid bone growth, as it often occurs during adolescence when children are growing rapidly [1]. Other risk factors for osteosarcoma include genetic predisposition, exposure to radiation, and certain genetic conditions such as Li-Fraumeni syndrome or Rothmund–Thomson syndrome and hereditary retinoblastoma.

The diagnosis of OS typically involves a combination of imaging tests, such as X-rays, CT scans, and MRIs, and a biopsy to confirm the presence of cancer cells [21]. Once the diagnosis is confirmed, additional tests may be done to determine the extent of the disease and whether it has spread to other parts of the body.

The treatment of OS typically involves a combination of surgery, chemotherapy, and sometimes radiation therapy [21]. The primary goal of treatment is to remove the tumour and prevent it from spreading to other parts of the body. Surgery is usually the first line of treatment and involves removing the tumour and a margin of healthy tissue around it to ensure that all cancer cells are removed. Chemotherapy is often used before and after surgery to shrink the tumour and kill any remaining cancer cells. Radiation therapy may be used in some cases to kill cancer cells that remain after surgery or to relieve pain and other symptoms.

The prognosis for OS depends on several factors, including the stage of the disease, the location and size of the tumour, and the patient's age and overall health. In general, the prognosis is better for patients with localized disease that has not spread to other parts of the body. With advances in treatment, the survival rates for osteosarcoma have improved significantly in recent years, with many patients achieving long-term remission or cure. Stage I is rare and has a 90% survival rate, while stage III has a 30% survival rate.

Based on the predominant matrix production, tumours can be further divided into fibroblastic, chondroblastic, osteoblastic and telangiectatic OS [20], and based on the aggressiveness of the tumour into low, intermediate and high grade OS [22].

Chondroblastic osteosarcoma

Chondroblastic osteosarcoma is a subtype of osteosarcoma that is characterized by the presence of a significant amount of cartilage in the tumour. It is one of the most common subtypes of OS. The treatment for chondroblastic OS typically involves a combination of surgery to remove the tumour and chemotherapy to kill any remaining cancer cells. Despite treatment, the prognosis for this subtype of osteosarcoma is often poorer than for other subtypes, due to its aggressive nature and tendency to spread to other parts of the body.

Osteosarcoma stem cells

Osteosarcoma stem cells are a subtype of cancer stem cells found in osteosarcoma. These cells have the ability to self-renew and generate the diverse cells that make up the tumour. They play a key role in tumour growth, spread, and resistance to therapy.

Osteosarcoma stem cells exhibit several key characteristics that distinguish them from other tumour cells. They are capable of long-term self-renewal, can differentiate into various types of tumour cells, and show resistance to many conventional cancer therapies. They also express specific markers, such as CD133 and Stro-1, which can be used to identify and isolate these cells [23].

Understanding osteosarcoma stem cells can provide new insights into cancer therapy. Since these cells are resistant to many conventional therapies, targeting osteosarcoma stem cells could be an effective way to treat OS. Research is ongoing to develop therapies that specifically target these cells, with the aim of improving outcomes for patients with OS.

1.3   Tumour Spheroids

Tumour spheroids are three-dimensional cell culture models that are very useful tools for studying tumour biology. They can be generated from a variety of tumour cell lines and are used in cancer research to study tumour growth, invasion, metastasis, and drug resistance.

Formation of tumour spheroids

The multicellular tumour spheroids, hereafter MTS, culture system offers a particularly convenient experimental paradigm for studying tumours. They are formed when mono-layer tumour cells are grown by different *in vitro* methods [24]. Oxygen and nutrition come through the surface of tumour spheroids. A three-layered model of MTS is illustrated in Figure 1.7: necrotic core layer a, quiescent or nonproliferating cells layer b, and proliferating cells layer c. These three layers are distinct. In the necrotic region, which is the inner part of the tumour, the deficiency of oxygen and nutrients causes starvation-induced cell death. The quiescent layer refers to an area where cells are at rest and are not in an active process of division or proliferation, while the proliferating layer refers to the area of the tumour where cells are actively proliferating and dividing.
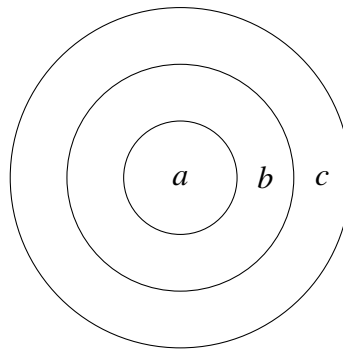
Figure 1.7: Multicellular tumour spheroid

Advantages of using tumour spheroids in research

Tumour spheroids offer several advantages over traditional two-dimensional cell culture models. They more accurately replicate the physical and biochemical characteristics of tumours in the body [1]. This makes them a valuable tool for studying tumour growth, invasion, and metastasis. Tumour spheroids are also more resistant to chemotherapy and radiation therapy than monolayer cultures, making them useful for studying drug resistance mechanisms. Additionally, tumour spheroids can be easily manipulated and imaged, allowing researchers to study tumour behavior in real-time. These advantages make tumour spheroids an invaluable tool for cancer research and drug development.

The importance of research in the field of tumour spheroids

Research in the field of tumour spheroids is essential for a lot of reasons. Firstly, researching tumours helps uncover the causes of cancer, i.e. the genetic predisposition, environmental factors, and lifestyle choices. Understanding causes can reduce the risk of developing cancer. Secondly, the better we understand different types of tumours, the more competent we become in their diagnosis and treatment. Early detection of tumours increases the chances of successful treatment and improves patient prognosis. Thirdly, studying tumours can help with the development of new therapies, including targeted therapies, immunotherapies, and other innovative approaches to cancer treatment. This can result in improved treatment options and increased survival rates. Fourthly, each tumour is unique, necessitating personalized treatment approaches that take into account the specificities of each patient and their tumour. Studying tumours help in the development of personalized therapeutic approaches. Finally, if we understand the causes and mechanisms of tumour development, we can develop preventive strategies that can reduce cancer incidence. This may involve lifestyle changes, screening programs, and other interventions that mitigate exposure to risk factors. In the end, tumour research is essential to progress-

ing the war against cancer, boosting cancer patient quality of life and reducing mortality through improved prevention, diagnosis, and treatment.

1.4 Tumour Growth Phases and the Role of Spheroid Cultures

Solid tumours progress through two distinct phases of growth: **the avascular phase** and **the vascular phase**. Understanding these phases is crucial for developing effective models and treatments.

The initial phase of tumour growth is avascular, meaning the tumour lacks its own blood supply. During this phase, the tumour cells rely on diffusion to receive nutrients and oxygen. This phase can be effectively studied in the laboratory using three-dimensional multicellular spheroids. These spheroids mimic the early stages of tumour growth and have growth kinetics similar to *in vivo* tumours. Typically, avascular tumour nodules grow to a few millimeters in diameter [25].

As the tumour grows, it transitions to the vascular phase, where it can stimulate the formation of new blood vessels, a process known as angiogenesis. This transition is critical as it allows the tumour to receive a sufficient blood supply, supporting rapid growth and the potential for metastasis. Tumours secrete various chemical substances to induce surrounding tissues to sprout new blood vessels toward the tumour.

The vascular phase is marked by significant changes in the tumour's ability to invade surrounding tissues and spread to distant parts of the body. Endothelial cells (EC) from nearby capillaries are activated, degrade their basal lamina, and migrate toward the tumour, forming new capillary sprouts. These sprouts eventually fuse, forming a network that penetrates the tumour and establishes circulation [25].

Spheroid cultures are crucial for modeling the avascular phase of tumour growth and studying the initial stages of angiogenesis. They help researchers understand how tumours survive without a blood supply and what triggers the transition to the vascular phase. By examining spheroids, scientists can gather data on the effects of various treatments on tumour growth and the effectiveness of anti-angiogenic therapies.

In summary, the avascular and vascular phases of tumour growth represent key stages in the development of solid tumours. Spheroid cultures provide a valuable tool for studying these phases, offering insights into tumour biology and aiding in the development of new therapeutic strategies. Understanding these phases helps researchers tackle the complex challenges of tumour growth and metastasis.

Chapter 2

Mathematical Models

2.1   Introduction to Mathematical Models

Mathematical modeling has a distinctive philosophy and methodology, which stands in contrast to the empirical approach of deducing mechanisms from a limited (usually small) set of observations.
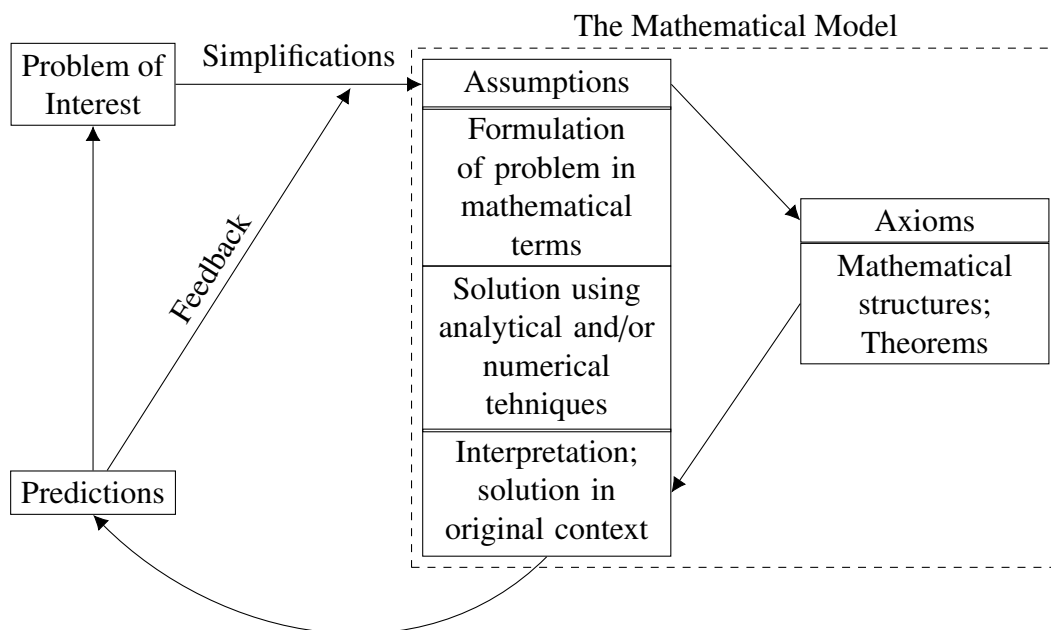
Figure 2.1: Mathematical modeling process

In Figure 2.1, we can see diagrammatic representation of the fundamental stages of the

mathematical modeling process and their interconnections.

In this paper, I will develop mathematical models to describe the growth of avascular tumours *in vitro* and fit these models to experimental data to estimate the model parameters. The goal is to gain insights into the mechanisms that drive tumour growth and to develop predictive models that can guide future research and treatment strategies.

What is a mathematical model?

Mathematical models are an essential tool in tumour research. Following [37], one possible definition of a mathematical model is;

Definition 2.1.1. *Mathematical model is an abstract, simplified mathematical construction related to part of the "real world" and created for a specific purpose.*

It is aimed to mirror reality through the use of mathematics. The purpose of a mathematical model is to capture the essential features of a system and to make predictions about its behavior. Mathematical models can take many forms, including differential equations, difference equations, and agent-based models. They can be used to describe a wide range of systems, from physical systems like the motion of planets to biological systems like the growth of tumours. In the context of tumour research, mathematical models are used to describe the growth and spread of tumours, the response of tumours to treatment, and the interactions between tumour cells and the immune system.

Why are mathematical models useful?

Mathematical models are useful in tumour research for several reasons. Firstly, they provide a quantitative framework for understanding tumour growth and behavior. By using mathematical models, researchers can simulate the behavior of tumours under different conditions, predict how tumours will grow and respond to treatment, and identify new therapeutic targets. Secondly, mathematical models can help to integrate data from different sources, such as experimental data, clinical data, and imaging data, to gain a more comprehensive understanding of tumour biology. Thirdly, mathematical models can be used to test hypotheses and generate new insights into tumour biology. By comparing model predictions with experimental data, researchers can refine their understanding of tumour growth and behavior. Finally, mathematical models can be used to guide experimental design and optimize treatment strategies. By using mathematical models to predict the effects of different treatments, researchers can identify the most effective treatment strategies and optimize patient outcomes.

Limitations of mathematical models

While mathematical models are a powerful tool in tumour research, they also have limitations. Some of limitations are;

- **Simplifications:** Mathematical models are simplifications of reality and may not capture all the complexities of tumour biology. For example, many models assume homogeneous tumour growth, whereas tumours are often heterogeneous with different cell types and microenvironments. Also, simple models are easy to manage, but they may not be accurate enough so the results can be misleading and that is the reason why more complex models are needed.

