TUM School of Computation, Information and Technology

TECHNISCHE UNIVERSITÄT MÜNCHEN

# Mathematical modeling and lung cancer: quantifiable prognostic models improve clinical routine care

Pirmin Schlicke

Vollständiger Abdruck der von der TUM School of Computation, Information and Technology der Technischen Universität München zur Erlangung eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Daniel Matthes
Prüfer der Dissertation:
1. Prof. Dr. Christina Kuttler
2. Prof. Dr. Heiko Enderling
3. Prof. Dr. Felix Krahmer

Die Dissertation wurde am 02.05.2022 bei der Technischen Universität München eingereicht und durch die TUM School of Computation, Information and Technology am 16.11.2022 angenommen.

Remember to look up at the stars and not down at your feet. Try to make sense of what you see and wonder about what makes the universe exist. Be curious. And however difficult life may seem, there is always something you can do and succeed at.

It matters that you don't just give up.

- Stephen Hawking

*In loving memory of my mother*

# Acknowledgments

First and most importantly I thank Prof. Dr. Christina Kuttler for introducing me to the exciting world of biomathematics, especially cancer dynamics, and for the amazing opportunity to perform research and teaching under her supervision. The years that I worked on this thesis and at TUM were full of her inestimable support, not exclusively related to this work. Thank you for always having an open door for me and my questions and always being there when help in any direction was needed.

I am tremendously grateful to Prof. Dr. Heiko Enderling and Prof. Dr. Felix Krahmer for taking their valuable time reviewing this thesis and to take part in the examination committee as well as Prof. Dr. Daniel Matthes for taking the chair of the committee.

Meeting Sébastien Benzekry, PhD, HDR at a conference during the early stage of my PhD studies turned out to be incredible luck for me. Even though our on-site cooperation in Bordeaux was cancelled on short notice due to the Covid-pandemic, our planned project remotely took place and turned out to produce very interesting results. Thank you so much for introducing me to novel methods and constantly providing constructive comments and help throughout our project.
I am incredibly thankful to Prof. Dr. Schumann, Dr. Pascale Tomasini and Eléonore Simon as medical doctors providing access to clinical data that the method in this work was calibrated to - despite their full schedule. Without their help (that even took place during the Covid-pandemic!) the clinical use of these approaches would probably never have been unveiled.

The great working atmosphere at the chair of mathematical modeling (M6) played a huge role in the success of my PhD studies. I thank all current and former colleagues contributing to this environment. Special thanks go to Dr. Carl-Friedrich Kreiner, happening to be the teaching assistant in my very first lectures when I started my bachelor's studies at TUM ten years ago, ending up sharing his lunch time with me.

I want to thank my friends and my brother Marcel for their help and their open ears, especially towards the end of my PhD studies, reminding me that there is life happening off the thesis.
I am deeply grateful to my parents Manfred and Antonie for their encouragement and letting me chose independently what to do in my life, for enabling me to study mathematics and their constant support during all that time. Unfortunately, it was not granted to my mother to see me finishing my PhD studies.
But I know you were with me, all the time.

# Abstract

The thesis at hand examines different approaches to mathematically model the metastatic process in cancer diseases with ordinary and partial differential equations. Extensions allow to incorporate treatment possibilities and comparison to clinical data. Statistical evaluations prove the increased predictive value of the approaches in identifying high-risk patients on an individual basis.

Die vorliegende Arbeit untersucht verschiedene Ansätze, um den Metastasierungsprozess bei Krebserkrankungen mathematisch zu modellieren und nutzt dabei gewöhnliche sowie partielle Differentialgleichungen. Erweiterungen des Modells erlauben die Erfassung von Therapiemöglichkeiten und den Abgleich mit klinischen Daten. Statistische Auswertungen beweisen den erhöhten prognostischen Nutzen dieser Ansätze um Hochrisikopatienten auf einer individuellen Ebene zu erkennen.

# Publications by the author

[Ben+22] S. Benzekry*, P. Schlicke*, P. Tomasini, and E. Simon. "Mechanistic modeling of brain metastases in NSCLC provides computational markers for personalized prediction of outcome." In: *submitted* (2022).

[SKS21] P. Schlicke, C. Kuttler, and C. Schumann. "How mathematical modeling could contribute to the quantification of metastatic tumor burden under therapy: insights in immunotherapeutic treatment of non-small cell lung cancer." In: *Theoretical Biology and Medical Modelling* 18.1 (Dec. 2021), p. 11. ISSN: 1742-4682. DOI: 10.1186/s12976-021-00142-1.

# Contents

# 1 Introduction

## 1.1 Motivation

Cancerous diseases are among the most frequent diseases in the Western world. In the year 2013 about half a million Germans have been diagnosed cancer and about quarter a million Germans have died directly related to a cancerous disease, with tendency lately increasing [Koc16]. Despite cancer as a disease group is responsible for about 10 million deaths annually according to the WHO, the disease itself remains incompletely understood.

The technical and scientific advances of the last decades yield more treatment options and higher treatment success than ever before, culminating in public attention with the 2018 Nobel prize in Physiology or Medicine awarded to James P. Allison and Tasuku Honjo for the identification of the PD-1/PD-L1 pathway as target structure for several immunotherapeutic drugs.
The metastatic process is yet not fully understood nor predictable, but about 90% of cancer-related deaths are directly caused by metastatic lesions, most of which are not detectable with modern clinical imaging techniques at early stage diagnosis [CW11; LPW17]. Remarkably, even though the technical advances of the last decades yield better diagnosis accuracy and treatment success than ever before, the quantitative forecasting possibilities for cancerous diseases are to date relatively poor. Many mathematical models have shed light into specific dynamics of emerging cancerous diseases and optimized treatment. However, there is to date still no framework capable of integrating the vast amount of clinical data for proper disease prognosis [GM03]. This draws attention to mathematical tools for improved estimation of the not detectable metastatic threat to aim at improved diagnostic and prognostic methodologies and thus well-informed clinical decisions for improved treatment strategies and patient care. In this work, several approaches are examined that not only allow for quantified estimations of number and size of metastases during the clinical time course but also show the prognostic potential identifying high-risk patients already early in treatment.

## 1.2 Outline

After a short introduction into the biological background of this work, i.e. a brief presentation into the main lung cancer histology types, the metastatic process, treatment possibilities and the data sets analysed in the corresponding publications of this thesis, an overview over the mathematical tools is given. These include formal definitions of ordinary and partial differential equations, different local and global sensitivity analysis approaches and statistical evaluations of survival analysis with the Kaplan-Meier-method, Cox proportional hazards model and concordance indices to determine the predictive value of these regression models. All mathematical models analysed and used on the data are properly built up in chapter four: the untreated and the treated growth dynamics for the primary tumor alone as well as the metastatic density distribution modeling approach with and without therapy. A slightly simplified model is also introduced for this purpose for the statistical evaluations. All these models are properly presented in terms of implementation and general simulation as well as analysed for their dynamical behavior in chapter five. Chapter six shows the respective models applied to clinical data and presents how well they are able to describe these data. The statistical analysis at the end of that chapter gives proof of the significant increase in clinical prediction possibility for these models applied in clinical routine and evaluation of their model parameters as computationally generated biomarkers pointing on helpful clinical information. A short discussion is presented thereafter.

# 2 Biological principles

The data sets examined in this work consist of patients that had lung cancer diagnosed. This chapter briefly introduces the important biological background to understand the dynamics of lung cancer, the metastasizing process and the currently available treatment options before the two examined data sets are described. The chapter strongly builds on the overview published in [SKS21].

## 2.1 Lung cancer

The globally leading cancer-related mortality cause is lung cancer. Most of the primary malignant lung tumors are carcinomas that can be divided into two major groups: non-small cell lung cancer (NSCLC, about 85% of all lung cancer cases) and small cell lung cancer (SCLC). The five year survival probability for NSCLC is approximately 22% [Bra+18]. Typically, NSCLC can be subdivided into the three main histological subgroups of adenocarcinoma (ADC), squamous cell carcinoma (SCC) and large cell carcinoma [Ken+08; Tra+15], that are determined by immunohistological testing. Other, less frequent subgroups such as neuroendocrine tumors are not considered here; throughout this work only patient data with NSCLC is examined.

**Adenocarcinoma.** This most frequent histology type of lung cancer is usually derived from secretory epithelial cells and therefore mainly located in the lung periphery [Wei06; HTC20]. It is also the most frequent tumor type of lung cancer patients without smoking history [Wak+07]. Its genetic profile is rather heterogeneous but several biomarkers have been identified useful for diagnosis. The pneumocyte marker TTF1 is expressed in about 80% of primary lung ADC [Hec+01]. If this biomarker test is negative, the marker Napsin A is recommended for another biomarker test to identify ADC [Suz+05].
Further identified biomarkers such as p53-expression and KRAS or EGFR mutations correlate with worse survival prognosis and worse efficacy of classic chemotherapy [Sle+90; Lee+16; Lyn+04]. On the other hand, genetic alterations in EGFR, EML4-ALK and Ros1 (among others) show increased efficacy and better survival prognosis during respective targeted therapy approaches [Lyn+04; Mak+18; HTC20].

**Squamous cell carcinoma.** SCC is the second-most common histology type of lung cancer. Its incidence is strongly associated with smoking history and higher ratio for men than for women [Der+15]. Morphologically, SCC is defined by showing cornification or intracellular bridges and its proliferation rate is high [HTC20]. Molecular markers to identify differentiation within the group of SCC are cytokeratins such as CK5/6, CK14 and proteins such as p40 and p63. Testing recommendations in clinical routine include the ADC biomarkers TTF1 and Napsin A for negativity, to exclude ADC. Differential diagnosis towards SCLC can be performed with immunohistochemical analysis of the Chromogranin A, Synaptophysin and CD56 biomarkers [HTC20].

**Adenosquamous carcinoma.** Adenosquamous carcinoma (ASC) are tumors that show components of both ADC and SCC. By clinical definition, the tumor has to consist of at least 10% of its volume by these two carcinoma, respectively [Tra+15]. Therefore, the diagnosis of ASC is only possible considering the tumor resectate [Tra+11]. ASC is most frequently diagnosed in male smokers [Ish+92] and shows survival prognosis worse than those of ADC or SCC alone [Fil+11].