- **Assumptions:** Mathematical models rely on assumptions that may not always hold true in practice. For example, many models assume that tumour growth is exponential, whereas tumours often exhibit more complex growth patterns. Also, the assumptions can be wrong and that can lead to wrong results.

- **Data:** Mathematical models require data to parameterize and validate them. If the data used to develop the model is incomplete or inaccurate, the model predictions may be unreliable. Also, the data can be limited and that can lead to wrong results.

- **Computational complexity:** Mathematical models can be computationally intensive and require specialized software and expertise to develop and analyze. This can be a barrier to their use in tumour research. Also, the complexity of the models can be a problem because they can be hard to manage and analyze.

- **Parameters:** Realistic simulations require many parameters, which are often difficult to estimate. The uncertainty in parameter values can lead to uncertainty in model predictions. Also, the parameter space can be too large to have meaningful conclusions.

- **Choosing the model:** Selecting the appropriate mathematical model is crucial for accurate results. The choice depends on the specific research question, the available data, and the biological processes that need to be captured. However, there is often a trade-off between model complexity and interpretability. While complex models may capture the biological processes more accurately, they can be more difficult to analyze and interpret. On the other hand, simpler models may be easier to understand and analyze, but they may not capture all the relevant biological processes. Therefore, choosing the right model requires careful consideration and expert knowledge because incorrect models may well fit limited data.

- **Interpretation:** Mathematical models can be difficult to interpret, especially for non-experts. Understanding the assumptions, parameters, and predictions of a model can be challenging, and misinterpretation can lead to incorrect conclusions.

Despite these limitations, mathematical models are a valuable tool in tumour research and can provide important insights into tumour biology and treatment. Let's take a look at a couple of interesting models used in tumour research, and after that, at a general model of growth.

2.2   Classical Growth Models: Logistic, von Bertalanffy, and Gompertz

There are several mathematical models that have been developed for various purposes but also can be used to describe tumour growth too. These models are based on different assumptions and capture different aspects of tumour biology. Some of the most well-known models include the logistic model, the von Bertalanffy model, and the Gompertz model. Each of these models has its strengths and limitations, and they can be used to describe different aspects of tumour growth.

Logistic model

The logistic model was introduced in a series of three papers by Pierre François Verhulst between 1838 and 1847, who devised it as a model of population growth by adjusting the exponential growth model. It is most often characterized by one of the following forms:

$$y' = \alpha_L y \left(1 - \frac{y}{K_L} \right) = \alpha_L y - \beta_L y^2, \quad \beta_L = \frac{\alpha_L}{K_L}.$$

It is somewhat unclear why is this model named logistic (French: *logistique*), but it is believed that it was concocted as a contrast to *logarithmic*. This is because, at the time, the term *logarithmic* was used to describe the today's exponential growth. And thus, the term *logistic* was used to describe the growth that is bounded by a certain limit.

The logistic model is a simple model that describes the growth of a population or tumour over time. It is based on the assumption that the growth rate of the population or tumour is proportional to the current population size, but that the growth rate decreases as the population or tumour approaches a maximum size. The logistic model is characterized by a sigmoidal growth curve, with an initial period of exponential growth followed by a period of decelerating growth as the population or tumour reaches its carrying capacity, which we will define later on. The logistic model is widely used in tumour research to describe the growth of avascular tumours.

Von Bertalanffy model

The von Bertalanffy model is a growth model that was first proposed by Ludwig von Bertalanffy in 1938. It is a non-linear model that describes the growth of an organism over time. The model is based on the assumption that the growth rate of an organism is proportional to the difference between the organism's maximum growth rate and its current growth rate. The von Bertalanffy model is given by the equation:

$$y' = \alpha_B y^\mu - \beta_B y^\nu.$$

Several variations of the von Bertalanffy model have been developed to describe different aspects of growth, such as the growth of tumours, fish, and other organisms. The model characterized by $\mu = \frac{2}{3}$ and $\nu = 1$ is based on the *surface rule* and is often named after von Bertalanffy. We call this model *classical von Bertalanffy model*. The surface rule states that the surface area of an organism is proportional to the square of the cube root of its volume. This rule is based on the observation that the surface area of an organism determines its metabolic rate, while the volume of an organism determines its growth rate. As an organism grows, its volume increases faster than its surface area. This changes the surface area to volume ratio, which is crucial for metabolic processes.

Gompertz model

Various mathematicians have developed models to describe tumour growth over the years, some with more success than others. Many of the early models were based on exponential growth, but it was soon realized that tumour growth is more complex than that. A simple exponential model predicts unbounded growth regardless of how small the growth factor is. In reality, tumours do not grow indefinitely but reach a maximum size.

Earlier I have listed some limitations of mathematical models, and one of them is data - mathematical models require data to parameterize and validate them. If the data used to develop the model is incomplete or inaccurate, the model predictions may be unreliable. This was one of the main reasons for the lack of success of the early models. However, modern experimental tehniques have provided researchers with more accurate data on tumour growth, allowing them to develop more sophisticated models that better capture the complexities of tumour biology.

One of the most successful models is the Gompertz model, which was first proposed by Benjamin Gompertz in 1825. In its conception, the Gompertz model was not used to describe growth, but rather to describe human mortality rates. Gompertz developed this model to explain the observed increase in mortality rates with age, which he noted followed a specific exponential pattern. If we denote the distribution of human age with $M(t)$, then the Gompertz model is given by the equation:

$$M(t) = pe^{-e^{a-bt}} = pA^{B^t},$$

with constants a, b, and p, where $p > 0, b < 0$, and $A, B > 0$. In the early 20th century, biologists like L.J. Winsor and Raymond Pearl recognized the potential of the Gompertz model for describing biological growth, particularly in populations and tumour growth. Winsor specifically published work in 1932 using the Gompertz curve to model growth, highlighting its applicability beyond mortality.

Mechanism of adaptation of the Gompertz model to growth

To see that the Gompertz model is a good fit for tumour growth, we need to understand how it works. We start with a simple growth equation:

$$\frac{dN}{dt} = aN,$$

where N is the magnitude of a growing quantity, t is time, and a is the intristic growth coefficient. The quantity $\frac{dN}{dt}$ is termed *rate of growth*.

Now we assume that the growth coefficient a is not constant, but rather changes with time in the following way:

$$\frac{da}{dt} = -ka,$$

where k is a positive constant, so-called *rate of decay* of the growth coefficient. With the initial condition $a(0) = a_0$, the solution to this differential equation is:

$$a(t) = a_0 e^{-kt}.$$

Substituting this into the original growth equation, we get:

$$\frac{dN}{dt} = a_0 e^{-kt} N.$$

This is a separable differential equation, which can be solved by integrating both sides:

$$\int \frac{dN}{N} = \int a_0 e^{-kt} dt$$

$$\ln |N| = -\frac{a_0}{k} e^{-kt} + C, \quad C \in \mathbb{R}$$

$$N(t) = e^{-\frac{a_0}{k} e^{-kt} + C}.$$

With the initial condition $N(0) = N_0$, we get:

$$N(t) = N_0 e^{\frac{a_0}{k}(1 - e^{-kt})}.$$

This expression defines the Gompertz distribution function, or Gompertz growth equation.

Gompertz model for tumour growth

As I've mentioned earlier, the early models of tumour growth were based on exponential growth, which predicted unbounded growth. However, it was soon realized that tumour growth is more complex than that and that tumours reach a maximum size. A noteworthy study by Laird [40] in 1965 compared the growth rates of various tumours and extrapolated the growth curve to one cell.

In her study, Laird gathered extensive empirical data on the growth of tumours in mice. This involved measuring tumour volumes over time under controlled conditions. She compared several mathematical models of growth, including the exponential growth model, the cubic growth model, and the Gompertz model, to see which best fit the empirical data. After applying these models to her data and assessing their fit by examining how well they could predict tumour growth over time, she found that Gompertz model, characterized by its sigmoidal shape, matched the observed tumour growth patterns more closely than other models.

This conclusion was no surprise, as the Gompertz model's theoretical basis also supported its application to biological growth. The model's form, which represents a decelerating growth rate as the tumour size increases, corresponded well with the biological understanding of tumour growth constraints, such as limited resources and increasing cell death rates as tumours grow larger.

2.3   General Model of Growth

While the Gompertz model is a good fit for many tumour growth data, it is not the only model that can be used to describe tumour growth. I have already mentioned the logistic model and the classical von Bertalanffy model, which are also commonly used in tumour research. There are many other models that have been developed to describe tumour growth, each with its strengths and limitations.

However, in this paper, I will use just one additional model that generalizes Gompertz model;

$$x' = (\alpha - \beta h(x, \gamma))x,$$

where

$$h(x, \gamma) = \begin{cases} \frac{x^\gamma - 1}{\gamma} & \gamma \neq 0 \\ \ln(x) & \gamma = 0, \end{cases}$$

where $\beta > 0$, $\alpha\gamma + \beta > 0$ and $\gamma > -1$. Later we will show that it generalizes logistic and classical von Bertalanffy models too.
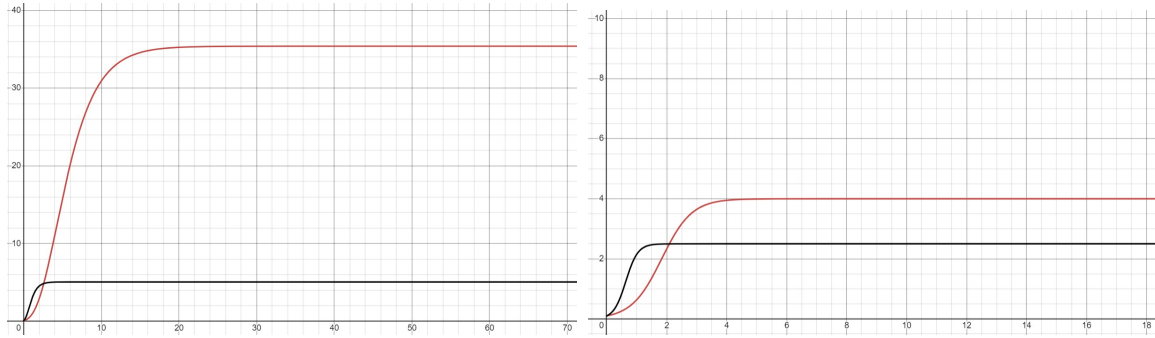
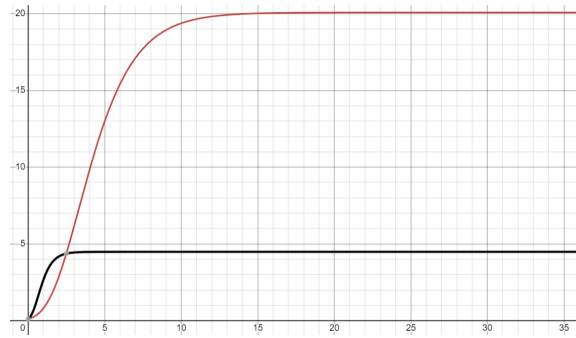Figure 2.2: The graph of the model given $x_0 = 0.1$ and $\gamma \neq 0$.

Figure 2.3: The graph of the model given $x_0 = 0.1$ and $\gamma = 0$.

In Figure 2.2 we can see the graph of the model given $x_0 = 0.1$ and $\gamma \neq 0$. Speaking of left figure, the red color represents the graph with $\alpha = 1.5, \beta = 0.5, \gamma = -0.1$, while black color represents the graph with $\alpha = 3, \beta = 2, \gamma = -0.1$. In the right figure, the difference is that $\gamma = 1$ and it is scaled differently.

In Figure 2.3 we can see the graph of the Gompertz model given $x_0 = 0.1$. The red color represents the graph with $\alpha = 1.5, \beta = 0.5$, while black color represents the graph with $\alpha = 3, \beta = 2$.