## 2.2 Metastatic process

Lung cancer is very likely to metastasize. Cancer-related deaths are mostly driven by existence of metastases [Ste16]. They form from cancer cells that leave the tumors via blood and lymphatic vessels and move to other distant sites in the body [Wei06]. After extravasation at these sites the metastatic cells construct tumor micro environments and induce angiogenesis to boost their proliferation [Fid03]. This process is referred to as colonization and the established daughter tumors are called metastases [Wei06].
Due to immune system activity, only a minority of these metastatic cells even survive the travel through the body's vessel. However, a diagnosis of metastases in cancerous diseases results in significantly lowered survival times [Ste16; Luz+98]. But most metastases remain unidentified at primary diagnosis of the disease due to technical limitations [Wei06; GM06; CW11]. The metastases are simply too small to be traceable with imaging techniques such as computed tomography (CT) and magnetic resonance imaging (MRI). Still, they can be of considerable number already at primary diagnosis and their overall summed volume (the total metastatic burden) is of high importance for treatment options and survival outcomes [PCF99; Oh+09].

## 2.3 Treatment

Classic local treatment on primary tumors is surgical excision. However, treatment options including all metastases as targets at the same time are systemic treatments. For lung cancer, these are mainly chemo- and immunotherapy. If the mutational profile consists of non-altered genes, proteins and pathways that would allow a targeted therapy, usually a PD-1/PD-L1-pathway targeting immunotherapy is administered right away (first line therapy) or subsequently after a platinum-based chemotherapy (second-line therapy) [HTC20].

**Chemotherapy.** Chemotherapeutic agents trigger cytostatic or cytotoxic effects on different pathways of cells, ideally on cancerous cells with minimal adverse effects to healthy cells [JRL02; Meh15]. As adverse effects in chemotherapy application are usually very common, many different combination therapies of classical drugs and newer targeted therapy methods have been used as a standard of care for different cancerous diseases [Meh15]. Still, the common theory behind the quantitative effects of chemotherapy is originally formed by the so-called 'log cell kill' theory [SSW64; Ski86]. It states that an applied chemotherapy dose kills a constant fraction of tumor cells independent of the actual tumor size during a certain fixed amount of time. Experimental regimes have shown this behavior in mice, in very contrast to human cancer settings [Nor+76]. For human tumors, it was shown that the chemotherapy exhibits effects in tumor regression proportional to the growth rate of an untreated tumor of this size and that the tumor size also depends on the integrated drug effect during the course of actual treatment [NS86; SN06]. This allows for an explanation of refractory effects termed as 'kinetic resistance' that are widely observed in clinical chemotherapy applications on tumors reaching a small size [Fru+10].

**Immunotherapy.** Growth dynamics of tumors and their metastatic seeding process is not only subject to the individual cell characteristics. In living organisms, it was shown that these behaviors are additionally influenced by the immune system [HW11]. Indirect outcomes on the tumor size by enhanced antitumoral activity of the immune system has become a major key point of modern cancer treatment, the so-called immunotherapy [Egg12; He+15; SFR16]. For different cancer types immunotherapy was proven to show significant clinical benefits in patients with advanced stages of cancer and is well-established as standard treatment [RK15; Kea16; Iwa+17]. Generally, the immunotherapeutic treatment can be distinguished into four major subgroups: the active non-specific immunotherapy (treatment e.g. via cytokines), the active specific immunotherapy (vaccines), the passive immunotherapy (monoclonal antibodies) and approaches to block immune escape mechanisms (e.g. CTLA-4) [Ber+17]. For lung

cancer, immunotherapeutic treatment is often times passive immunotherapy, i.e., monoclonal antibodies that regulate immune checkpoint inhibitors. These inhibitors are expressed by tumor cells to manipulate immune checkpoints to decrease T-cell activity and suppress immune responses [Par12; QP13].
The PD-1/PD-L1-pathway is a specific immune checkpoint that regulates T-cell activity in the effector phase of the immune response. Its downregulation caused by PD-1 activation in the peripheral tissue usually prevents collateral damage during an immune response [QP13; Rib12]. Tumor cells can manipulate this pathway by expressing the ligands PD-L1 and PD-L2. These ligands bind to the T-cells PD-1 receptor and inactivate the T-cell to decrease the immune response towards the tumor cells [Par12; QP13; Roz+12; PK15].
Many cancer cells express PD-L1. This made it an attractive target for the immunotherapeutic treatment approach [TDP15]. An overview of some clinically approved immunotherapeutic monoclonal antibodies and their respective characteristic measures are shown in table 2.1.

Table 2.1: Some clinically approved monoclonal antibodies used as immunotherapeutic drugs targeting the PD-1/PD-L1 pathway. Taken from [SKS21].

| Antibody | Type | Year of approval (FDA) | Molar Mass [kDa] | Approved Dosages | Half-life [d] |
|---|---|---|---|---|---|
| Atezolizumab | humanized | 2016 | 145 | 1200mg q3w | 27 |
| Avelumab | human | 2017 | 143 | 800mg q2w | 6.1 |
| Durvalumab | human | 2017 | 146 | 10mg/kg q2w | 18 |
| Nivolumab | human | 2014 | 146 | 240mg q2w or 480mg q4w | 26.7 |
| Pembrolizumab | humanized | 2014 | 146 | 200mg q3w or 400mg q6w | 22 |
| Cemiplimab | human | 2018 | 144 | 350mg q3w | 19.4 |

q2w = quaque two weeks, every two weeks; analogous for other numbers

Non-detected metastases influence therapeutic outcomes since decisions in the treatment protocol are also subject to the number and size of metastases as well as their properties [Wei06; PCF99; Koy+08; Boe+09]. Therefore, proper estimation of the total tumor burden in a patient's body is of uttermost importance.

## 2.4 Data presentation

The first data set (published in [SKS21]) contains three patients with NSCLC and features volumetric size measurements of primary tumors and clinically detected metastases as well as the therapy schedule. Some details are shown in table 2.2.

Table 2.2: Data set one: patient-specific clinical parameters. Taken from [SKS21].

| Patient | KE-01 | KE-02 | KE-03 |
|---|---|---|---|
| Sex | F | M | M |
| Histology | Adenosquamous carcinoma | Adenocarcinoma | Adenocarcinoma |
| Molecular Pathology | EGFR-, PD-L1 5, CK7+, TTF1+, p63+, Chromogr-, Syn- | EGFR-, ALK-, KRAS-, BRAF-, PD-L1 50, Ros1- | EGFR-, ALK-, KRAS-, BRAF-, PD-L1 5 |
| Size of pt at diagnosis [ml] | 240.74 | 14.27 | 71.75 |

The three patients with NSCLC of this data set were routinely treated with a prior Cisplatin/Pemetrexed Chemotherapy and/or 1L/2L immunotherapy (Pembrolizumab or Nivolumab), as highlighted in table 6.1. The molecular pathology features diverse histological testing of tumor markers as follows: EGFR = Epidermal Growth Factor Receptor, PD-L1 = Programmed Death Ligand 1, CK7 = Cytokeratin 7, TTF1 = Thyroid Transcription Factor 1, p63 = Tumor Protein 63, Chromogr = Chromogranin A, Syn = Synaptophysin, ALK = EML4-ALK fusion protein, KRAS = Kirsten Rat Sarcoma, BRAF = rapid accelerated fibrosarcoma (B-Type), Ros1 = rather often translocated in sarcoma. The sign "+" or "-" indicates a positive respective negative test result for the gene translocations and upregulations.

All three patients were routinely treated in the Clinic of Pneumology, Thoracic Oncology, Sleep and Respiratory Critical Care of the Klinikverbund Allgäu, Germany. Data use was approved a priori by the ethical commission of BLAEK (Ethik-Kommission der Bayerischen Landesärztekammer), reference number 19021. The corresponding volumetric data were calculated from the appraisal environment *syngo.CT* LCAD of *syngo.Via* VB40, Siemens Healthineers, from the CT slices that were acquired in routine clinical care. These volumetric data can be rescaled into measurements of cell numbers by applying the conversion rule $10^{-3}$ ml $= 1$ mm$^3$ $= 10^6$ cells [Kle09; SMS95].

The second data set (submitted in [Ben+22]) contains 31 patients with NSCLC. Here, longitudinal diameter measurements were determined from corresponding CT and

MRI images of primary tumors and their corresponding brain metastases as well as the therapy schedule and critical time points such as time to relapse and time of death. Three of these patients had no known primary tumor size measurement at primary diagnosis and had to be dropped during the evaluation phase. The longitudinal measurements were converted to cell number measurements by assuming a spheroid shape, calculating the volume as $vol(\delta) = \frac{4}{3}\pi \left(\frac{\delta}{2}\right)^3$ with $\delta$ the respective diameter measurement. Again, cell numbers were evaluated from this volumetric measurement using the conversion rule 1 $mm^3 \mathrel{\widehat{=}} 10^6$ cells. All these clinical histories are listed in the appendix of [Ben+22], an example patient's clinical history is shown in Figure 6.3, subplot A. Further details to these patients are shown in table 2.3.
Also here, all patient data were collected in clinical routine care at the Multidisciplinary Oncology and Therapeutic Innovations Department, Assistance Publique - Hôpitaux de Marseille, Aix Marseille University in Marseille, France. The non-interventional retrospective study presented in [Ben+22] did not require opinion of a CPP in accordance with the requirements of the Jarde 2016 law regarding studies qualified as internal by the CNIL.

Table 2.3: Data set two: patient overview. Taken from [Ben+22].

| | | |
|---|---|---|
| Sex | Female | N=10 (32%) |
| | Male | N=21 (67%) |
| Age at diagnosis | $\leq 40$ years | N=2 (6%) |
| | $> 40, \leq 60$ years | N=13 (42%) |
| | $> 60, \leq 80$ years | N=14 (45%) |
| | $> 80$ years | N=2 (6%) |
| Histology of primary tumor | adenocarcinoma | N=20 (65%) |
| | squamous cell carcinoma | N=9 (29%) |
| | other | N=2 (6%) |
| Stage at diagnosis | I/II | N=20 (65%) |
| | III | N=11 (35%) |
| Size of primary tumor at diagnosis (diameter value) | $\leq 25$ mm | N=14 (45%) |
| | $> 25, \leq 50$ mm | N=6 (19%) |
| | $> 50, \leq 75$ mm | N=7 (23%) |
| | $> 75$ mm | N=1 (3%) |
| | unknown | N=3 (10%) |
| Number of brain metastases at relapse | 1 | N=18 (58%) |
| | 2 | N=5 (16%) |
| | 3 | N=8 (26%) |
| Volume of brain metastases at relapse (sum of diameters) | $\leq 20$ mm | N=18 (58%) |
| | $> 20, \leq 40$ mm | N=9 (29%) |
| | $> 40, \leq 60$ mm | N=3 (10%) |
| | $> 60$ mm | N=1 (3%) |
| Surgery on primary tumor (before relapse) | yes | N=22 (71%) |
| | no | N=9 (29%) |

# 3 Mathematical background

This chapter discusses the mathematical tools used. Deterministic mathematical models as dynamic and time-dependent processes are often formulated as systems of differential equations. A brief overview for those is given in the two first sections. The explanation for the tools for sensitivity analysis and statistical survival analysis are presented thereafter.

## 3.1 Ordinary Differential Equations

A time-dependent process can be expressed with the model output's change between two different time points. Let $x(t) : \mathbb{R} \to \mathbb{R}^n$ be the model output at time $t \in \mathbb{R}$ with $n \in \mathbb{N}$. The difference $x(t_1) - x(t_0)$ of the model output's value at two different times $t_1$ and $t_0$ can be expressed by its change

$$\frac{x(t_1) - x(t_0)}{t_1 - t_0}. \tag{3.1}$$

The derivative of $x(t)$ with respect to $t$ evaluated at time $t_0$ is the growth at time $t_0$ and defined as the corresponding limit

$$\dot{x}(t_0) := \lim_{t_1 \to t_0} \frac{x(t_1) - x(t_0)}{t_1 - t_0}, \tag{3.2}$$

if the limit on the right hand side exists. This derivative is frequently also written as $\frac{dx}{dt} = \dot{x}(t)$.
A first order ordinary differential equation of inhomogeneous type with an initial condition $x(t_0) = x_0 \in \mathbb{R}^n$ is of the form

$$\dot{x}(t) = g(t)x(t) + h(t) \tag{3.3}$$

with some continuous functions $g : D \to \mathbb{R}$ and $h : D \to \mathbb{R}^n$ on an open domain of definition $D \subset \mathbb{R}$. This problem has an unique solution defined on $D$ which can be calculated from the method of *variation of constants* and is given as

$$x(t) = x_0 e^{\int_{t_0}^{t} g(\tau)d\tau} + \int_{t_0}^{t} e^{\int_{t_0}^{t} g(\tau)d\tau - \int_{t_0}^{s} g(\tau)d\tau} h(s)ds. \tag{3.4}$$

A proof for this method can be found in various literature on ordinary differential equations, e.g. [Wal93].

## 3.2 Partial Differential Equations

Partial differential equations are used to model dynamics, where the change of some model output $u(t,x) : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^n$ with $m, n \in \mathbb{N}$ is subject to multiple variables such as time $t \in \mathbb{R}$ and space $x \in \mathbb{R}^m$. In this case, the derivatives read

$$\begin{aligned} \frac{\partial u}{\partial t}(t,x) &= \lim_{\Delta t \to 0} \frac{u(t+\Delta t, x) - u(t,x)}{\Delta t} \\ \frac{\partial u}{\partial x}(t,x) &= \lim_{\Delta x \to 0} \frac{u(t, x+\Delta x) - u(t,x)}{\Delta x} \end{aligned} \tag{3.5}$$

and the general form of a linear partial differential of first order reads

$$f(t,x)\frac{\partial u}{\partial t}(t,x) + g(t,x)\frac{\partial u}{\partial x}(t,x) + h(t,x)u(t,x) = 0. \tag{3.6}$$

To solve this analytically, one can apply the so-called *method of characteristics* on this equation with an initial condition $u(t_0, x) = u_{t_0}(x)$ and a boundary condition $u(t, x_0) = u_{x_0}(t)$ with $t_0, x_0 \in \mathbb{R}^n$ [Zau89]. The method of characteristics directly uses of the so-called characteristics of such an equation, which roughly correspond to the projection of the solution $u(t,x)$ on the $(t,x)$-plane.

## 3.3 Sensitivity Analysis

Sensitivity analysis is a mighty methodology to determine the influence of parameter uncertainty intervals and the effects of parameter variation on the behavior of dynamical systems, in this context outputs and predictions of a mathematical model. It is of high interest to know the model behavior after perturbations in the input parameter values. Saltelli et al. [Sal04] define sensitivity analysis as *'the study of how uncertainty in the output of a model [...] can be apportioned to different sources of uncertainty in the model input'*, in contrast to uncertainty analysis that in turn aims at quantifying uncertainty in the model input [Sal08].
Sensitivity analysis can be of local or of global nature, depending on the point of reference that sensitivity is calculated of. For local sensitivity analysis, this is one single point for which changes in one or more parameter values are examined. However, this brings limitations in sensitivity interpretation, since the sensitivities are only calculated with respect to exactly this base line point. Any slight deviation from this base line could potentially lead to very different sensitivity estimations. Still, for models of ordinary differential equations, it is the method of choice by its simple implementation and fast application as well as the background, that the base line reference value can be reviewed with the technical interpretation in mind. In contrast to that, global sensitivity

analysis determines conclusions for the global parameter range space. This of course comes at a larger computational time needed for estimation as well as with more complex methods of implementation.

### 3.3.1 Local sensitivity analysis

A typical way to investigate local sensitivity for ordinary differential equation models is the so-called derivative-based local method. It investigates the changes in the model output under local first-order changes of a certain parameter, which coincide with the first derivative evaluated at the point of interest, i.e., the reference base line point. This is of course highly applicable to a model compounded of differential equations. There exists a huge variety of methods to consider small parameter changes and their relation to the resulting effects on the equation's trajectories.
In the following, a direct method (also known as *variational method*) from the field of mathematical theory based on the representation of sensitivities as time-dependent trajectories is introduced [TV72; DG76; RCM15].
Assume the surveyed model is a system of general first order ordinary differential equations depending on a vector of parameters $\Theta \in \mathbb{R}^m, m \in \mathbb{N}$. Let the system be of the form

$$\frac{\partial X(t,\Theta)}{\partial t} = f(X(t,\Theta),\Theta). \tag{3.7}$$

A solution of this equation is a vector $X(t,\Theta) \in \mathbb{R}^n, n \in \mathbb{N}$, forming the model output and consisting of $n$ state variables $x_j(t,\Theta) \in \mathbb{R}, j \in \{1,...,n\}$ that depend on time $t$ and a vector $\Theta \in \mathbb{R}^m, m \in \mathbb{N}$ which in turn consists of $m$ different time-independent parameters $\Theta_i \in \mathbb{R}, i \in \{1,...,m\}$. The initial condition $X(t_0,\Theta) = X_0$ of the ordinary differential equation (3.7) can be interpreted as a parameter of the model, i.e. $\Theta_{k_i} = X_0 \in \mathbb{R}^n$ for some $i$ with $k_i \in \{1,...,m\}$ $\forall 1 \leq i \leq m$. By this, the effects of varying initial conditions on the model output can be determined as well.
A parameter whose small variations result in large changes in the solution of equation (3.7) is considered a *sensitive* parameter and a *nonsensitive* one otherwise.
If the ordinary differential equation (3.7) cannot be solved explicitly one can use a workaround as follows: by Taylor's formula a sufficiently small change[1] with respect to a certain parameter vector component $\Delta\Theta_i$ within the system of ordinary differential equations can be expressed as

$$X(t,\Theta + \Delta\Theta_i) \approx X(t,\Theta) + \sum_{i=1}^{m} \frac{\partial X(t,\Theta)}{\partial \Theta_i} \Delta\Theta_i \tag{3.8}$$

[1] Small enough in terms of linear extrapolation.

The sensitivity of the solution $X(t,\Theta)$ of equation (3.7) depending on perturbations of the parameter $\Theta_i$ is a vector defined as

$$S_{\Theta_i}(t) = \frac{\partial X(t,\Theta)}{\partial \Theta_i}. \tag{3.9}$$

For an analytical expression of $S_{\Theta_i}$ one can use its time-dependence:

$$\dot{S}_{\Theta_i} = \frac{\partial S_{\Theta_i}}{\partial t} = \frac{\partial \left(\frac{\partial X(t,\Theta)}{\partial \Theta_i}\right)}{\partial t} = \frac{\partial \left(\frac{\partial X(t,\Theta)}{\partial t}\right)}{\partial \Theta_i} \tag{3.10}$$

by the theorem of Fubini, the time-independence of the parameter vector $\Theta$ and each of its components $\Theta_i$. Use equations (3.7,3.10) as well as the chain rule to arrive at

$$\begin{aligned} \dot{S}_{\Theta_i} &= \frac{\partial (f(X(t,\Theta),\Theta))}{\partial \Theta_i} = \frac{\partial f(X(t,\Theta),\Theta)}{\partial \Theta_i} + \sum_{i=1}^{n} \frac{\partial f(X(t,\Theta),\Theta)}{\partial X(t,\Theta)} \frac{\partial X(t,\Theta)}{\partial \Theta_i} = \\ &= \frac{\partial f(X(t,\Theta),\Theta)}{\partial \Theta_i} + \sum_{i=1}^{n} \frac{\partial f(X(t,\Theta),\Theta)}{\partial X(t,\Theta)} S_{\Theta_i} = f_{\Theta_i} + J\, S_{\Theta_i} \end{aligned} \tag{3.11}$$

with a vector $f_{\Theta_i} \in \mathbb{R}^n$ having components $\frac{\partial f_j}{\partial \Theta_i}$ and the Jacobian matrix $J \in \mathbb{R}^{n\times n}$ of the original system in equation (3.7). The $(k,l)$ elements of $J$ are given by $\frac{\partial f_k}{\partial x_l}$. This yields the scalar representations[2]

$$S_{x_j\Theta_i} = \frac{\partial x_j(t,\Theta_i)}{\partial \Theta_i} \text{ and} \tag{3.12}$$

$$\dot{S}_{x_j\Theta_i} = \frac{\partial f_j(X(t,\Theta),\Theta)}{\partial \Theta_i} + \sum_{k=1}^{n} \frac{\partial f_j(X(t,\Theta),\Theta)}{\partial x_k(t,\Theta)} S_{x_k\Theta_i}, \tag{3.13}$$

of which the latter one can be computed numerically if an analytic solution of equation (3.12) is not available.
Sensitivities of different parameters are difficult to compare with each other as the absolute value of the parameters and the model outputs might be subject to huge differences in orders of magnitude. To receive a measurement about the effect of a relative change of certain objective parameters introduce the concept of elasticities that can be interpreted as relative sensitivities. The elasticity of a positive differentiable function $X_j$ of a positive parameter $\Theta_i$ describes how fast the relation of the parameter of interest and the objective function changes and is defined as

[2]Recall the indices $i \in \{1,...,m\}, m \in \mathbb{N}$ and $j \in \{1,...,n\}, n \in \mathbb{N}$.

$$E_{X_j\Theta_i} = \lim_{\Delta\Theta_i \to 0} \frac{\Delta X_j / X_j}{\Delta\Theta_i / \Theta_i} = \frac{\Theta_i}{X_j} \frac{\partial X_j(t, \Theta_i)}{\partial \Theta_i} = \frac{\Theta_i}{X_j} S_{\Theta_i}(t), \tag{3.14}$$

by applying the definition of sensitivity of equation (3.9).