Proposition 2.3.1. *The general solution of the given differential equation with initial condition $x(0) = x_0$ when $\gamma = 0$ is:*

$$x(t) = e^{\frac{\alpha - (\alpha - \beta \ln(x_0)) e^{-\beta t}}{\beta}}.$$

And when $\gamma \neq 0$:

$$x(t) = \left(\frac{\beta - \left(\beta - (\alpha\gamma + \beta) x_0^{-\gamma} \right) e^{-(\alpha\gamma + \beta)t}}{\alpha\gamma + \beta} \right)^{-\frac{1}{\gamma}}.$$

Proof. We split the proof into two cases:

$\boxed{\gamma = 0}$: In this case, the differential equation becomes:

$$x' = (\alpha - \beta \ln(x))x$$

$$\frac{x'}{(\alpha - \beta \ln(x))x} = 1$$

This equation is separable so we can solve it by integrating both sides. Let's start with the left side:

$$\int \frac{dx}{x(\alpha - \beta \ln(x))} = \begin{bmatrix} y = \ln(x) \\ dy = \frac{1}{x}dx \end{bmatrix} =$$

$$= \int \frac{dy}{\alpha - \beta y} =$$

$$= -\frac{1}{\beta} \ln|\alpha - \beta y| + C_1 =$$

$$= -\frac{1}{\beta} \ln|\alpha - \beta \ln(x)| + C_1.$$

The right side is just $t + C_2$ so we have:

$$-\frac{1}{\beta} \ln|\alpha - \beta \ln(x)| = t + C_3, \quad C_3 \in \mathbb{R}.$$

$$\ln|\alpha - \beta \ln(x)| = -\beta t + C_4, \quad C_4 \in \mathbb{R}.$$

$$\alpha - \beta \ln(x) = C_5 e^{-\beta t}, \quad C_5 \in \mathbb{R}.$$

$$\ln(x) = \frac{\alpha - C_5 e^{-\beta t}}{\beta}, \quad C_5 \in \mathbb{R}.$$

$$x(t) = e^{\frac{\alpha - C_5 e^{-\beta t}}{\beta}}, \quad C_5 \in \mathbb{R}.$$

Using $x(0) = x_0$, we can find C_5:

$$x_0 = e^{\frac{\alpha - C_5}{\beta}} \implies C_5 = \alpha - \beta \ln(x_0).$$

So the solution is:

$$x(t) = e^{\frac{\alpha - (\alpha - \beta \ln(x_0))e^{-\beta t}}{\beta}}.$$

Notice that this is Gompertz model.

$\boxed{\gamma \neq 0}$: In this case, the differential equation becomes:

$$x' = \left(\alpha - \beta\frac{x^\gamma - 1}{\gamma}\right)x$$
$$= \left(\alpha - \frac{\beta x^\gamma}{\gamma} + \frac{\beta}{\gamma}\right)x$$
$$= \alpha x - \frac{\beta}{\gamma}x^{\gamma+1} + \frac{\beta}{\gamma}x$$
$$= -\frac{\beta}{\gamma}x^{\gamma+1} + \left(\alpha + \frac{\beta}{\gamma}\right)x.$$

We can see that this is a *Bernoulli differential equation* with $n = \gamma + 1$. This means that we need to divide the equation with $x^{\gamma+1}$:

$$x' = -\frac{\beta}{\gamma}x^{\gamma+1} + \left(\alpha + \frac{\beta}{\gamma}\right)x$$
$$\frac{x'}{x^{\gamma+1}} = -\frac{\beta}{\gamma} + \left(\alpha + \frac{\beta}{\gamma}\right)x^{-\gamma}$$

and substitute $y = x^{-\gamma}$ and, consequently, $y' = -\gamma x^{-\gamma-1}x'$:

$$-\frac{1}{\gamma}y' = -\frac{\beta}{\gamma} + \left(\alpha + \frac{\beta}{\gamma}\right)y$$
$$y' = \beta - (\alpha\gamma + \beta)y$$
$$\frac{y'}{\beta - (\alpha\gamma + \beta)y} = 1$$

This is a separable differential equation, so we can solve it by integrating both sides:

$$\int \frac{dy}{\beta - (\alpha\gamma + \beta)y} = \int dt$$
$$\frac{-\ln(\beta - (\alpha\gamma + \beta)y)}{\alpha\gamma + \beta} = t + C_1, \quad C_1 \in \mathbb{R}$$
$$\ln(\beta - (\alpha\gamma + \beta)y) = -(\alpha\gamma + \beta)t + C_2, \quad C_2 \in \mathbb{R}$$
$$\beta - (\alpha\gamma + \beta)y = C_3 e^{-(\alpha\gamma+\beta)t}, \quad C_3 \in \mathbb{R}$$
$$y = \frac{\beta - C_3 e^{-(\alpha\gamma+\beta)t}}{\alpha\gamma + \beta}, \quad C_3 \in \mathbb{R}$$

Now, using $y = x^{-\gamma}$, we get:

$$x(t) = \left(\frac{\beta - C_3 e^{-(\alpha\gamma+\beta)t}}{\alpha\gamma + \beta} \right)^{-\frac{1}{\gamma}}, \quad C_3 \in \mathbb{R}.$$

All that is left is to use the initial condition $x(0) = x_0$ to find C_3:

$$x_0 = \left(\frac{\beta - C_3}{\alpha\gamma + \beta} \right)^{-\frac{1}{\gamma}} \implies C_3 = \beta - (\alpha\gamma + \beta)x_0^{-\gamma}.$$

So the solution is:

$$x(t) = \left(\frac{\beta - \left(\beta - (\alpha\gamma + \beta)x_0^{-\gamma} \right) e^{-(\alpha\gamma+\beta)t}}{\alpha\gamma + \beta} \right)^{-\frac{1}{\gamma}}.$$

$\square$

As can be seen, the solution to the given differential equation is quite complex. The model is based on the observation that tumour growth is initially exponential but slows down as the tumour reaches a certain size. Since the tumour can't grow indefinitely, it is useful to introduce the concept of *carrying capacity*, which is often denoted with K.

Definition 2.3.2. *Carrying capacity is the size which the population of a species tends to approach, given the food, water and other necessities available in the environment.*

In Figure 2.4 we can see that for any initial value x_i, the population size $x(t)$ approaches the carrying capacity K as $t \to \infty$. This means that carrying capacity is stable *equilibrium point* of the model. Equilibrium point is a state of balance and *stable* means that the system will return to this state if it is slightly perturbed.
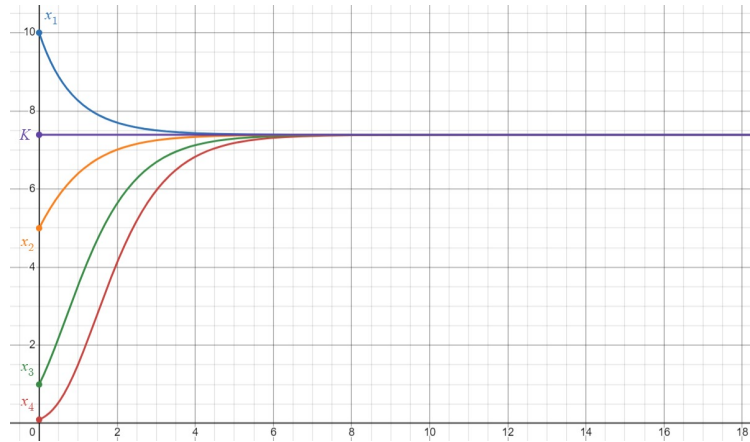
Figure 2.4: The graph representation for the model given $\alpha = 2, \beta = 1$ and $\gamma = 0$. The blue color represents the graph with initial value $x_1 = 10$, orange color with $x_2 = 5$, green color with $x_3 = 1$, while red color represents the graph with $x_4 = 0.1$. Purple colored is carrying capacity K.

Proposition 2.3.3. *The carrying capacity of the given model, assuming that* $0 < x_0 < K$ *is:*

$$K = \begin{cases} \left(1 + \frac{\alpha\gamma}{\beta}\right)^{\frac{1}{\gamma}} & \gamma \neq 0 \\ e^{\frac{\alpha}{\beta}} & \gamma = 0 \end{cases}$$

Proof. It can easily be seen that $x(t)$ is an increasing function both when $\gamma = 0$ and $\gamma \neq 0$. That means that the carrying capacity is the limit of $x(t)$ when $t \to \infty$. We split the proof into two cases:

$\boxed{\gamma = 0}$:

$$K = \lim_{t \to \infty} x(t)$$
$$= \lim_{t \to \infty} e^{\frac{\alpha - (\alpha - \beta \ln(x_0))e^{-\beta t}}{\beta}}$$
$$= e^{\frac{\alpha - (\alpha - \beta \ln(x_0)) \cdot 0}{\beta}}$$
$$= e^{\frac{\alpha}{\beta}}.$$

$\boxed{\gamma \neq 0}$:

$$K = \lim_{t \to \infty} x(t)$$

$$= \lim_{t \to \infty} \left(\frac{\beta - \left(\beta - (\alpha\gamma + \beta)x_0^{-\gamma}\right) e^{-(\alpha\gamma+\beta)t}}{\alpha\gamma + \beta} \right)^{-\frac{1}{\gamma}}$$

$$= \left(\frac{\beta - \left(\beta - (\alpha\gamma + \beta)x_0^{-\gamma}\right) \cdot 0}{\alpha\gamma + \beta} \right)^{-\frac{1}{\gamma}}$$

$$= \left(1 + \frac{\alpha\gamma}{\beta} \right)^{\frac{1}{\gamma}}.$$

$\square$

I've mentioned earlier that the early models of tumour growth were not successful because they predicted unbounded growth. By introducing the concept of carrying capacity, the model can now capture the fact that tumours can't grow indefinitely. In the previous proposition I've found the values of carrying capacity for both cases when $\gamma = 0$ and $\gamma \neq 0$, thus proving that the model does not predict unbounded growth.

Since the model is a generalization of the Gompertz model, it is expected to exhibit similar behavior. The Gompertz model is characterized by its sigmoidal shape, which represents a decelerating growth rate as the tumour size increases. This is consistent with the biological understanding of tumour growth constraints, such as limited resources and increasing cell death rates as tumours grow larger. The model is also based on the observation that tumour growth is initially exponential but slows down as the tumour reaches a certain size. This is consistent with the empirical data on tumour growth, which shows that tumours do not grow indefinitely but reach a maximum size.

Before we try to fit the model to the data, we will find the point where the exponential growth stops and the decelerating growth starts. This point is called the *inflection point*. Before we can define the inflection point, we need to define when the function is *concave* and *convex*.

Definition 2.3.4. *A real function $f(x)$ is **convex** on an interval $I \subset \mathbb{R}$ if:*

$$(\forall x_1, x_2 \in I)\,(\forall x_1 < x_2) \implies \left(f\left(\frac{x_1 + x_2}{2}\right) \leq \frac{f(x_1) + f(x_2)}{2} \right).$$

*A real function $f(x)$ is **concave** on an interval $I \subset \mathbb{R}$ if:*

$$(\forall x_1, x_2 \in I)\,(\forall x_1 < x_2) \implies \left(f\left(\frac{x_1 + x_2}{2}\right) \geq \frac{f(x_1) + f(x_2)}{2} \right).$$

Definition 2.3.5. *Let $I \subset \mathbb{R}$ be an interval and $f(x)$ a real function on I. The point $x_0 \in I$ is called an **inflection point** of $f(x)$ if there exists $\delta > 0$ such that $f(x)$ is convex on $\langle x_0 - \delta, x_0 \rangle$ and concave on $\langle x_0, x_0 + \delta \rangle$, or vice versa.*

These definitions do not provide a direct way to find the inflection point, but the following lemma does.

Lemma 2.3.6. *Let $f(x)$ be a real function on an interval $I \subset \mathbb{R}$ and $x_0 \in I$. If $f(x)$ is twice differentiable at x_0, then x_0 is an inflection point of $f(x)$ if $f''(x_0) = 0$ and $f''(x)$ changes sign at x_0.*

Using these definitions and the given result, one can find that the inflection points are:

- When $\gamma = 0$, the inflection point is $(t_*, x(t_*)) = \left(\frac{-1}{\beta} \ln\left(\frac{\beta}{\alpha - \beta \ln(x_0)} \right), e^{\frac{\alpha - \beta}{\beta}} \right)$.

- When $\gamma \neq 0$, the inflection point is $(t_*, x(t_*)) = \left(\frac{-1}{\alpha\gamma + \beta} \ln\left(\frac{\beta\gamma}{(\alpha\gamma + \beta)x_0^{-\gamma} - \beta} \right), \left(\frac{\alpha\gamma + \beta}{\beta\gamma + \beta} \right)^{\frac{1}{\gamma}} \right)$.

Transformation of solutions using convex combinations

In this subsection, we will transform the solutions of the differential equations obtained earlier using convex combinations of certain basis functions. This reformulation is particularly advantageous for the practical part of my paper, where leveraging the properties of convex combinations will facilitate more efficient and interpretable analyses.

Definition 2.3.7. *Let $f(x)$ and $g(x)$ be real functions on an interval $I \subset \mathbb{R}$. A **convex combination** of $f(x)$ and $g(x)$ is a function $h(x)$ of the form:*

$$h(x) = \lambda f(x) + (1 - \lambda)g(x),$$

where $\lambda \in [0, 1]$.