### 3.3.2 Global sensitivity analysis

Though the above presented approach based on derivatives is without doubt very efficient in terms of computation and thus frequently used, the crucial limitation of it is the fact, that it needs a reference point at which the derivatives can be computed. The derivative-based approach does not provide any information on the behavior of other combinations in the parameter space even slightly off from this reference point. But this is the most interesting question which is especially of concern under uncertainty in the input parameters in nonlinear systems [Sal08]. Today, there is some variety of different methods to conduct such a *global* sensitivity analysis, depending on the analysis' objectives [Sal04; IL15; CSC11].
In the following we use an adjusted version of the elementary effects method introduced by Morris in 1991 [Mor91]. The method adheres to the concept of local variation around a previously defined reference point, but allows wider ranges of variations of the input parameters and is not restricted on a certain sample point. It allows to classify the inputs into three different groups: the group of inputs of negligible effects, inputs of large linear effects without interactions and inputs of large non-linear or interaction effects. As the method's name might indicate, it forms a special type of a *one-at-a-time design* and relies on the determination of *elementary effects* by varying one parameter of interest on a previously discretized range in the reasonable parameter space and keeping all other parameters fixed.

For $m, n \in \mathbb{N}$ consider a model function $f : \mathbb{R}_0^+ \times \mathbb{R}^m \to \mathbb{R}^n$ returning a measurable model output feature $X \in \mathbb{R}^n$ of the form

$$X = f(t, \Theta) \tag{3.15}$$

with an input vector $\Theta$ that consists of $m$ independent input parameters $\Theta_i \in \mathbb{R}$ for $i \in \{1, ..., m\}$. Now discretize the input parameter space $\mathbb{R}^m$ into $k \in \mathbb{N}$ selected levels in each dimension, i.e. into a $m$-dimensional $k$-level grid $\Omega$, and sample $r \in \mathbb{N}$ realizations of parameter vectors $\Theta^{(j)} \in \Omega$ with $j \in \{1, ..., r\}$.
For a model of the form (3.15) the $j$'th realization for an elementary effect $EE_i^{(j)}$ is defined in terms of equation (3.16) using the perturbed input parameter component $\Theta_i^{(j)}$ for a fixed sampled parameter vector $\Theta^{(j)}$. We need that the parameter vector

perturbed in the $i$'th component, $(\Theta^{(j)} + \Delta e_i)$, is still contained in the discretized input parameter space $\Omega$:

$$EE_i^{(j)} = \frac{f(\Theta^{(j)} + \Delta e_i) - f(\Theta^{(j)})}{\Delta}, \tag{3.16}$$

with $\Delta$ denoting a predetermined multiple of $\frac{1}{k-1}$ as perturbation distance and $e_i$ the canonical unit vector that consists purely of zeros but a one on the $i$'th component.

The distribution $F_i$ of the elementary effects $EE_i$ with perturbed input parameter component $\Theta_i$ can be computed by random sampling of different parameter vectors $\Theta$ from the interior grid points of the set $\Omega$ [3]. The sensitivity measures for perturbations in the component $\Theta_i$ proposed by Morris are the mean $\mu_i$ and the standard deviation $\sigma_i$ of the distribution $F_i$ of those elementary effects. This means that to ensure the reliance of this approach, one needs a large sample size of different perturbations for each input parameter component.
The interpretation of the obtained sensitivity measures is more or less intuitive: the mean $\mu_i$ explains the overall and expected influence of the perturbation in parameter vector component $i$ on the model output $X$ while the standard deviation $\sigma_i$ is a measure of nonlinear or interaction influences of the $i$'th input parameter vector component on the model output $X$. A high standard deviation $\sigma_i$ indicates that the elementary effects relative to the input parameter component $\Theta_i$ deviate strongly from each other, thus the elementary effect highly differs from the respective baseline sample points. This means that the model outcome will depend notably on the choice of the other input parameter components. A low standard deviation $\sigma_i$ however displays similar values of elementary effects, which in turn implies that the effects of perturbations of the input parameter component $\Theta_i$ is almost independent of the choice of values for the other input parameter components.

To prevent type II errors identifying sensitive components $\Theta_i$, i.e., to prevent the non-identification of an influential parameter component, Morris suggested to always consider the measures $\mu_i$ and $\sigma_i$ simultaneously. As a type II error would occur when perturbations in one input parameter component yield different signs in the elementary effect, the absolute value of $\mu_i$ is very low whereas the value of $\sigma_i$ is remarkable. As argued by [CCS07], this potentially leads to problems for complex and large model settings with many model outputs. They provided an alternative measure $\mu_i^*$, which is the mean of the absolute values of elementary effects $EE_i$. Still, it is suggested to

[3] The original Morris method provides a scheme of parameter trajectories, such that the elementary effects are always computed in a random parameter component $\Theta_i$ from each iteration step. This can be conferred to as a 'random walk' on the grid points of the discretized parameter space. We stick to random sampling in the parameter space, as computational cost was not the primary point of interest.

compute the other sensitivity measures aswell to gather the maximum of information out of the elementary effects method [Sal08]. For example, if the value of $\mu_i$ and $\mu_i^*$ are both high, the perturbation of the $i$'th input parameter component $\Theta_i$ is not only high, but also the sign of these effects is the same. In contrast, if $\mu_i^*$ is high and $\mu_i$ is low, then the perturbation mentioned has high effects with different signs depending on the points in the space $\Omega$ at which these effects are computed. An overview of the formal definition of all three mentioned sensitivity measures is given in the equation set (3.17).

$$\begin{aligned} \mu_i &= \frac{1}{r}\sum_{j=1}^{r} EE_i^{(j)} \\ \mu_i^* &= \frac{1}{r}\sum_{j=1}^{r} |EE_i^{(j)}| \\ \sigma_i &= \sqrt{\frac{1}{r}\sum_{j=1}^{r}\left(EE_i^{(j)} - \mu_i\right)^2} \end{aligned} \tag{3.17}$$

The measures $\mu_i$ and $\mu_i^*$ return the average value and the average absolute value of the qualitative influence of the parameter $\Theta_i$ on the model output respectively and therefore yield valuable insight in the model behavior under parameter perturbations. A high value of $\mu_i^*$, i.e. a value fundamentally different from zero, indicates that the parameter $\Theta_i$ has an important influence on the reference model output. The measure of spread $\sigma_i$ gives an idea about nonlinear effects of the parameter $\Theta_i$ or interaction effects of parameter $\Theta_i$ with other parameters. A high value of $\sigma_i$ indicates that parameter $\Theta_i$ highly affects the reference model output in a manner of nonlinear effects or interactions with other parameters $\Theta_j \neq \Theta_i$. Two different parameters are said to interact with each other if their combined effect on the model output value of interest cannot be expressed as a sum of both single effects [Sal04; Sal08]. When parameters with small mean absolute values $\mu_i^*$ are identified, the respective parameters may be fixed to any value of their uncertainty range without losing too many information of the model output. But a concrete quantification of the nonlinear or interaction effects aside from the standard deviation of a certain parameter is not possible with this approach.

To overcome this restriction, introduce the concept of the decomposition of variances, i.e. the variance for every parameter determined with the approach above gets split into every possible interaction of first order, second order and higher order. The total order sensitivity describes the model output variance depending on all these perturbations [Sob01]. The underlying theory is the following:
Let $f : \mathbb{R}^m \to \mathbb{R}^n$ with $f \in L^2$, i.e. let $f$ be a square-integrable function, and be defined on the hypercube $[0,1]^m$ with $m \in \mathbb{N}$. By [Hoe48] this function can be rewritten as a

sum of elementary functions, where every factor is a function depending solely on its index:

$$f(\Theta) = f_0 + \sum_{i=1}^{m} f_i(\Theta_i) + \sum_{i<j}^{m} f_{i,j}(\Theta_i, \Theta_j) + ... + f_{1,2,...,m}(\Theta) \tag{3.18}$$

with

$$\begin{aligned} f_i(\Theta_i) &= \int f(\Theta) \prod_{k \neq i} dx_k - f_0, \\ f_{i,j}(\Theta_i, \Theta_j) &= \int f(\Theta) \prod_{k \neq i,j} dx_k - f_0 - f_i(\Theta_i) - f(\Theta_j) \end{aligned} \tag{3.19}$$

and so on. Following [Sob93] and using $\{i_1, ..., i_s\} \subseteq \{1, ..., m\}$ this decomposition is unique if and only if

$$\int_0^1 f_{i_1,...,i_s}(\Theta_{i_1,...,i_s}) d\Theta_{i_k} = 0, \text{ with } 1 \leq k = \{i_1, ..., i_s\} \leq s. \tag{3.20}$$