The particular reformulation we will use has the following form:

$$x(t) = f\left(f^{-1}(x_0)e^{-\phi t} + f^{-1}(K)\left(1 - e^{-\phi t} \right) \right),$$

where $f(x)$ is an invertible function, x_0 is the initial condition, K is the carrying capacity, and ϕ is a constant.

Proposition 2.3.8. *The reformulation of the solutions of the general model are given below:*

$$f(x) = \begin{cases} x^{-\frac{1}{\gamma}}, & \gamma \neq 0 \\ e^x, & \gamma = 0 \end{cases}, \qquad \phi = \begin{cases} \alpha\gamma + \beta, & \gamma \neq 0 \\ \beta, & \gamma = 0 \end{cases}$$

Proof. We split the proof into two cases:

$\boxed{\gamma = 0}$:

$$x' = \alpha x - \beta x \ln(x)$$

$$\frac{x'}{x} = \alpha - \beta \ln(x)$$

$$\frac{d}{dt} \ln(x) + \beta \ln(x) = \alpha$$

$$\frac{d}{dt} \left(e^{\beta t} \ln(x) \right) = \alpha e^{\beta t}$$

Integrating both sides from 0 to t, we get:

$$e^{\beta t} \ln(x(t)) - e^{\beta \cdot 0} \ln(x_0) = \int_0^t \alpha e^{\beta s} ds$$

$$e^{\beta t} \ln(x(t)) - \ln(x_0) = \left[\frac{\alpha}{\beta} e^{\beta s} \right]_0^t$$

$$e^{\beta t} \ln(x(t)) - \ln(x_0) = \frac{\alpha}{\beta} \left(e^{\beta t} - 1 \right)$$

$$\ln(x(t)) = \ln(x_0) e^{-\beta t} + \frac{\alpha}{\beta} \left(1 - e^{-\beta t} \right)$$

$$x(t) = e^{\ln(x_0)e^{-\beta t} + \frac{\alpha}{\beta} \left(1 - e^{-\beta t} \right)}.$$

This means that $f(x) = e^x$ and $\phi = \beta$.

$\boxed{\gamma \neq 0}$:

$$x' = \alpha x - \beta x \frac{x^\gamma - 1}{\gamma}$$

$$x' = \alpha x + \frac{\beta}{\gamma} x - \frac{\beta}{\gamma} x^{\gamma+1}$$

$$x' x^{-1-\gamma} = \left(\alpha + \frac{\beta}{\gamma} \right) x^{-\gamma} - \frac{\beta}{\gamma}$$

Now we can substitute $y = x^{-\gamma}$ and, consequently, $y' = -\gamma x^{-\gamma-1} x'$ and $y_0 = x_0^{-\gamma}$:

$$-\frac{1}{\gamma} y' = \frac{\alpha\gamma + \beta}{\gamma} y - \frac{\beta}{\gamma}$$

$$y' e^{(\alpha\gamma+\beta)t} + (\alpha\gamma + \beta) y e^{(\alpha\gamma+\beta)t} = \beta e^{(\alpha\gamma+\beta)t}$$

$$\left(y e^{(\alpha\gamma+\beta)t} \right)' = \beta e^{(\alpha\gamma+\beta)t}$$

Integrating both sides from 0 to t, we get:

$$y e^{(\alpha\gamma+\beta)t} - y_0 = \frac{\beta}{\alpha\gamma + \beta} \left(e^{(\alpha\gamma+\beta)t} - 1 \right)$$

$$y = y_0 e^{-(\alpha\gamma+\beta)t} + \frac{\beta}{\alpha\gamma + \beta} \left(1 - e^{-(\alpha\gamma+\beta)t} \right)$$

$$x = \left(x_0^{-\gamma} e^{-(\alpha\gamma+\beta)t} + \frac{\beta}{\alpha\gamma + \beta} \left(1 - e^{-(\alpha\gamma+\beta)t} \right) \right)^{-\frac{1}{\gamma}}.$$

This means that $f(x) = x^{-\frac{1}{\gamma}}$ and $\phi = \alpha\gamma + \beta$.

$\square$

Remark 2.3.9. *One additional advantage of this reformulation is that it allows us to easily check if the carrying capacity values we have obtained earlier are correct. By substituting $t = \infty$ into the reformulated solutions, we can see that the carrying capacity values are indeed correct:*

$$f_1^{-1}(K_1) = \frac{\alpha}{\beta} \implies K_1 = e^{\frac{\alpha}{\beta}},$$

$$f_2^{-1}(K_2) = \frac{\beta}{\alpha\gamma + \beta} \implies K_2 = \left(1 + \frac{\alpha\gamma}{\beta} \right)^{\frac{1}{\gamma}}.$$

Why is the general model called general?

In the previous sections, the general model of growth was introduced and its solutions were derived. Following this, the solutions were transformed using convex combinations of certain basis functions. This reformulation, mentioned earlier, is beneficial for the practical part of this paper, and the reasons for its usefulness will now be explained.

To do so, I will first reformulate the logistic and classical von Bertalanffy models using convex combinations of certain basis functions.

Proposition 2.3.10. *The reformulations of the logistic model* $x' = \alpha_L x - \beta_L x^2$ *and the classical von Bertalanffy model* $x' = \alpha_B x^{\frac{2}{3}} - \beta_B x$ *are given below:*

$$f(x) = \begin{cases} x^{-1}, & \text{logistic model} \\ x^3, & \text{classical von Bertalanffy model} \end{cases}, \qquad \phi = \begin{cases} \alpha_L, & \text{logistic model} \\ \frac{\beta_B}{3}, & \text{classical von Bertalanffy model} \end{cases}$$

Proof. We will prove these claims separately:

Logistic Model :

$$x' = \alpha_L x - \beta_L x^2$$
$$\frac{x'}{x^2} = \frac{\alpha_L}{x} - \beta_L$$

Substituting $y = x^{-1}$ and $y' = -x^{-2} x'$, we get:

$$y' + \alpha_L y = \beta_L$$
$$y' e^{\alpha_L t} + \alpha_L y e^{\alpha_L t} = \beta_L e^{\alpha_L t}$$
$$(y e^{\alpha_L t})' = \beta_L e^{\alpha_L t}$$

Integrating both sides from 0 to t, we get:

$$y e^{\alpha_L t} - y_0 = \int_0^t \beta_L e^{\alpha_L s} ds$$
$$y = y_0 e^{-\alpha_L t} + \frac{\beta_L}{\alpha_L} \left(1 - e^{-\alpha_L t}\right)$$
$$x = \left(\frac{1}{x_0} e^{-\alpha_L t} + \frac{\beta_L}{\alpha_L} \left(1 - e^{-\alpha_L t}\right)\right)^{-1}.$$

This means that $f(x) = x^{-1}$ and $\phi = \alpha_L$.

Classical von Bertalanffy Model :

$$x' = \alpha_B x^{\frac{2}{3}} - \beta_B x$$
$$x' x^{-\frac{2}{3}} = \alpha_B - \beta_B x^{\frac{1}{3}}$$

Substituting $y = x^{\frac{1}{3}}$ and $y' = \frac{1}{3}x^{-\frac{2}{3}}x'$, we get:

$$y' + \frac{\beta_B}{3}y = \frac{\alpha_B}{3}$$

$$y'e^{\frac{\beta_B}{3}t} + \frac{\beta_B}{3}ye^{\frac{\beta_B}{3}t} = \frac{\alpha_B}{3}e^{\frac{\beta_B}{3}t}$$

$$\left(ye^{\frac{\beta_B}{3}t}\right)' = \frac{\alpha_B}{3}e^{\frac{\beta_B}{3}t}$$

Integrating both sides from 0 to t, we get:

$$ye^{\frac{\beta_B}{3}t} - y_0 = \int_0^t \frac{\alpha_B}{3}e^{\frac{\beta_B}{3}s}ds$$

$$y = y_0e^{\frac{-\beta_B}{3}t} + \frac{\alpha_B}{\beta_B}\left(1 - e^{-\frac{\beta_B}{3}t}\right)$$

$$x = \left(x_0^{\frac{1}{3}}e^{-\frac{\beta_B}{3}t} + \frac{\alpha_B}{\beta_B}\left(1 - e^{-\frac{\beta_B}{3}t}\right)\right)^3.$$

This means that $f(x) = x^3$ and $\phi = -\frac{\beta_B}{3}$.

$\square$

Now let's take a closer look at some of the reformulated solutions:

- **General model when $\gamma \neq 0$:** $x = \left(x_0^{-\gamma}e^{-(\alpha\gamma+\beta)t} + \frac{\beta}{\alpha\gamma+\beta}\left(1 - e^{-(\alpha\gamma+\beta)t}\right)\right)^{-\frac{1}{\gamma}}$,

- **Logistic model:** $x = \left(x_0^{-1}e^{-\alpha_L t} + \frac{\beta_L}{\alpha_L}\left(1 - e^{-\alpha_L t}\right)\right)^{-1}$,

- **Classical von Bertalanffy model:** $x = \left(x_0^{\frac{1}{3}}e^{-\frac{\beta_B}{3}t} + \frac{\alpha_B}{\beta_B}\left(1 - e^{-\frac{\beta_B}{3}t}\right)\right)^3$.

It is now fairly easy to see why the general model is called general. By using $\alpha = \alpha_L - \beta_L$, $\beta = \beta_L$, and $\gamma = 1$ in the general model, we get the logistic model, and by using $\alpha = \alpha_B - \beta_B$, $\beta = \frac{1}{3}\alpha_B$ and $\gamma = -\frac{1}{3}$, we get the classical von Bertalanffy model. This means that the general model is a generalization of both the logistic and classical von Bertalanffy models, among many others.

Remark 2.3.11. *I have mentioned earlier that this reformulation gives us an easy way to check if our carrying capacity values are correct. It also allows us to easily calculate the carrying capacity values for the logistic and classical von Bertalanffy models:*

$$K_L = \lim_{t\to\infty} x(t) = \lim_{t\to\infty} \left(x_0^{-1} e^{-\alpha_L t} + \frac{\beta_L}{\alpha_L} \left(1 - e^{-\alpha_L t} \right) \right)^{-1} = \frac{\alpha_L}{\beta_L},$$

$$K_B = \lim_{t\to\infty} x(t) = \lim_{t\to\infty} \left(x_0^{\frac{1}{3}} e^{-\frac{\beta_B}{3} t} + \frac{\alpha_B}{\beta_B} \left(1 - e^{-\frac{\beta_B}{3} t} \right) \right)^3 = \left(\frac{\alpha_B}{\beta_B} \right)^3.$$

Another useful property of this reformulation is that it allows us to easily find the inflection points of the logistic and classical von Bertalanffy models:

Inflection point of the logistic model: $\quad (t_*, x(t_*)) = \left(-\frac{1}{\alpha_L} \ln \left(\frac{-\beta_L}{\beta_L - \alpha_L x_0^{-1}} \right), \frac{\alpha_L}{2\beta_L} \right),$

Inflection point of the classical von Bertalanffy model:

$$(t_*, x(t_*)) = \left(-\frac{3}{\beta_B} \ln \left(\frac{\alpha_B}{3\alpha_B - 3\beta_B x_0^{\frac{1}{3}}} \right), \left(\frac{2\alpha_B}{3\beta_B} \right)^3 \right).$$

Chapter 3

Fitting the Models to the Data

3.1 Data Collection

As I previously mentioned, the data used in this paper is from the Croatian Science Foundation project "Methods Developing in Mathematical Modelling in Biology and Medicine". It consists of time series of average sizes of avascular tumours grown *in vitro*, measured at discrete time points up to the lifetime of the tumour spheroid in culture. It was collected using a combination of experimental techniques, including cell culture, microscopy, and image analysis.

Data preparation

After obtaining ethical approval and informed consent, I have analyzed chondroblastic osteosarcoma taken from a young male patient. The tumour was sampled after the chemotherapy. To have a data set that can be used for further analysis, the data was preprocessed. The process of tumour *in vitro* growth was done by a biologist. The process includes the following steps:

1. The frozen tumour cells were kept in crio tubes on -80°C.

2. Tumour cells were thawed for the experiment and placed in adherent Petri dishes.

3. Tumour cells were grown in a culture medium DMEM/F-12 (Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12) with 10% FBS and 1% penicillin-streptomycin.

4. The next steps were done in a laminar flow hood under sterile conditions.

5. When 80-90% confluence was reached, cells were washed with PBS (10mL) to remove inhibitors and dead cells.

6. The 1mL of trypsin-EDTA was added so the cells could detach from the Petri dish and left at 37°C in a humidified atmosphere with 5% CO2 for 4 minutes.