Property (3.20) is equivalent to the condition that the expansion's mean is equal zero, i.e. $\int f(\Theta_i) d\Theta_i = 0$. Then we have that $\int f(\Theta_i) f(\Theta_j) d\Theta_i d\Theta_j = 0$ for all parameter pairs $\Theta_i, \Theta_j$ with $i \neq j$, which in turn means that we can calculate the terms of $f_0, f_i, f_{i,j}$ etc. by observing the conditional expectation of the model output [Sal08].
In practice, the function $f$ can always be defined on the hypercube $[0,1]^m$ with $m \in \mathbb{N}$ by tranformation of the parameter vector range to the interval $[0,1]$ in each component. Define $V_{i_1,...,i_s} := \int f^2_{i_1,...,i_s}(\Theta_{i_1,...,i_s}) d\Theta_{i_1} ... d\Theta_{i_s}$ as the partial variance of the parameter subset $\{\Theta_{i_1}, ..., \Theta_{i_s}\}$, i.e. $V_{i_1,...,i_s}$ is the variance of $f_{i_1,...,i_s}(\Theta_{i_1,...,i_s})$. This implies that

$$\begin{aligned} V_i &= \mathrm{Var}(f_i(\Theta_i)) = \mathrm{Var}(\mathbb{E}(f(\Theta)|\Theta_i)) \\ V_{i,j} &= \mathrm{Var}(f_{i,j}(\Theta_i, \Theta_j)) = \\ &= \mathrm{Var}(\mathbb{E}(f(\Theta)|\Theta_i, \Theta_j)) - \mathrm{Var}(\mathbb{E}(f(\Theta)|\Theta_i)) - \mathrm{Var}(\mathbb{E}(f(\Theta)|\Theta_j)) \\ V_{i,j,k} &= \mathrm{Var}(f_{i,j,k}(\Theta_i, \Theta_j, \Theta_k)) = \\ &= \mathrm{Var}(\mathbb{E}(f(\Theta)|\Theta_i, \Theta_j, \Theta_k)) - V_{i,j} - V_{i,k} - V_{j,k} \\ &\quad - \mathrm{Var}(\mathbb{E}(f(\Theta)|\Theta_i)) - \mathrm{Var}(\mathbb{E}(f(\Theta)|\Theta_j)) - \mathrm{Var}(\mathbb{E}(f(\Theta)|\Theta_k)) \end{aligned} \tag{3.21}$$

and so on. Here we can interpret the value $V_i$ as the first order effect of $\Theta_i$ on the model output, $V_{i,j}$ as the second-order effect on the model output, i.e. the joint interaction effect of two independent parameter components $\Theta_1$ and $\Theta_j$ and analogous interpretations for the higher orders. A model is called to be *additive*, if it has no other coefficients than $V_i$ of the form (3.21) that are unequal to zero.

Squaring and integrating both sides of equation (3.18) using condition (3.20) and plugging in the just defined definition of variances yields

$$V = \sum_{i=1} V_i + \sum_{i<j} V_{i,j} + \sum_{i<j<k} V_{i,j,k} + ... + V_{1,2,...,m} \tag{3.22}$$

which i.e. forms a decomposition of the unconditional variance $V = \mathrm{Var}(f(\Theta))$, which is denoted as the measure of uncertainty of the model output. The decomposition is often denoted as functional ANOVA [ES81].
The *variance-based sensitivity indices* as introduced by Sobol [Sob93], also called *Sobol sensitivity indices* for a certain parameter subset $\{\Theta_{i_1}, ..., \Theta_{i_s}\}$ are special cases of Pearson's correlation ratio [Pea05] and defined as

$$S_{i_1,...,i_s} = \frac{V_{i_1,...,i_s}}{V}. \tag{3.23}$$

The amount of different indices depends on the number of parameters in the model and is equal to $2^m - 1$. By this definition, one can interpret the value

$$S_i = \frac{V_i}{V} = \frac{\mathrm{Var}(\mathbb{E}(f(\Theta)|\Theta_i))}{\mathrm{Var}(f(\Theta))} \tag{3.24}$$

as the normalized first-order involvement of parameter $\Theta_i$ to the model output variance, the value

$$S_{i,j} = \frac{V_{i,j}}{V} = \frac{\mathrm{Var}(\mathbb{E}(f(\Theta)|\Theta_i, \Theta_j)) - V_i - V_j}{\mathrm{Var}(f(\Theta))} \tag{3.25}$$

as the normalized second-order involvement of the interaction of the parameters $\Theta_i$ and $\Theta_j$, and so on until order $m$ [Sal08].

Dividing equation (3.22) by $V$ yields that all sensitivity indices have to sum up to 1,and that the indices give a percentual measure of their contribution of certain parameters and parameter combinations to the model output variance under perturbations:

$$1 = \sum_{i=1} S_i + \sum_{i<j} S_{i,j} + \sum_{i<j<k} S_{i,j,k} + ... + S_{1,2,...,m} \geq 0. \tag{3.26}$$

The total-order sensitivity indices $S_{T_i}$ as introduced by [HS96] is defined as the sum over all these interaction indices for a certain parameter $\Theta_i$, using $\#i$ as notation indicating all indices associated to $\Theta_i$, therefore

$$S_{T_i} = S_i + \sum_{j \neq i} S_{i,j} + \sum_{j \neq i, k \neq i, j<k} S_{i,j,k} + ... = \sum_{l \in \#i} S_l \geq 0. \tag{3.27}$$

The value of $S_{T_i}$ quantifies the overall effect of one perturbed parameter vector component $\Theta_i$ on the model output of interest. By (3.26) and (3.27) we receive the two inequalities

$$1 \geq S_{T_i} \geq 0. \tag{3.28}$$

If $S_{T_i} = 0$ for a parameter $\Theta_i$, the respective parameter is said to be noninfluential on the model output. For values of $S_{T_i}$ relatively close to zero, the parameter $\Theta_{i_1}$ can be fixed to any value in its predefined range without influencing the model output variance. The approximation error resulting of this simplification is relative to the value of $S_{T_i}$ [Sob+07].

The following implementation of this method follows a Monte Carlo based numerical approach extended from [Sob93; HS96; Sal02] and presented in [Sal08] (Note that the formulae of equations (3.34) and (3.35) were wrong in the printed book version!).
Assume to have a model of $k$ factors, i.e., of $k$ different input parameters. Now generate a matrix of dimension $(N, 2k)$ that consists of random numbers. Split up this matrix into two sub-matrices $A$ and $B$ of dimension $(N, k)$ each, such that

$$A = \begin{bmatrix} x_1^{(1)} & x_2^{(1)} & \cdots & x_i^{(1)} & \cdots & x_k^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \cdots & x_i^{(2)} & \cdots & x_k^{(2)} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_1^{(N-1)} & x_2^{(N-1)} & \cdots & x_i^{(N-1)} & \cdots & x_k^{(N-1)} \\ x_1^{(N)} & x_2^{(N)} & \cdots & x_i^{(N)} & \cdots & x_k^{(N)} \end{bmatrix} \tag{3.29}$$

$$B = \begin{bmatrix} x_{k+1}^{(1)} & x_{k+2}^{(1)} & \cdots & x_{k+i}^{(1)} & \cdots & x_{2k}^{(1)} \\ x_{k+1}^{(2)} & x_{k+2}^{(2)} & \cdots & x_{k+i}^{(2)} & \cdots & x_{2k}^{(2)} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{k+1}^{(N-1)} & x_{k+2}^{(N-1)} & \cdots & x_{k+i}^{(N-1)} & \cdots & x_{2k}^{(N-1)} \\ x_{k+1}^{(N)} & x_{k+2}^{(N)} & \cdots & x_{k+i}^{(N)} & \cdots & x_{2k}^{(N)} \end{bmatrix} \tag{3.30}$$

Now further define $k$ matrices $C_i$ that are a copies of matrix $B$ but interchanged their $i$'th column with matrix $A$:

$$C = \begin{bmatrix} x_{k+1}^{(1)} & x_{k+2}^{(1)} & \cdots & x_i^{(1)} & \cdots & x_{2k}^{(1)} \\ x_{k+1}^{(2)} & x_{k+2}^{(2)} & \cdots & x_i^{(2)} & \cdots & x_{2k}^{(2)} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{k+1}^{(N-1)} & x_{k+2}^{(N-1)} & \cdots & x_i^{(N-1)} & \cdots & x_{2k}^{(N-1)} \\ x_{k+1}^{(N)} & x_{k+2}^{(N)} & \cdots & x_i^{(N)} & \cdots & x_{2k}^{(N)} \end{bmatrix} \tag{3.31}$$

The model output for each of the input rows of the matrices $A$,$B$ and $C_i$ is now calculated, which yields $k+2$ vectors of dimension $N$ each:

$$y_A = f(A), \quad y_B = f(B), \quad y_{C_i} = f(C_i). \tag{3.32}$$

Now define

$$f_0^2 = \left(\frac{1}{N}\sum_{j=1}^{N} y_A^{(j)}\right)\left(\frac{1}{N}\sum_{j=1}^{N} y_B^{(j)}\right). \tag{3.33}$$

The first order sensitivity indices as introduced earlier can then be estimated with

$$S_i = \frac{V\left[E(Y|X_i)\right]}{V(Y)} = \frac{\frac{1}{N}\sum_{j=1}^{N} y_A^{(j)} y_{C_i}^{(j)} - \frac{1}{N^2}\sum_{j=1}^{N} y_A^{(j)} y_B^{(j)}}{\frac{1}{N}\sum_{j=1}^{N}\left(y_A^{(j)} y_B^{(j)}\right) - f_0^2}. \tag{3.34}$$

Analogously, the total-order sensitivity indices can be estimated as

$$S_{T_i} = 1 - \frac{V\left[E(Y|X_{\sim i})\right]}{V(Y)} = 1 - \frac{\frac{1}{N}\sum_{j=1}^{N} y_B^{(j)} y_{C_i}^{(j)} - f_0^2}{\frac{1}{N}\sum_{j=1}^{N}\left(y_A^{(j)} y_B^{(j)}\right) - f_0^2}. \tag{3.35}$$

## 3.4 Statistical Survival Analysis

The statistical tools used in evaluation of clinical benefit for model-derived parameter values are briefly presented in this section. We consider time to event analyses to evaluate whether model-generated computational parameters are indicators of clinical outcomes.

### 3.4.1 Survival estimation

In many medical studies the interpretation of time to event analyses plays a major role. These events can be any events expressed in a binary formulation. If for some patient the event did not happen until the time at which the observation ended, the respective patient is censored. This can either happen in a way that the event is simply unknown, the patient was lost from the observation or that a follow-up examination was impossible [AB98]. Either way, the information for observed times is not lost: since the patient did not encounter an event until the observation's end, this information can be integrated into the survival analysis.
Patients are usually not recruited at the same time, therefore the observation phase can be of different length for different patients. An important assumption for studies that compare survival data is that the prognosis for different patients is the same,

independent of the time of recruitment for the study.
For our purpose, we examine two different types of events: overall survival (OS) as the time from primary diagnosis to the patient's death (induced by the disease examined) and progression-free survival (PFS) as the time from primary diagnosis to disease progression. Despite some censoring for several patients of a data set, the survival probability (in terms of OS and PFS) can be estimated.
A classical method to estimate the survival probability is the well-established Kaplan-Meier-method [KM58; BA98]. This method considers observation time intervals with respect to events: the conditional probability of a patient to survive a certain time interval is calculated if the patient already survived until the start of the respective interval. The probability to survive until a certain time is simply the product of the previous conditional probabilities [ZLB07c]. Patients that had to be censored at a certain time are not considered in the calculation of the future conditional probabilities. An example for such a Kaplan-Meier-method to examine patient survival times of the second data set (submitted in [Ben+22]) depending on the histology of the primary tumor of NSCLC is shown in figure 3.1.