7. Medium DMEM/F-12 (5mL) was added to block the action of trypsin.

8. The cells were resuspended so they could be lifted from the surface.

9. They were transferred to a falcon tube of 15mL and centrifuged at 1300rpm for 5 minutes.

10. The supernatant was removed and the medium for spheroid growth (1mL of double concentrated DMEM/F-12 without FBS (Fetal Bovine Serum), with ITS (Insuline, Transferrin, Selenium), putrescine and progesterone) was added to the cells.

11. The cells were resuspended again.

12. The cell suspension was aliquoted in test tubes of $20\mu L$ and $20\mu L$ of trypan blue was added. It was resuspended again.

13. The $13\mu L$ of the cell suspension was put in the hemocytometer and the number of cells was counted and concentration was calculated.

14. The cell suspension was made at a concentration of 10 cells/μL. Cells were seeded in 96 well plates in a total volume of $100\mu L$. Respectively, $50 \mu L$ of cell suspension and $50\mu L$ of 2% methylcellulose. Growth factors FGF and EGF were added to the cells.

15. The plate was incubated for 24 hours.

16. The tumour cells started to grow and form a spheroid.

17. The spheroids were monitored by z-slices of $5\mu m$ thickness using a confocal microscope Leica SP8 x FLIM. 40x magnification was used. They were monitored three times a week (Monday, Wednesday, and Friday) for nearly two months. Every Monday and Friday FGF and EGF were added to the medium.

The pictures of z-slices of each spheroid were taken, and the size of the spheroid was measured using the ImageJ software. The volume of each spheroid was then calculated using the ReViMS software.
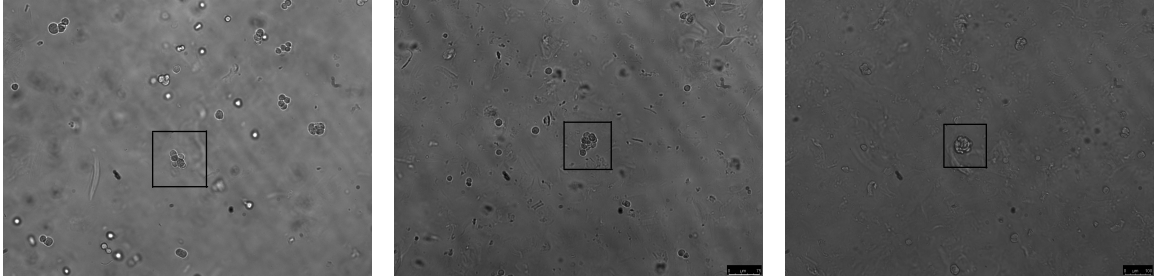

Figure 3.1: Images of the tumour spheroid at the beginning, middle, and end of the experiment

In Figure 3.1, we can see the images of the same tumour spheroid at the beginning, middle, and end of the experiment. The spheroid grows over time, and its size increases. The images were taken on the 1st, 12th and 28th day of the experiment, respectively. Each image represent one z-slice of the spheroid.

Conjecture 3.1.1. *The size of the tumour spheroid in a contolled environment follows a sigmoidal (S-shaped) curve over time.*

The S-shaped curve is a curve that resembles the letter "S". It has three parts: the initial phase where the growth is slow, the exponential phase where the growth is rapid, and the plateau phase where the growth slows down [39].

ImageJ software

After receiving the pictures of z-slices of the spheroids in .tif format, the goal was to measure the size of each spheroid. I used the ImageJ software, which is a free image-processing program developed by the National Institutes of Health. The software is widely used in scientific research for image analysis and processing. The first step for each z-slice of each spheroid was to transform the .tif image into a binary image. The threshold tool was used to segment the area of interest from the background. After filling the holes in the segmented region, the area of the spheroid z-slice was calculated by counting the number of pixels in the segmented region. The area of interest of the spheroid in each slice was then recorded in the data set. The pictures with areas of interest were saved and used for further analysis.

In Figure 3.2, we can see the z-slice of a spheroid taken with a confocal microscope. The area of interest in the z-slice obtained by the ImageJ software is shown in Figure 3.3.
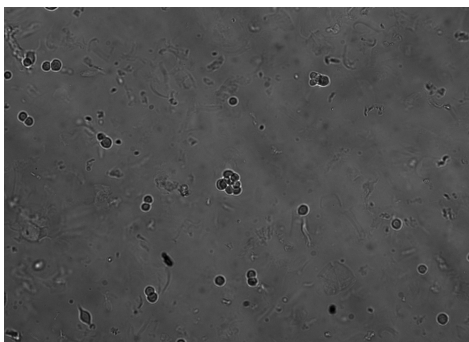
Figure 3.2: Z-slice of an osteosarcoma spheroid imaged on a Leica TCS SP8 X confocal microscope. Magnification 40x.
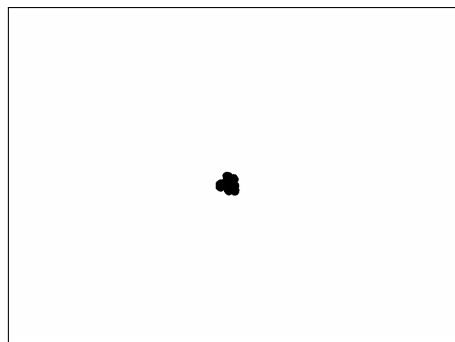

Figure 3.3: Area of interest in the picture of z-slice of a spheroid after processing in ImageJ software.

ReViMS software

After measuring the size of each spheroid in each z-slice, the next step was to calculate the volume of each spheroid. The ReViMS software was used for this purpose. ReViMS is a software tool for reconstructing 3D models from 2D images, widely used in scientific research for visualizing and analyzing complex structures in 3D. The first step was to import the pictures of the spheroid z-slices processed in the ImageJ software into the ReViMS. The software then reconstructed a 3D model of the spheroid based on the 2D images. The volume of each spheroid was calculated by the software and recorded in the data set. Reconstruction method was linear interpolation. Parameters used for the reconstruction were 0.645 μL for x-pixels and y-pixels, and 5 μm for z-distance between consecutive section.

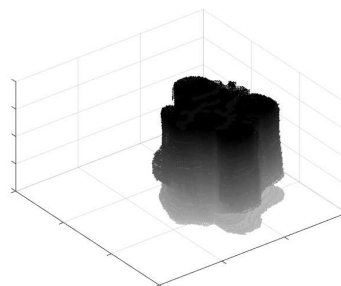Figure 3.4: Area of interest of z-slice of a tumour spheroid.

Figure 3.5: 3D model of a tumour spheroid.

Figure 3.4 shows the area of interest in the z-slice of a spheroid processed using the ImageJ software. The 3D model of the spheroid obtained by the ReViMS software is shown in Figure 3.5.

3.2 Data Set

As a basis for the further analysis, we assume that the number of cells in the spheroid is proportional to its size.

The data set consists of 16 tumour spheroids that survived (almost) the entire experiment. Each spheroid was analyzed 3 times to improve *the accuracy* of the measurements. The accuracy of a measurement is the degree of closeness of the measurement to the true value of the quantity being measured.

For each spheroid, *the average size* was calculated and recorded in the data set. The average size is the sum of all sizes divided by the number of measurements.

Average volumes (in μm^3) for 16 spheroids are represented in Table 3.1. The average volume of all spheroids at each time point is given in the last column and denoted by $\overline{V}$. Here, values are rounded to the nearest whole number. Missing values are denoted by a dash and were not included in the calculations because those spheroids were not measured at that time points. The data set from the Table 3.1 was divided by 10^4 to make the calculations easier and is used for further analysis and model fitting.

| Volume (μm^3) | Spheroid | | | | | | | | | | | | | | | | |
Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	$\overline{V}$
1.	-	23705	36349	26409	42815	10371	19319	24432	-	9110	22305	15131	21006	20093	28721	14395	**22440**
4.	-	36956	44448	31590	24520	26817	39820	25091	25499	-	19856	17939	43546	24886	25301	40138	**30458**
6.	-	55479	46190	50053	50158	53871	48753	51174	40295	40372	33303	50573	59753	43547	77397	36630	**49170**
8.	76893	87892	32990	105248	81646	50199	39806	50668	81446	72301	63817	-	104052	39432	84324	59139	**68657**
11.	58687	68369	85603	126069	61764	43346	93235	99736	65122	81847	74367	58767	97276	51582	82310	56069	**75259**
13.	75227	63724	79314	109122	76465	235573	56767	126578	75327	62726	59641	51689	55017	61460	97437	126725	**88299**
15.	59212	54491	64895	123310	57094	247182	-	56424	69041	97648	83984	46673	53986	74497	65493	50850	**80319**
18.	60498	28017	76462	122903	59986	123633	-	51555	51685	60822	28207	39934	47261	64311	142925	53710	**67461**
20.	50871	56735	64955	68933	50992	117555	-	60439	44946	136130	81070	43543	52415	77680	69215	33819	**67286**
22.	40062	41803	103796	63844	39783	106102	103785	48157	33978	-	54124	-	55953	56352	90259	76387	**65313**
25.	28807	50386	81990	32680	28607	52271	100007	65768	25590	38470	46056	84780	62956	39167	84200	63246	**55311**
27.	28245	44382	58865	35348	29894	50385	67462	58041	31494	39115	52941	95934	51501	46094	82020	73026	**52797**
29.	26079	76647	77202	17424	27846	60011	-	37709	20694	50912	87365	70234	56224	66148	81933	64225	**54710**
32.	26936	66144	100341	20565	15201	38964	-	52178	17485	42073	78774	72940	-	71362	76914	67710	**53399**
34.	14986	69570	83685	20384	15370	31348	-	-	-	24850	60987	45697	-	41169	61046	42791	**42657**
36.	16849	68356	83023	15044	16902	35880	-	-	-	37704	42059	51602	-	63665	62631	36039	**44146**
43.	11948	-	57400	9936	12189	39159	-	-	-	19087	45529	-	-	65272	53742	38143	**35240**
48.	16028	-	76663	-	17428	37972	-	-	-	25393	-	-	-	50941	53518	42641	**40073**
50.	12189	-	49259	-	15606	-	-	-	-	30356	-	-	-	33866	53465	48857	**34800**

Table 3.1: Data set of the average volume size of tumour spheroids

3.3   Data Visualization

To better understand the data set, I have visualized the data in several ways, but first I will present some of the notations used in the following sections; if n_i is the number of spheroids measured at ith time point, then $\overline{V_i}$ is the mean volume of n_i spheroids at time point t_i;

$$\overline{V_i} = \frac{1}{n_i} \sum_{j=1}^{n_i} V_{ij},$$

where V_{ij} is the volume of jth spheroid at ith time point.

The average volumes of all MTS at each time point are illustrated in Figure 3.6. Time points (in days) are represented on the x-axis, while the y-axis indicates the corresponding average volumes.

In my investigation of spheroid growth, I have chosen to focus on the first 15 days of the experiment, which span the initial 7 time points. Beyond this period, the spheroids cease to grow and begin to enter a declining phase. Examining these 7 time points, the plot shows that the average volume really follows a S-shaped curve over time.
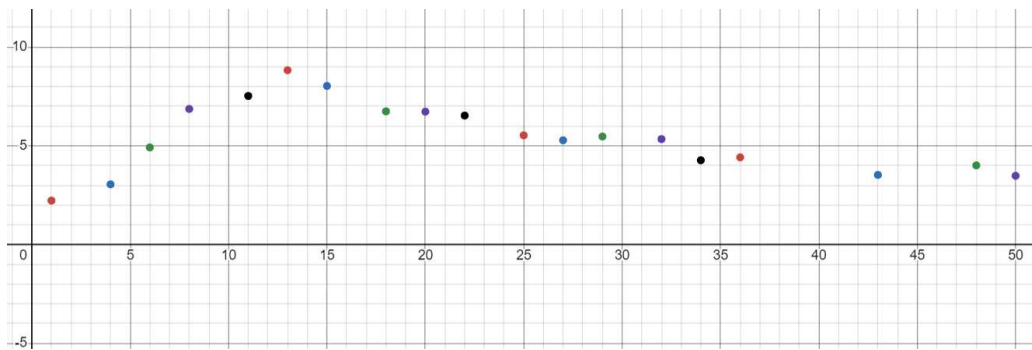

Figure 3.6: Average volume of tumour spheroids at each time point

Another interesting plot is shown in Figure 3.7. It shows the volumes of the individual 16 tumour spheroids. Each line represents the volume of one spheroid over time.
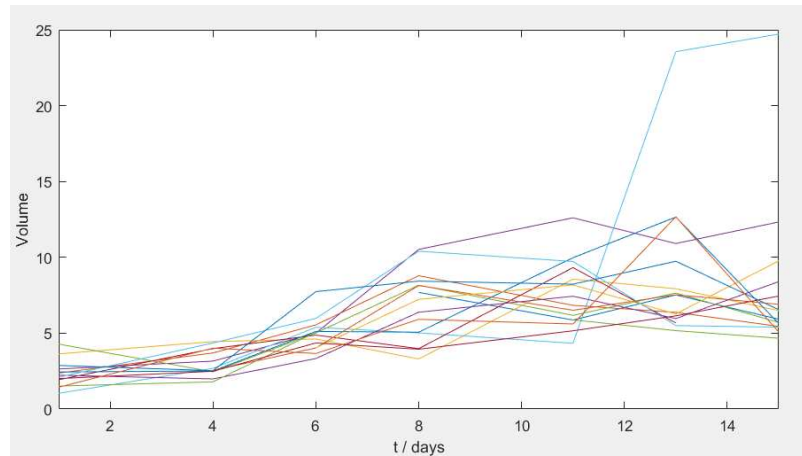
Figure 3.7: Volumes of 16 individual MTS

For further analysis, I need to list some statistical measures of the data set.