Since special attention is given on the comparison of survival of two different groups within the data set, it is important to determine statistically whether the mortality risk in two groups may be considered statistically different. For this, the Kaplan-Meier-method is applied for each of the two subgroups of the patient collective with respect to some covariate of interest, e.g., the median value of a parameter - thus the one group consists of all patients that have this covariate value larger than the median value and the other group consists of patients that have this covariate value smaller or equal than the median value of the respective parameter.
A standard and frequently applied method to perform this analysis is the log-rank test [BA04]. If one of the two groups has an advantage in survival, then in this group the events would occur later in time. If not, then the events would occur in random order over both groups independently. The observed and the expected events of the respective groups are used to determine the test statistic

$$\Lambda = \frac{(O_I - E_I)^2}{E_I} + \frac{(O_{II} - E_{II})^2}{E_{II}} \tag{3.36}$$

with $O_I$, $O_{II}$ the number of observed events in the first and second group and $E_I$, $E_{II}$ the number of expected events in the first and second group, respectively. This means that the log-rank test equally weights all events. For a sufficiently large number of events the test statistic is approximately $\chi^2$-distributed with one degree of freedom [ZLB07a]. This allows to determine the corresponding p-value. For a significance level of 5%, p-values<0.05 are considered statistically significant, i.e., this results in statistically

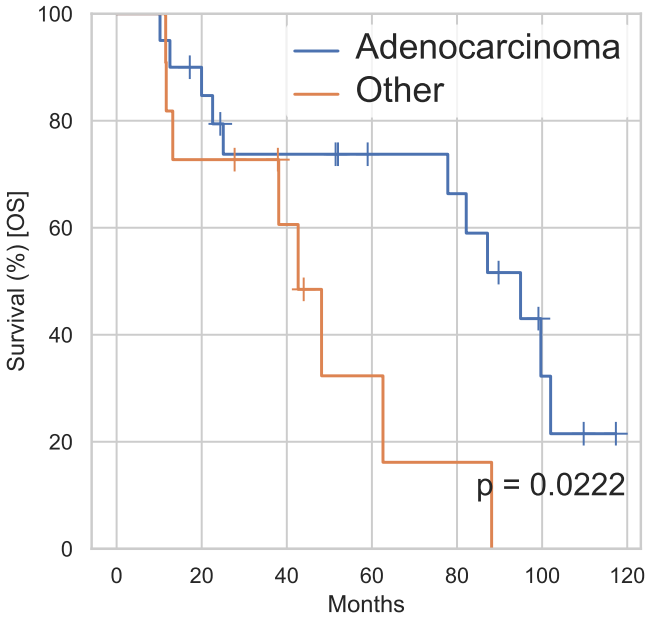


Figure 3.1: Kaplan-Meier curves with respect to overall survival (OS) for the second patient data set controlling for primary tumor histology, dividing the patient set in adenocarcinoma and other histologies. The examination with a log-rank test (explained in the text) yields statistically significant differences in survival curves when controlling for this covariate ($p = 0.0222$).

significant different survival properties for the two groups that were compared. This significance level will be used throughout this thesis.

### 3.4.2 Cox regression

The Cox regression model [Cox72], also known as the *Cox proportional hazards model*, is one of the most popular methods to analyse survival data. It aims at evaluating effects of potentially multiple different covariates on the censored survival data. The important assumption of this model states that the effects of different covariates on the survival are constant over time. This means, in particular, that the effects along a certain scale are additive [ZLB07b].
The proportional hazards model uses hazard functions $h(t)$ to be properly evaluated. A hazard function is the probability per time unit that an event happens to an individual patient in a short time interval $\Delta t$ conditioned that the event did not happen until the start of this interval. It can therefore be interpreted as the risk per time for the event to happen at time $t$ which can be formalized as

$$h(t) = \frac{E(t)}{N(t)\Delta t} \tag{3.37}$$

with $E(t)$ the number of individuals that the event happens to in the time interval $[t, t + \Delta t]$ and $N(t)$ the number of individuals that are alive at time $t$.
In terms of the Cox model, the hazard function is described in dependence to $n \in \mathbb{N}$ input covariate values $X_1, ..., X_n$ and a so-called baseline-hazard $h_0(t)$:

$$h(t) = h_0(t) \cdot e^{\beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n}. \tag{3.38}$$

The baseline-hazard explains the risk for an event, if all covariate values are equal to zero, i.e., all factors are absent. The coefficients $\beta_i$ are the regression coefficients of the input covariates. If an input covariate $X_i$ changes, the coefficient $\beta_i$ explains how the expected hazard changes with respect to the change of $X_i$.
The hazard function is often times also formulated in its logarithmic expression as

$$\ln h(t) = \ln h_0(t) + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n. \tag{3.39}$$

Let now, as assumed, the fraction of hazard functions of two different groups be constant over time, then the hazard ratio HR can be expressed as

$$HR = \frac{h_I(t)}{h_{II}(t)}. \tag{3.40}$$

This value is now a constant and, in particular, time-independent by assumption. If this value is smaller one, the first group shows a decreased hazard compared to the second

group and therefore an increased survival time. The opposite of this argumentation holds vice versa.
An interpretation for hazard ratios and their consequences on survival can be explained with an example as follows: Suppose two groups A and B have a HR of 0.6, i.e. the hazard for group A is lower than the one of group B. Let the (known) survival probability of group B for a certain time interval, e.g. one year, be denoted by $0 < x < 1$. Then the hazard for group A is reduced by $(1 - HR) = (1 - 0.6) \cdot 100\% = 40\%$. The expected survival probability for group A in the same time interval (here: one year) would then be $x^{0.6} > x$ and clearly increased compared to the survival of group B.

### 3.4.3 Predictive power of Cox model covariates

To estimate how well a statistical model can predict unseen data, cross-validation can be applied [RN03]. For some known data set, a certain fraction $k$ of the data is excluded. The remaining fraction $1 - k$ is used to form prediction hypotheses using a Cox proportional hazards model, as presented before. These hypotheses are then tested for their prediction accuracy on the other fraction $k$ of the data. For the cross-validation used later in this thesis, a value of $k = 3$ is used. This means that a Cox proportional hazards is formed on a sampled two thirds fraction of the data and tested for its predictive accuracy on the remaining third of the data.
This prediction accuracy of the Cox proportional hazards model can, for instance, be evaluated by determination of Harrel's concordance index, also called c-index [Har82; HLM96]. This index is defined as

$$c = \frac{n_c}{n_c + n_d} \tag{3.41}$$

with $n_c$ the number of concordant pairs of patients and $n_d$ the number of discordant pairs. To properly understand this formulation, some definitions have to be introduced, confer [SWZ16]:
Assume a data set of a patient collective with $n$ patients is given. For a patient $i$, $1 \leq i \leq n$, the event response is denoted by $T_i$. We have that $T_i = 1$ if the event occurred to patient $i$ during the time of observation and $T_i = 0$ otherwise. The Cox proportional hazards model assigns a risk score $\eta_i$ for each patient $i$. For a correct estimation of the statistical model, a higher risk score would yield a shorter time to event $t_i$.
Now examine the model's risk scores and times to event for each pair of patients $i$ and $j$. If $T_i + T_j = 2$, both patients had the respective event. Then the pair $(i, j)$ is considered as a concordant pair if $\eta_i > \eta_j$ and $t_i < t_j$. In contrast to that, the pair is considered discordant, if $\eta_i > \eta_j$ and $t_i \geq t_j$.
If $T_i + T_j = 1$, then only one event was observed. Without loss of generality let $i$ be

the patient to which the event occurred. Then we know that $t_i$ is the time to event for patient $i$ and that $t_j$ is the time until which patient $j$ was censored (without an observed event). If $t_j > t_i$, then the event first occurred to patient $i$. If this patient also had the higher risk score, i.e., if $\eta_i > \eta_j$, then the pair $(i, j)$ is considered a concordant pair. For the case where the patient $i$ had the lower risk score such that $\eta_i < \eta_j$, the pair is considered discordant. If $t_j < t_i$, then the patient $j$ is censored too early to determine whom the event occurred first to. In that case, the pair $(i, j)$ is not considered in the calculation.
Finally, if $T_i + T_j = 0$, then no event happened to both of the patients. The pair $(i, j)$ is therefore also not considered in the calculation.
These thoughts allow us to reformulate the expression for calculation of the concordance index as

$$c = \frac{\sum_{i \neq j} \mathbb{1}_{\eta_i < \eta_j} \mathbb{1}_{t_i > t_j} T_j}{\sum_{i \neq j} \mathbb{1}_{t_i > t_j} T_j}. \tag{3.42}$$

The values of the concordance index return a probability of correct prediction in ranking when comparing individuals. This means that it can only attain values between zero and one. A coin flip with a 50% chance to predict the correct ranking would yield a c-index of 0.5, therefore a c-index larger than 0.5 might already be considered as helpful in estimating time to events with respect to the risk scores generated by Cox proportional hazard's model.

# 4 Mathematical models

Mathematical models applied to tumor growth and metastases' development are relatively recent [AM04]. The first models formulated to study solid tumor growth with application to experimental data date back to the 1930's [May32], identifying a linear growth law for late stage solid Jensen's rat sarcoma after administration of radiation. This linear growth behavior is a consequence of the observation that with growing size of the tumor, the proliferating fraction reduces from the whole tumor to some outer rim of tissue with even decreasing thickness. The growth dynamic starts with some exponential growth, eventually decelerating and arriving at a linear growth dynamic for large tumor sizes.
In 1955, Thomlinson and Gray [TG55] defined a model considering oxygen diffusion and metabolism in lung cancer cells growing in rods, and noted that from the inside to the outer rim of a large tumor that usually shows necrotic core formations, some gradient in oxygen tension has to be present. Further, cells with lowest oxygenation levels are less damaged by radiation than those with higher oxygenation levels.
Since resources for tumor growth are limited in any living organism, modeling approaches quickly introduced an environmental carrying capacity to decelerate growth dynamics at larger sizes. The logistic growth equation introduced by Verhulst [Ver45] is an ordinary differential equation used in modeling approaches backed with experimental data with certain flaws [KT85]. However, the central awareness to incorporate growth-limiting dynamics and carrying capacities should be vital for further approaches with other growth dynamic equations, such as the generalized two-parameter equation and the von-Bertalanffy equation [Ber57; MB93; Sav79; VA82].
The important milestone concerning tumor growth dynamics with carrying capacities in living organisms, especially humans, dates back to 1825 [Gom25], though formed in another modeling context. The general idea of the Gompertz growth function in the sense of tumor growth postulates that the growing fraction of the tumor size, which is the fraction of proliferating tumor cells, is a function that decreases in time from an initially exponential growth. It was Laird to first fit the analytical solution of the Gompertz equation to volumetric tumor size measurements of animal experiment data [Lai64; Lai65] with remarkable good results. The biological explanation of the equations was given several years afterwards [Xu87]. Since its formulation and definition the Gompertz equation has been successfully applied to describe measured growth

dynamics of different experimental and clinical tumors for different histologies [SS72; Aka78; Dem80; Nor88; Ben+14; Ben+16; Ben+17; Bil+19; SKS21].
The uttermost reason of cancer-related deaths is well-proven: 90% of them do not occur by presence of the primary tumor but are direct consequences of the existence of metastases that greatly reduce treatment prognosis [NBM09; CW11; CN12; Ste16]. However, quantitative insights into metastasis formation and possible treatment approaches to counter certain formation mechanisms are relative sparse, possibly due to limited availability of data [ALM15].
Different approaches established are models based on deterministic ordinary differential equation formulations that focus on the competition between cancerous and healthy cells [SLK76], dormancy and cell dynamics of metastatic cancer cells [End+09; Bil+19] or on genetic and phenotypic diversity of primary tumor and metastatic tumor cells [Alm+14]. Other modeling approaches have shown that physical effects on and of the extracellular matrix in turn influence the tumor growth [PK12; Sch+22], drug delivery [Mic+11] and treatment success [Jai13]. Stochastic Markov chain models have been implemented to serve as computational frameworks and to estimate treatment response [CG86; HM10; New+12; Hae+12].
However, applicability to clinical settings on an individual patient basis remain meager. Based on a landmark publication [IKS00] a von Foerster partial derivative equation system is used to describe a density of metastases subject to growth and seeding dynamics including effects from a primary tumor. Several applications with adjustments have been showing astonishing possibilities to potentially fill the gap for clinical application [Ben+16; Ben+17; Bil+19; SKS21; Ben+22].

## 4.1 Tumor growth model

Introduce a time-dependent function $x(t)$ as the size of a tumor at time $t \geq 0$. The Gompertz growth equation combines, as previously stated, the initial exponential growth phase that is slowed down by exponential decay of the growth parameter to finally reach an asymptotic value for $t \to \infty$ [SKS21]. This asymptotic value is known as the tumor carrying capacity (the maximum size a tumor can reach) and is denoted by $K$. Then, by definition, $x(t) \xrightarrow{t\to\infty} K$. For some initial tumor size $x(0) = x_0 > 0$ the Gompertz model reads

$$\frac{dx}{dt} = re^{-at}x(t) \tag{4.1}$$

The parameters $r, a > 0$ can be interpreted as the initial exponential growth rate and the exponential decay factor, respectively.

Reformulating the growth rate as

$$r = a \ln \frac{K}{x_0}, \tag{4.2}$$

we arrive at an alternative way frequently used to express the Gompertz growth dynamics as

$$\frac{dx}{dt} = a \ln (K/x_0) e^{-at} x(t) \tag{4.3}$$

The Gompertz equation in its two formulations can be solved analytically to calculate the tumor size $x(t)$ at time $t \geq 0$ for $x_0 > 0$:

$$x(t) = x_0 e^{\left(\frac{r}{a}\left(1-e^{-at}\right)\right)} = x_0^{e^{-at}} K^{1-e^{-at}}. \tag{4.4}$$

The 'one renegade cell' theory states that the cancer's origin is one single mutated cell [Wei06; Bor+09], i.e. $x_0 = 1$. In that case, the Gompertz equation can be reformulated as

$$\begin{aligned} \frac{dx}{dt} &= a \ln (K/x(t)) x(t) \\ x(0) &= 1 \end{aligned} \tag{4.5}$$

Using again the initial tumor size $x(0) = 1$, the equations (4.1) and (4.5) can be solved analytically for

$$x(t) = e^{\left(\frac{r}{a}\left(1-e^{-at}\right)\right)} = K^{1-e^{-at}}. \tag{4.6}$$

The age of a tumor $T$, which is the time that a tumor of size $x_0 = 1$ needs to grow to a size $x(T)$, can be derived by solving equation (4.6) for time as a function depending on size $x$ [SKS21]. Let $T(x)$ be this function, then

$$T(x) = -\frac{1}{a} \ln \left(1 - \frac{a}{r} \ln (x)\right) = -\frac{1}{a} \ln \left(1 - \frac{\ln (x)}{\ln (K)}\right). \tag{4.7}$$

The clinical characterization of the tumor volume doubling time (TVDT) originally introduced for exponential tumor growth [Sch61; SS76] can also be applied to the Gompertz growth law. Unlike in the case of exponential growth, the TVDT in case of Gompertz growth is not a constant value but time-dependent. The TVDT is defined as the time a tumor of size $x > 1$ needs to grow to size $2x$ and can be estimated clinically [Ver+12]. Using equation (4.7) one receives

$$\begin{aligned} TVDT(x) = T(2x) - T(x) &= -\frac{1}{a} \ln \left(\frac{a \ln (2x)}{a \ln (x) - r} - \frac{r}{a \ln (x) - r}\right) \\ &= -\frac{1}{a} \ln \left(1 - \frac{\ln (2x)}{\ln (K)}\right) + \ln \left(1 - \frac{\ln (x)}{\ln (K)}\right). \end{aligned} \tag{4.8}$$

A main critic of the Gompertz growth dynamics is that for very small sizes the growth rate will diverge: $\frac{dx}{dt} \to \infty$ for $x \to 0$. But since the model formulation starts at an initial tumor size of one and the application to data later in this work focuses on clinically detectable tumors only (i.e., tumors much larger a size of one), this critic can be neglected in the context of this thesis. This is consistent with the observation that the Gompertz function well describes tumor growth especially in mid and later stages [Ged79; Sot+10; TD15; Rei+17].

## 4.2 Tumor growth model with therapeutic effects

Validating model feedback with clinical routine data proves to be difficult as there are rarely patients that have completely untreated cancerous diseases. Thus corresponding therapy effects should be included into the growth dynamics. The following ideas have been published in [SKS21] and are proven for their solvability.

### 4.2.1 Chemotherapy

Let the dynamics of chemotherapy be introduced as follows: $\mathbb{1}_{C_i}(t)$ denotes the characteristic function returning a logical value that indicates the application of a chemotherapeutic drug $C_i$ at time $t \geq 0$. This idea put into a mathematical formulation yields

$$\mathbb{1}_{C_i}(t) = \begin{cases} 0, \text{ if chemotherapy is not applied} \\ 1, \text{ if chemotherapy is applied} \end{cases} \quad \text{at } t \geq 0 \text{ using pharmaceutical } i. \tag{4.9}$$

In the following, let this function have finitely many jump discontinuities, i.e., the number of switches in therapy is finite. This obviously corresponds to clinical treatment protocols in reality.
Further assume that on days without chemotherapy, the concentration of chemotherapeutic drugs within the body and thus their effect is negligible. This is realistic for chemotherapeutic drugs, as the half-life of these pharmaceuticals is relatively small and lies in ranges of several hours. By this definition $\mathbb{1}_{C_i}(t)$ has a finite amount of jump discontinuities if both the application time of chemotherapeutic drugs and the amount of applications of the respective drug is measurable but finite. Let $\mu_i(t)$ denote the fractional effect of chemotherapy administration of drug $i$ on the tumor size. The refractory effect of drug administration can be expressed based on an idea of [Cla+09] as an ODE $\frac{d\mu_i}{dt}$ reducing the effects in a constant fraction $\mu_i^* \geq 0$ when the drug is applied, accounting for a fraction of the tumor to show resistance towards the drug,

using $\mu_i(0) = \mu_{i,0} \in [0,1]$:

$$\frac{d\mu_i}{dt} = \begin{cases} 0 & \text{, if } \mathbb{1}_{C_i}(t) = 0 \\ -\mu_i^* \mu_i(t) & \text{, if } \mathbb{1}_{C_i}(t) = 1 \end{cases} \tag{4.10}$$

This corresponds to an exponential decay of efficacy during application. We can now compute the effect of a pharmaceutical with initial therapy outcome $\mu_{i,0}$ on a later day within the therapy. The therapy effect will decrease recursively to $(1 - \mu_i^*)$ of its original value after each application day in active chemotherapy intervals. For times t not in the interval of active chemotherapy, the drug effect will stay the same since the drug is not applied. It is further assumed that if there was a recovery cycle after a treatment interval, the efficacy of the same pharmaceutical will not change until the next treatment interval.
The value of $\mu(t)$ can be calculated analytically, if $\mathbb{1}_{C_i}(t)$ and, in particular, the jump discontinuities are known:

**Lemma 4.2.1.** *For some index $i \in \mathbb{N}$ and time $t \in \mathbb{R}_0^+$ let $\mu_i(t)$ be a function $\mu_i(t) : \mathbb{R}_0^+ \to (0,1)$ that fulfills the equation (4.10), using a function $\mathbb{1}_{C_i}(t) : \mathbb{R}_0^+ \to \{0,1\}$ in the form of definition (4.9), a constant $\mu_i^* \in [0,1]$ and a constant $\mu_{i,0} \in [0,1]$. In particular, let $\mathbb{1}_{C_i}(t)$ have only a finite amount of jump discontinuities. Let $\int_0^t \mathbb{1}_{C_i}(\xi)d\xi$ denote the total time in days of applied chemotherapy in the context of equation (4.9). Then $\mu_i(t)$ can be computed as*

$$\mu_i(t) = \mu_{i,0}\, e^{-\mu_i^* \int_0^t \mathbb{1}_{C_i}(\xi)d\xi} \tag{4.11}$$

*Proof:* Distinguish the two possibilities of chemotherapy application. First assume $\mathbb{1}_{C_i}(t) = 0\ \forall t \in \mathbb{R}_0^+$, i.e. examine the case of no application. By equation (4.10) we receive that $\dot{\mu}_i(t) = 0$. Basic calculus yields $\mu_i(t) = \mu_{i,0}\ \forall\ t \in \mathbb{R}_0^+$ with $\mu_{i,0} \in \mathbb{R}$. This means that under no chemotherapeutic treatment the effectivity of drug $i$ stays at a constant value.