The *variance* is a measure of the dispersion or spread of a set of data points around the mean. The *standard deviation* is the square root of the variance. Variance is calculated as the average of the squared differences between each data point and the mean. The formula for the variance is:

$$\text{Var}(X) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2,$$

where x_i are data points, $\overline{x}$ is the mean of the data set, and n is the number of data points. For *sample variance* we divide by $n - 1$ instead of n.

In this paper, the standard deviation is calculated as:

$$\text{std}_i = \sqrt{\frac{1}{n_i - 1} \sum_{j=1}^{n_i} (V_{ij} - \overline{V_i})^2}$$

In Figure 3.8 we can see the *coefficients of variation* for each time point. The coefficient of variation is a measure of the relative variability of the data set around the mean. It is calculated as the ratio of the standard deviation to the mean. The plot shows that the coefficient of variation does not exhibit a consistent pattern of increase or decrease but rather fluctuates up and down.

To check if mean volumes are linearly dependent on *standard error*, I have plotted the mean volumes against the standard errors in Figure 3.9 and used *linear regression* to plot the best-fit line. Standard error (SE) is a measure of the dispersion of the sample mean around the true population mean. It is calculated as the ratio of the standard deviation to the square root of the sample size. Linear regression is a statistical method to model the relationship between a dependent variable and one or more independent variables by fitting

a linear equation to the observed data. The plot shows that there exists a linear relationship between these two variables.
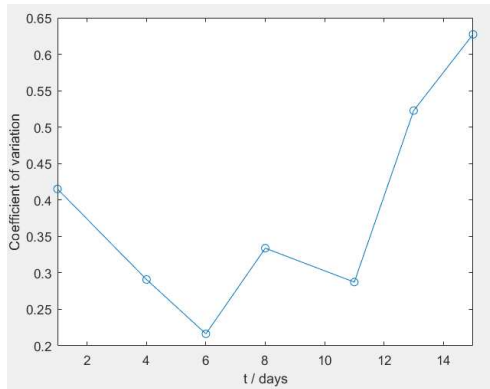

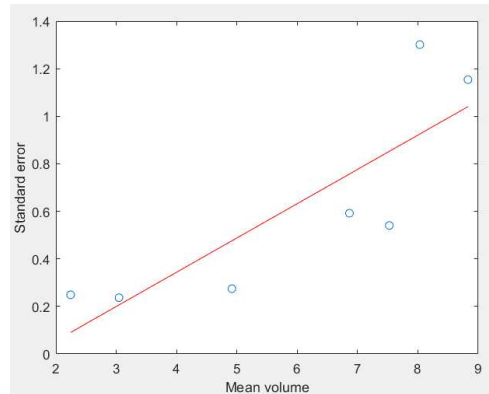Figure 3.8: Coefficient of variation for each time point

Figure 3.9: Standard error vs. mean volume

Figure 3.10, presented prior to model fitting, depicts the mean volume with intervals of $\pm 2 \cdot$SE against time on the x-axis. The goal is to identify a model that accurately encompasses these intervals. Based on this plot, it appears that the inflection point $(t_*, x(t_*))$ is in the close proximity of the point $(6, 5)$, suggesting a critical transition in growth dynamics. Moreover, the carrying capacity K seems to lie within the range of $[7, 10]$.
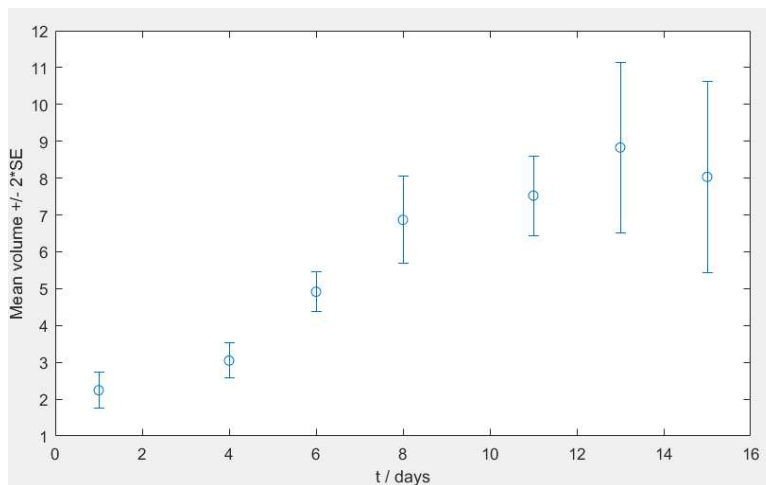

Figure 3.10: Mean volumes of 16 MTS $\pm 2 \cdot$standard errors

3.4 Methods of Fitting

In this section I will describe the methods used to fit models to the data. The analysis was conducted using the *Matlab* programming language. Given that the function $x(t|\theta)$ is nonlinear with respect to the parameters $\theta = (\alpha, \beta, x_0, \gamma)$, the method used for fitting is the *weighted nonlinear least squares method* [39]. This method is used to estimate the parameters of a nonlinear model by minimizing the sum of the squares of the differences between the observed values and the values predicted by the model. For a model given by $y = x(t|\theta)$, where y represents the predicted value, t denotes the independent variable, and θ is the vector of parameters in parameter space Θ, the method identifies the values of θ that minimize the function:

$$L(\theta) = \sum_{i=1}^{n} w_i \left(Y_i - x(t_i|\theta)\right)^2,$$

where n is the number of time points, w is the weight and Y_i is the observed value at time t_i.

If $Y_i = \overline{V}_i$, then the weight is $w_i \sim \frac{1}{\text{SE}(\overline{V}_i)^2}$, where $\text{SE}(\overline{V}_i) = \frac{std_i}{\sqrt{n_i}}$. Furthermore, if we assume that $\text{SE}(\overline{V}_i) \sim \overline{V}_i$, then we can take $w_i = \frac{1}{\overline{V}_i^2}$.

Hence, the function to minimize becomes:

$$L(\theta) = \sum_{i=1}^{n} \frac{1}{\overline{V}_i^2} \left(\overline{V}_i - x(t_i|\theta)\right)^2.$$

This method gives the vector of parameters $\hat{\theta} = (\hat{\alpha}, \hat{\beta}, \hat{x}_0, \hat{\gamma})$ such that:

$$L(\hat{\theta}) = \min_{\theta \in \Theta} L(\theta).$$

The method used in Matlab, called *fminsearch* [43], is a function designed to find the minimum of a scalar function with multiple variables, starting from an initial estimate. This function helps determine the parameters of models that best fit the data. *Fminsearch* employs the Nelder-Mead simplex algorithm, a direct search method that doesn't require the function's gradient. The algorithm operates based on a simplex, a geometric figure that extends the concept of a triangle to higher dimensions. It begins with an initial simplex, a set of points in the parameter space, and iteratively moves this simplex towards the function's minimum. During each iteration, the algorithm evaluates the function at the simplex's vertices and updates the simplex's position accordingly. The algorithm stops when the simplex converges to a small size or after a maximum number of iterations is reached. It returns the model parameters that minimize the function. This algorithm is both robust and efficient, making it suitable for fitting a wide variety of models to data.

3.5 Model Fitting

The idea was to find parameters α, β, x_0 and γ that best fit the data. Initially, I used the Gompertz, logistic and classical von Bertalanffy models to fit the data. After that, I used the general model and tried to find an even better fit. Since the *fminsearch* function in Matlab requires an initial guess for the parameters, I have chosen to use the estimated parameters of either the Gompertz, logistic or classical von Bertalanffy model in the general model, depending on which provided the best fit to the data.

For the initial guess of parameter x_0 for first three models, I used the mean volume of the spheroids at the first time point. Additionally, I applied a slightly different data transformation for each model to determine initial guesses for α and β. The *fminsearch* method was used with termination tolerance (TolFun) set to default value 10^{-4} together with maximum number of iterations (MaxIter) and function evaluations (MaxFunEvals) set both to 14000.

Gompertz model

Let's take a look on the results using the Gompertz model. The black color will be used for this model in all plots.
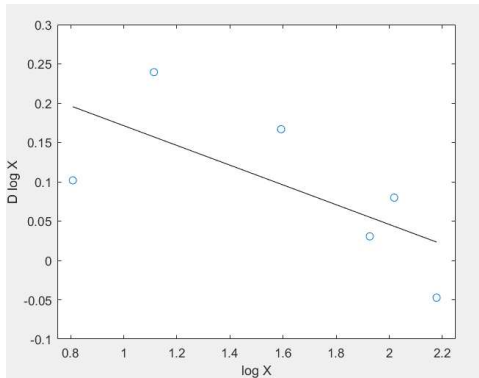

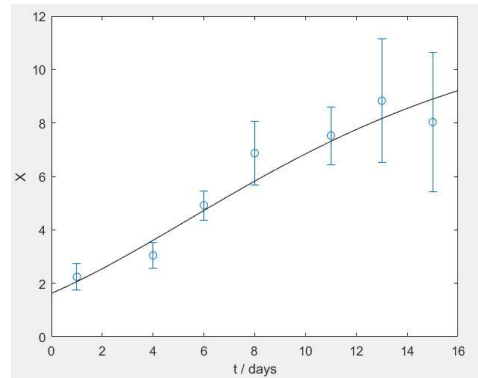Figure 3.11: Transformed data with linear regression line

Figure 3.12: Fitted Gompertz model

In Figure 3.11 we can see the line calculated using linear regression on transformed data. The intercept and slope of this line serve as initial parameters for α and β. Remember that $\gamma = 0$ in this model. X-axis represents the logarithmic transformation of mean volumes, while the y-axis shows the values calculated by the formula $y_i = \frac{\log(\overline{V}_{i+1}) - \log(\overline{V}_i)}{t_{i+1} - t_i}$.

Figure 3.12 shows the fit of the Gompertz model to the data. Parameters estimated by the *fminsearch* function are $\alpha_G = 0.3169$, $\beta_G = 0.1279$ and $x_{0_G} = 1.6249$. The value of the

criterion function is $L(\hat{\theta}_G) = 0.0831$. However, the model fits the data poorer than it was expected - it does not capture the second data point and fails to reach the plateau phase. Using the formula for carrying capacity from above, it is easy to calculate that $K_G = 11.91$ in the Gompertz model. This value could be too high for this data set. The inflection point is $(t_G, x(t_G)) = (5.3870, 4.3817)$ which is (perhaps) too early for this data set. Moreover, within this time interval, the model does not exhibit a clear S-shaped curve.

Logistic model

Considering the logistic model, which is colored in red in all plots, the results were as follows.

The line fitting the transformed data is shown in Figure 3.13. The intercept and slope of this line, along with x_0 mentioned earlier, serve as initial parameters. This model corresponds to the general model with $\gamma = 1$. In this plot, the data transformation uses the formula $y_i = \frac{\frac{1}{V_{i+1}} - \frac{1}{V_i}}{t_{i+1} - t_i}$ on the y-axis and $\frac{1}{V_i}$ on the x-axis.

In Figure 3.14, the fit of the logistic model to the data is shown. The parameters estimated by the *fminsearch* function are $\alpha_L = 0.2582$, $\beta_L = 0.0269$ and $x_{0_L} = 1.6601$, with a value of the criterion function $L(\hat{\theta}_L) = 0.0655$. The logistic model fits all data points and exhibits a nicer S-shaped curve compared to the Gompertz model. It also looks like its carrying capacity is lower than the Gompertz model's. If we calculate carrying capacity and inflection point, we get $K_L = 9.6$ which is nearly reached, and $(t_L, x(t_L)) = (6.0624, 4.8019)$. These values are more appropriate and suitable for this data set.
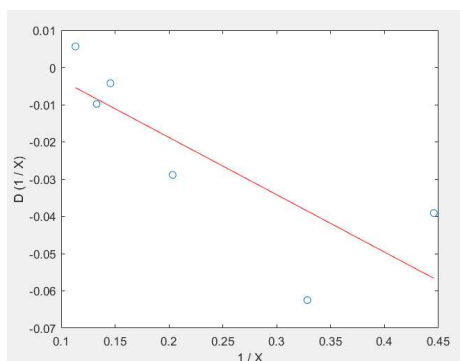

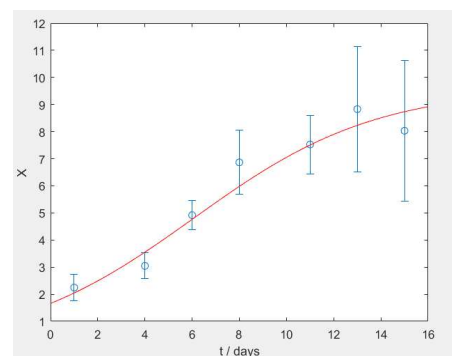Figure 3.13: Transformed data with linear regression line

Figure 3.14: Fitted logistic model

Classical von Bertalanffy model

Before we try to fit the general model, let's review the results using the classical von Berta-lanffy model, indicated by the color green in all plots.

Once again, the line fitting the transformed data is shown in Figure 3.15. The intercept and slope of this line, along with x_0, serve as initial parameters. This model is equivalent to the general model when $\gamma = -\frac{1}{3}$. The plot shows the transformation of the data using the formula $y_i = \frac{\overline{V}_{i+1}^{\frac{1}{3}} - \overline{V}_i^{\frac{1}{3}}}{t_{i+1} - t_i}$ on the y-axis and $\overline{V}_i^{\frac{1}{3}}$ on the x-axis.

The fit of the classical von Bertalanffy model to the data is shown in Figure 3.16. The parameters estimated by the *fminsearch* function are $\alpha_B = 0.6285$, $\beta_B = 0.2605$ and $x_{0_B} = 1.6106$. The value of the criterion function is $L(\hat{\theta}_B) = 0.0896$. However, this model does not capture second data point, it does not reach the plateau phase nor has a clear S-shaped curve. Its carrying capacity is $K_B = 14.04$, which is the highest among all models and, based on the plot, it is too high for this data set. The inflection point is $(t_B, x(t_B)) = (4.9906, 4.1613)$, which is, like in the Gompertz model, (perhaps) too early for this data set.
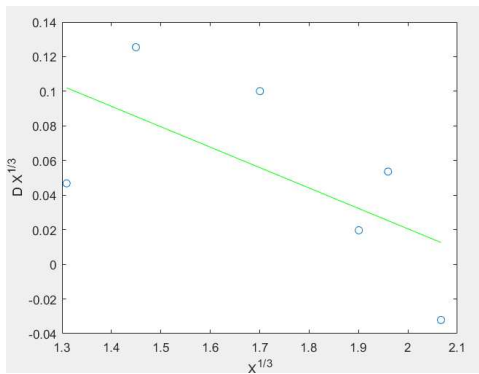

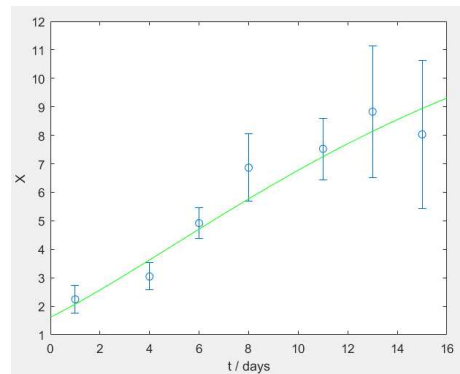Figure 3.15: Transformed data with linear regression line

Figure 3.16: Fitted classical von Berta-lanffy model

Comparison of the models

Upon initial inspection, the logistic model appears to best fit the data. To compare them easier, let's take a look at the plots of the fitted models in Figure 3.17. Once again, black represents the Gompertz model, red represents the logistic model, and green represents the classical von Bertalanffy model. The logistic model stands out as the most suitable because its carrying capacity is well-matched to the data and nearly reached, it fits all data points, exhibits the nicest S-shaped curve, and its inflection point appears most appropriate.

Another helpful criterion for decision-making is the analysis of *residuals*. Residuals are the differences between observed values and those predicted by the model. If a model fits the data well, residuals should be randomly scattered around zero. The standardized residual plots are shown in Figure 3.18. All residuals show a similar pattern, prompting further analysis to determine which model's residuals perform best.
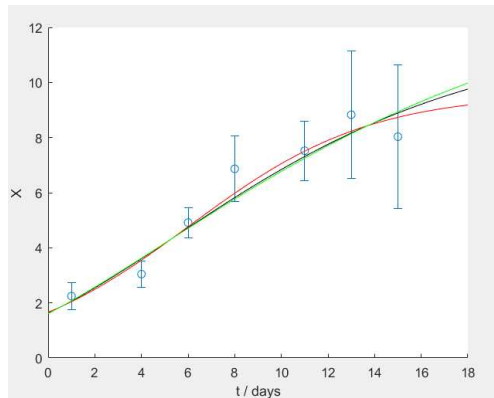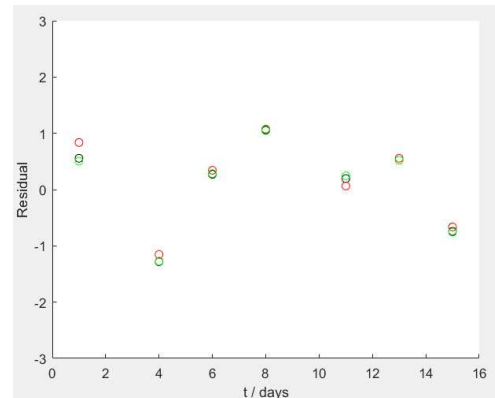

Figure 3.17: Fitted models

Figure 3.18: Standardized residuals of the fitted models

I have calculated the *sum of squared residuals (SSR)*, *mean squared error (MSE)* and *mean absolute error (MAE)* for each model. SSR measures the total sum of the squared differences between actual and predicted values; $SSR = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$. It is used to assess the overall error of the model. MSE represents the average of these squared differences; $MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$, providing a measure of the average squared error. MAE is the average of the absolute differences between actual and predicted values; $MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$, indicating the average absolute error of the model. Smaller values indicate better performance for all three metrics.

The results are summarized in Table 3.2. The logistic model has the smallest SSR, MSE and MAE, making it the best model for this data set.

Model	SSR	MSE	MAE
Gompertz	2.7021	0.3860	0.5308
Logistic	2.0650	0.2950	0.4606
classical von Bertalanffy	2.9920	0.4274	0.5616

Table 3.2: Results of the residuals analysis

General model

Given that residual analysis cleared up the preference, I have decided to use the logistic model as the initial guess for the general model.

Unlike previous models, the general model requires initial guesses for all four parameters instead of three. The initial guesses for α, β and x_0 are estimated parameters from logistic model using $\alpha = \alpha_L - \beta_L$, $\beta = \beta_L$ and $x_0 = x_{0_L}$, while the initial guess for γ is 1 because that is the value of γ in the logistic model. The general model is represented in blue in all plots.

Figure 3.19 shows the fit of the general model to the data, while Figure 3.20 represents both logistic and general model on the same plot. The values of the parameters are $\alpha = 0.1654$, $\beta = 3.5278e - 29$, $x_0 = 1.7620$ and $\gamma = 32.1934$ and the value of the criterion function is $L(\hat{\theta}) = 0.0351$. The plateau phase is reached and carrying capacity is $K = 8.0604$, while the inflection point is $(t, x(t)) = (8.5421, 7.2296)$ which is (perhaps) a bit late. Based on the plots, the general model fits the data a bit different than the logistic model. It reaches the plateau phase before and the inflection point is later, but both models fit all data points, and have a nice S-shaped curve.
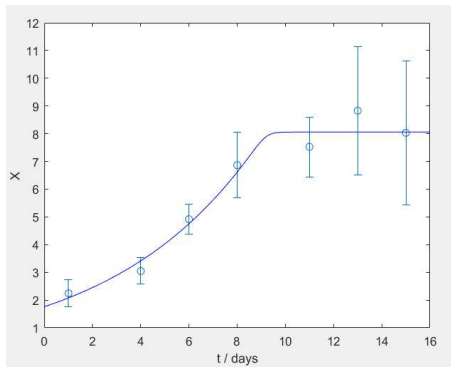

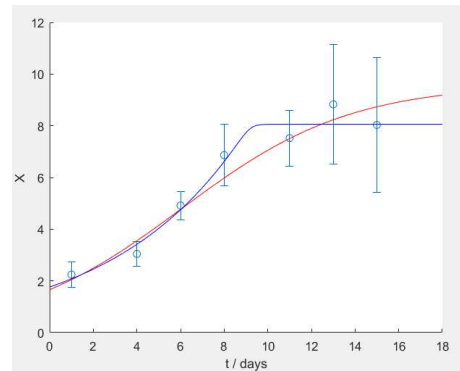Figure 3.19: Fitted general model

Figure 3.20: Fitted logistic and general models

Let's check the residuals in Figure 3.21. The SSE,, MSE and MAE are calculated and presented in Table 3.3. The general model has smaller SSR, MSE and MAE, making it better than the logistic model.
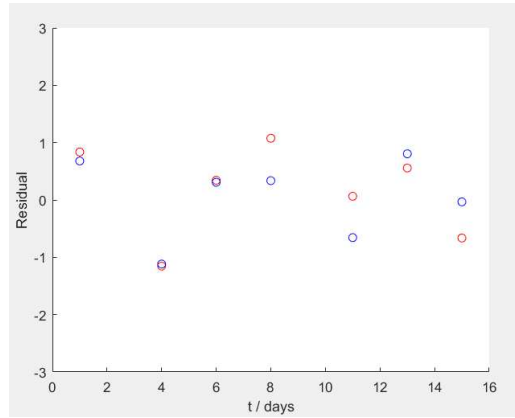
Figure 3.21: Standardized residuals of logistic and general model

Model	SSR	MSE	MAE
Logistic	2.0650	0.2950	0.4606
General	1.1313	0.1616	0.3258

Table 3.3: Results of the residuals analysis for logistic and general model

Since both models fit the data well, it is challenging to determine whether the general model is truly better than the logistic model for this data set, especially given the small number of data points. To check if there is a statistically significant difference between these two models, let's see the results of *post hoc* analysis.

3.6 *Post hoc* Analysis

The first analysis I have used is the *F-test* [39]. F-test is a statistical test used to compare the variances of two samples. The null hypothesis is that the parameter γ is equal to the logistic model's γ and the alternative hypothesis is that it is not;

$$H_0 : \gamma = 1$$
$$H_1 : \gamma \neq 1.$$

F-test is calculated as:

$$f = \frac{\frac{L_0(\hat{\theta}_0) - L(\hat{\theta})}{df1}}{\frac{L(\hat{\theta})}{df2}} \overset{H_0}{\sim} F(df1, df2),$$

where $L_0(\hat{\theta}_0)$ is the value of the criterion function for the logistic model with $\hat{\theta}_0 = (\hat{\alpha}_0, \hat{\beta}_0, \hat{x_0}_0)$, $L(\hat{\theta})$ is the value of the criterion function for the general model with $\hat{\theta} = (\hat{\alpha}, \hat{\beta}, \hat{x}_0, \hat{\gamma})$, $df1$

is the difference in the number of parameters between the general and logistic models, and $df2$ is the number of time points minus the number of parameters in the general model. Here, $df1 = 4 - 3 = 1$ and $df2 = 7 - 4 = 3$. Following that, we easily calculate f:

$$f = \frac{\frac{0.0655 - 0.0351}{1}}{\frac{0.0351}{3}} = 2.5983.$$

The critical value is $F(1, 3) = 10.128$ using the F-distribution table with the level of significance $\alpha = 0.05$. *P-value* is calculated as $P(F > f = 2.5983|H_0) = 0.2054$. A p-value is a measure that indicates the strength of evidence against the null hypothesis in statistical hypothesis testing. If the p-value is below a certain significance level (here 5%), it suggests that the observed data is statistically significant. This means there is sufficient evidence to reject the null hypothesis in favor of an alternative hypothesis, implying that the results are unlikely to have occurred by chance alone. Since $p > \alpha$, we fail to reject the null hypothesis. Also, the calculated f is less than the critical value, so we can conclude that there is no statistically significant difference between the general and logistic models in terms of their effectiveness.

There is one more interesting calculation to do. That is the 95% confidence intervals for the parameters θ. To do that, we used bootstrap method [42]. Using estimated parameters $\hat{\theta} = (\hat{\alpha}, \hat{\beta}, \hat{x}_0, \hat{\gamma})$ of the general model, we created new data:

$$V^*_{ki} = x(t_i|\hat{\theta}) + \hat{\sigma} \cdot \overline{V_i} \cdot \varepsilon^*_{ki},$$

where $k = 1, 2, ..., 999$ is the number of bootstrap samples, $i = 1, 2, ..., 7$ is the number of time points, $x(t_i|\hat{\theta})$ is the volume predicted by the general model at time t_i with parameters $\hat{\theta}$, $\hat{\sigma}$ is the square root of the criterion function in general model over the number of data points minus the number of estimated parameters; $(\hat{\sigma} = \sqrt{L(\hat{\theta})/(7 - 4)})$, $\overline{V_i}$ mean volume at time t_i and ε^*_{ki} is a random number from the standard normal distribution $N(0, 1)$, different for each k and i.

Then we used the *fminsearch* function to find the parameters of the general model for each bootstrap sample and stored it in the vector $\hat{\theta}^*$. The initial parameters used for *fminsearch* call were the estimated parameters of the general model with the change of $\hat{x}_0$ for V^*_{k1}. The MaxIter and MaxFunEvals were changed to 200000. We sorted the vector $\hat{\theta}^*$ for each parameter and calculated the 2.5% and 97.5% quantiles. To calculate the confidence intervals for θ, we used the formula:

$$\hat{\theta}^*_{(25)} - \hat{\theta} \le \hat{\theta} - \theta \le \hat{\theta}^*_{(975)} - \hat{\theta}.$$

Parameter	Lower bound	Upper bound
α	0.091503	0.216434
β	0	7.1e-29
x_0	1.330777	2.248796
γ	-1	61.301633

Table 3.4: Confidence intervals for the parameters of the general model

The results, using previously mentioned conditions $\beta > 0$ and $\gamma > -1$, are shown in Table 3.4. The confidence intervals for α, β and x_0 are narrow, while for γ it is wider. As we can see, $\gamma = 1$ is within limits of 95% CI for γ.

3.7 Conclusion

To conclude, both the logistic and general models fit the data better than Gompertz and classical von Bertalanffy models. To decide which of these two models is better, the *post hoc* analysis was used and showed that there is no statistically significant difference between them. The general model has 4 parameters, while the logistic model has 3, so the general model is more complex and it is better and easier to use the logistic model for further analysis. It is important to note that this analysis is based on only 7 time points and 16 MTS, suggesting the need for more data to make more meaningful conclusions.

Bibliography

[1] M. Kapalczynska, T. Kolenda, W. Przybyla, M. Zajaczkowska, A. Teresiak, V. Filas, M. Ibbs, R. Blizniak, L. Luczewski, and K. Lamperska. 2D and 3D cell cultures - a comparison of different types of cancer cell cultures. *Archives of Medical Science*, Vol. 14, No.4:910–919, 2018.

[2] J. J. van Middendorp, G. M. Sanchez, and A. L. Burridge. The Edwin Smith papyrus: a clinical reappraisal of the oldest known document on spinal injuries. *Spine*, Vol. 19, No.11:1815–1823, 2010.

[3] R. A. Norman. Dr Rudolph Virchow, the Father of Pathology. *Cambridge Scholars Publishing*, 2022.

[4] M. Sen and S. G. Honavar. Wilhelm Conrad Röntgen: Finding X. *Indian Journal of Ophthalmology*, Vol. 69, No.10:2570–2572, 2021.

[5] National Cancer Foundation. What is cancer? `https://www.cancerfoundation.org.au/what-is-cancer-.html`. Accessed: 2024-05-11.

[6] A. Fritz, C. Percy, A. Jack, K. Shanmugaratnam, L. Sobin, D. Max Parkin, and S. Whelan. International Classification of Diseases for Oncology. *World Health Organization*, 2013.

[7] Nacional Cancer Institute. Cancer classification. `https://training.seer.cancer.gov/disease/categories/classification.html#:~:text=Cancers%20are%20classified%20in%20two%20ways%3A%20by%20the,first%20method%3A%20cancer%20classification%20based%20on%20histological%20type`. Accessed: 2024-06-14.

[8] Stephanie Langmaid. Stages of Cancer. `https://www.webmd.com/cancer/cancer-stages`. Accessed: 2024-05-11.

[9] Douglas Hanahan and Robert A. Weinberg. Hallmarks of cancer: The next generation. *Cell,*, Vol. 144, No. 5:646–674, 2011.

[10] Mayo Clinic. Skin Cancer. `https://www.mayoclinic.org/diseases-conditi ons/skin-cancer/symptoms-causes/syc-20377605`. Accessed: 2024-05-12.

[11] Mayo Clinic. Lung Cancer. `https://www.mayoclinic.org/diseases-conditi ons/lung-cancer/symptoms-causes/syc-20374620`. Accessed: 2024-05-12.

[12] Mayo Clinic. Breast Cancer. `https://www.mayoclinic.org/diseases-condi tions/breast-cancer/symptoms-causes/syc-20352470`. Accessed: 2024-05-12.

[13] Mayo Clinic. Prostate Cancer. `https://www.mayoclinic.org/diseases-c onditions/prostate-cancer/diagnosis-treatment/drc-20353093`. Accessed: 2024-05-12.

[14] Mayo Clinic. Colon Cancer. `https://www.mayoclinic.org/diseases-condi tions/colon-cancer/symptoms-causes/syc-20353669`. Accessed: 2024-05-12.

[15] Mayo Clinic. Leukemia. `https://www.mayoclinic.org/diseases-condition s/leukemia/symptoms-causes/syc-20374373`. Accessed: 2024-05-12.

[16] Mayo Clinic. Osteosarcoma. `https://www.mayoclinic.org/diseases-condi tions/osteosarcoma/symptoms-causes/syc-20351052`. Accessed: 2024-05-12.

[17] Oncohema Key. Cancer Epidemiology: Incidence and Etiology of Human Neoplasms. `https://oncohemakey.com/cancer-epidemiology-incidence -and-etiology-of-human-neoplasms/`. Accessed: 2024-05-12.

[18] Kendall K. Morgan. Cancer Incidence Rates by Age. `https://www.webmd.com/ cancer/cancer-incidence-age`. Accessed: 2024-05-12.

[19] B. A. Lindsey, J. E. Markel, and E. S. Kleinerman. Osteosarcoma Overview. *Springer*, Vol. 4, No.1:25–43, 2017.

[20] H. C. Beird, S. S. Bielack, A. M. Flanagan, J. Gill, D. Heymann, K. A. Janeway, J. A. Livingston, R. D. Roberts, S. J. Strauss, and R. Gorlick. Osteosarcoma. *Nature Reviews Disease Primers*, Vol. 8, No.77, 2022.

[21] Mount Sinai. Osteosarcoma. `https://www.mountsinai.org/health-library /diseases-conditions/osteosarcoma`. Accessed: 2024-05-13.

[22] A. Righi, A. Paioli, A. P. Dei Tos, M. Gambarotti, E. Palmerini, M. Cesari, E. Marchesi, D. M. Donati, P. Picci, and S. Ferrari. High-grade focal areas in low-grade central osteosarcoma: high-grade or still low-grade osteosarcoma? *Clinical Sarcoma Research*, Vol. 5, No.23, 2015.

[23] H. K. Brown, M. Tellez-Gabriel, and D. Heymann. Cancer stem cells in osteosarcoma. *Cancer Letters*, Vol. 386:189–195, 2017. `https://www.sciencedirect.com/science/article/pii/S0304383516307078?via%3Dihub`, Accessed: 2024-05-16.

[24] M.T. Santini, G. Rainaldi, and P. L. Indovina. Multicellular tumour spheroids in radiation biology. *International Journal of Radiation Biology*, Vol. 75:787–799, 1999. `https://www.tandfonline.com/doi/abs/10.1080/095530099139845`, Accessed: 2024-05-15.

[25] M.A.J. Chaplain. Avascular growth, angiogenesis and vascular growth in solid tumours: The mathematical modelling of the stages of tumour development. *Mathematical and Computer Modelling*, Vol. 23, No.6:47–87, 1996.

[26] Wikipedia. History of cancer. `https://en.wikipedia.org/wiki/History_of_cancer`. Accessed: 2024-05-11.

[27] Nacional Cancer Institute. Cancer Staging. `https://www.cancer.gov/about-cancer/diagnosis-staging/staging`. Accessed: 2024-05-11.

[28] Cleveland Clinic. Tumor. `https://my.clevelandclinic.org/health/diseases/21881-tumor`. Accessed: 2024-05-11.

[29] National Cancer Institute. What is Cancer? `https://www.cancer.gov/about-cancer/understanding/what-is-cancer`. Accessed: 2024-05-11.

[30] American Cancer Society. What is Cancer? `https://www.cancer.org/cancer/understanding-cancer/what-is-cancer.html`. Accessed: 2024-05-11.

[31] A. I. Baba and C. Câtoi. *Comparative Oncology*, chapter 3 - TUMOR CELL MORPHOLOGY. 2007. `https://www.ncbi.nlm.nih.gov/books/NBK9553/`, Accessed: 2024-05-11.

[32] Wikipedia. Osteosarcoma. `https://en.wikipedia.org/wiki/Osteosarcoma`. Accessed: 2024-05-13.

[33] P. Mamachan, V. Dang, N. S. Bharadwaj, N. DeSilva, and P. Kant. Chondroblastic osteosarcoma—a case report and review of literature. 2019. `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7669408/`, Accessed: 2024-05-16.

[34] M. Barisam, M. S. Saidi, N. Kashaninejad, and N.-T. Nguyen. Prediction of necrotic core and hypoxic zone of multicellular spheroids in a microbioreactor with a u-shaped barrier. *Micromachines*, 2018. `https://www.ncbi.nlm.nih.gov/pmc/article s/PMC6187679/`, Accessed: 2024-05-15.

[35] Wikipedia. Cell cycle. `https://en.wikipedia.org/wiki/Cell_cycle`, Accessed: 2024-05-15.

[36] Ž. Bajzer, M. Marušić, and S. Vuk-Pavlović. Conceptual frameworks for mathematical modeling of tumor growth dynamics. *Mathematical and Computer Modelling,*, Vol. 23, No. 6:31–46, 1996.

[37] Ž. Bajzer, S. Vuk-Pavlović, and M. Huzak. Mathematical Modeling of Tumor Growth Kinetics. in J.A. Adam and N. Belomo (eds.) A survey of models for tumor-immune system dynamics. *Birkhauser*, 1997.

[38] R. B. Banks. *Growth and Diffusion Phenomena. Mathematical Frameworks and Applications*. Springer-Verlag, Berlin, 1994.

[39] M. Marušić. Mathematical models of tumor growth. *Mathematical Communications,*, Vol. 1, No. 2:175–188, 1996.

[40] A. K. Laird. Dynamics of tumour growth: Comparison of growth rates and extrapolation of growth curve to one cell. *British Journal of Cancer*, Vol. 19, No.2:278–291, 1965.

[41] B. Guljaš. Matematička analiza I & II. 2018, Zagreb.

[42] A. C. Davison and D. V. Hinkley. *Bootstrap Methods and Their Application*. Cambridge University Press, Cambridge, 1997.

[43] MathWorks. Fminsearch. `https://www.mathworks.com/help/matlab/ref/fm insearch.html`. Accessed: 2024-06-26.

Summary

In this thesis, the aim was to analyze the growth of tumour spheroids using the Gompertz, logistic, classical von Bertalanffy and general models of growth. The data set contains average volumes of 16 multicellular tumour spheroids of chondroblastic osteosarcoma tumours measured at 7 time points. The models were fitted to the data and compared. The Gompertz and classical von Bertalanffy models provided poor fits to the data, while the logistic and general models fit the data well. To decide which model is the best, the *post hoc* analysis was used. The F-test showed that there is no sufficient evidence to conclude that the models are significantly different in terms of their effectiveness. Since the logistic model has one parameter less than the general model, it is more parsimonious and should be used for further analysis.

Biography

I was born on April 30, 1998, in Zagreb. I attended Krunoslav Kuten Elementary School in Vrbovec from 2005 and Vrbovec High School, general gymnasium programme, from 2013. During my elementary and high school years, I enjoyed participating and achieved success in numerous mathematics competitions. Therefore, after completing high school education in 2017, I enrolled in the Undergraduate University Study of Mathematics at the Faculty of Science in Zagreb. I completed my undergraduate studies in 2022, thus becoming a Bachelor of Mathematics. After that, I enrolled in the Biomedical Mathematics graduate programme in English at the same faculty.