Assume now that $\mathbb{1}_{C_i}(t) = 1\ \forall\ t \in \mathbb{R}$, i.e., examine the case of active treatment. In this case equation (4.10) returns $\mu_i(t) = \mu_{i,0}e^{-\mu_i^* t}$ and therefore we get $\mu_i(t) = \mu_0 e^{-\mu_i^* t}$.
Let $\xi \in (0,1)$ denote the value of $\mu(t^*)$ at a time point $t^*$ at which a jump discontinuity of $\mathbb{1}_{C_i}(t)$ is happening, i.e. there is a switch in chemotherapy application. Then for $\varepsilon > 0$ one has for $\mathbb{1}_{C_i}(t^* + \varepsilon) = 1 \Rightarrow \mathbb{1}_{C_i}(t^* - \varepsilon) = 0 \wedge \mu(t^* + t) = \xi e^{-\mu_i^* t}$ and for $\mathbb{1}_{C_i}(t^* + \varepsilon) = 0 \Rightarrow \mathbb{1}_{C_i}(t^* - \varepsilon) = 1 \wedge \mu(t^* + t) = \xi$.

As $\mathbb{1}_{C_i}(t)$ is a step function by the form of (4.9), i.e. solely with nonzero returns if chemotherapy is applied, we know that the total number of days of chemotherapeutic application $\mathfrak{C}(t)$ in the time interval $[0,t]$ is given by the two expressions

$\mathfrak{C}(t) = \int_0^t \mathbb{1}_{C_i}(\xi)d\xi = \sum_{\xi=0}^{t} \mathbb{1}_{C_i}(\xi)$. Note that $\int_0^t \mathbb{1}_{C_i}(\xi)d\xi = t$ if $\mathbb{1}_{C_i}(\xi) \equiv 1$ and $\int_0^t \mathbb{1}_{C_i}(\xi)d\xi = 0$ if $\mathbb{1}_{C_i}(\xi) \equiv 0$.

Assume now for $\varepsilon > 0$ to observe different time points $0 < \xi < \tau < \vartheta$ with $|\mathbb{1}_{C_i}(\xi) - \mathbb{1}_{C_i}(\tau - \varepsilon)| = 0$, $|\mathbb{1}_{C_i}(\xi) - \mathbb{1}_{C_i}(\tau)| = 1$, $|\mathbb{1}_{C_i}(\tau) - \mathbb{1}_{C_i}(\vartheta - \varepsilon)| = 0$, $|\mathbb{1}_{C_i}(\tau) - \mathbb{1}_{C_i}(\vartheta)| = 1$ and $|\mathbb{1}_{C_i}(\xi) - \mathbb{1}_{C_i}(\vartheta)| = 0$, i.e. between the times $\xi$ and $\vartheta$ there are two jump discontinuities, one at time $\tau$, the other at time $\vartheta$. Assume further that $\mu_i(\xi) = \mu_{i,0}$. Distinguish again the two possible cases:
From $\mathbb{1}_{C_i}(\xi) = 0$ we have that $\mu_i(t) = \mu_i(\xi) = \mu_{i,0} = \mu_{i,0}e^{-\mu_i^* \int_0^t \mathbb{1}_{C_i}(s)ds}\ \forall t \in [\xi, \tau - \varepsilon)$, $\mathbb{1}_{C_i}(\tau) = 1$ and $\mu_i(t) = \mu_{i,0}e^{-\mu_i^* t} = \mu_{i,0}e^{-\mu_i^* \int_0^t \mathbb{1}_{C_i}(s)ds}\ \forall t \in [\tau, \vartheta - \varepsilon)$, as well as $\mathbb{1}_{C_i}(\vartheta) = 0$ and $\mu_i(t) = \mu_i(\vartheta) = \mu_{i,0}e^{-\mu_i^*(\vartheta - \tau)} = \mu_{i,0}e^{-\mu_i^* \int_0^t \mathbb{1}_{C_i}(s)ds}\ \forall t \geq \vartheta$.
From $\mathbb{1}_{C_i}(\xi) = 1$ follows that $\mu_i(t) = \mu_i(\xi)e^{-\mu_i^* t} = \mu_{i,0}e^{-\mu_i^* \int_0^t \mathbb{1}_{C_i}(s)ds}\ \forall t \in [\xi, \tau - \varepsilon)$, $\mathbb{1}_{C_i}(\tau) = 0$ and $\mu_i(t) = \mu_i(\xi)e^{-\mu_i^*(\tau - \xi)} = \mu_{i,0}e^{-\mu_i^* \int_0^t \mathbb{1}_{C_i}(s)ds}\ \forall t \in [\tau, \vartheta - \varepsilon)$, as well as $\mathbb{1}_{C_i}(\vartheta) = 1$ and $\mu_i(t) = \mu_i(\xi)e^{-\mu_i^*(\tau - \xi + t)} = \mu_{i,0}e^{-\mu_i^* \int_0^t \mathbb{1}_{C_i}(s)ds}\ \forall t \geq \vartheta$.
Overall for arbitrary and multiple time points of application we receive the final form

$$\mu_i(t) = \mu_{i,0}e^{-\mu_i^* \mathfrak{C}(t)} = \mu_{i,0}e^{-\mu_i^* \int_0^t \mathbb{1}_{C_i}(\xi)d\xi}.$$

□

Taking these effects for drug application into account for the growth dynamics of a tumor one has

$$\frac{dx}{dt} = \begin{cases} a\ln(K/x(t))x(t) & \text{, if } \mathbb{1}_{C_i}(t) = 0 \\ -\mu_i(t)x(t) & \text{, if } \mathbb{1}_{C_i}(t) = 1. \end{cases} \tag{4.12}$$

The ordinary differential equation is solvable not only for inactive chemotherapy regimen (cf. equation(4.4)), but also for the active chemotherapy regimen.

**Lemma 4.2.2.** *The initial value problem $\frac{dx}{dt} = -\mu_i(t)x(t)$ with $x(0) = x_0 > 0$ is analytically solvable for a function $\mu_i(t)$ of the form (4.11) with $\mathbb{1}_{C_i}(t) = 1\ \forall t > 0$. The solution reads*

$$x(t) = x_0 e^{\int_0^t -\mu_{i,0}e^{-\mu_i^* \int_0^\tau \mathbb{1}_{C_i}(\xi)d\xi}d\tau} \text{ for } t > 0. \tag{4.13}$$

*Proof:* By (4.11) we have that $\mu_i(t) = \mu_{i,0}\ e^{-\mu_i^* \int_0^t \mathbb{1}_{C_i}(\xi)d\xi} = \mu_{i,0}\ e^{-\mu_i^* t}$. Variation of constants directly yields

$$x(t) = x_0 e^{\int_0^t -\mu_i(\tau)d\tau} = x_0 e^{\int_0^t -\mu_{i,0}\ e^{-\mu_i^* \int_0^\tau \mathbb{1}_{C_i}(\xi)d\xi}d\tau} = x_0 e^{\int_0^t -\mu_{i,0}\ e^{-\mu_i^* \tau}d\tau} \tag{4.14}$$

□

For known therapy application times $\mathbb{1}_{C_i}(t)$ and therefore known jump discontinuity points of $\mathbb{1}_{C_i}(t)$ the solution $x(t)$ of equation (4.12) can be solved explicitly by a step-by-step approach from every jump discontinuity up to the next one. The solution procedure in particular depends on the initial therapy state $\mathbb{1}_{C_i}(0)$.

**Theorem 4.2.3.** *For some index $i \in \mathbb{N}$ and time $t \in \mathbb{R}_0^+$ let $\mu_i(t)$ be a function $\mu_i(t) : \mathbb{R}_0^+ \to (0,1)$ fulfilling condition (4.10) using a function $\mathbb{1}_{C_i}(t) : \mathbb{R}_0^+ \to \{0,1\}$ of the form (4.9) and let $\mu_i^* \in [0,1]$ as well as $\mu_{i,0} \in [0,1]$ be two constants, such that Lemma 4.2.1 may be applied. Let $x(t)$ be a function $x(t) : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ solving equation (4.12) and define the initial time point $\psi_0 = 0$ with $x(\psi_0) = 1$. If $\mathbb{1}_{C_i}(t)$ is known for all $t$, i.e. the values $\mathbb{1}_{C_i}(t) \in \{0,1\}$ are known for all $t$ and $\{\psi_1, ..., \psi_m\}$ denotes the set of all jump discontinuity time points $\psi_k > 0$ of the function $\mathbb{1}_{C_i}(t)$ with $k \in \{1, ..., m\}$ and $m \in \mathbb{N}$, then $x(t)$ can be computed at every time point $t$ between two respective jump discontinuity time points, i.e. $t \in (\psi_k, \psi_{k+1}]$.*
*For $\mathbb{1}_{C_i}(\psi_0) = 0$ we have that*

$$x(t) = \begin{cases} x(\psi_k)^{e^{-at}} K^{e^{-at}}, \text{ if } \mathbb{1}_{C_i}(t) = 0 \ (k \text{ even}) \\ x(\psi_k) e^{\int_0^t -\mu_{i,0} e^{-\mu_i^* \int_0^\tau \mathbb{1}_{C_i}(\xi) d\xi} d\tau}, \text{ if } \mathbb{1}_{C_i}(t) = 1 \ (k \text{ odd}) \end{cases} \tag{4.15}$$

*while for $\mathbb{1}_{C_i}(\psi_0) = 1$ we have that*

$$x(t) = \begin{cases} x(\psi_k)^{e^{-at}} K^{e^{-at}}, \text{ if } \mathbb{1}_{C_i}(t) = 0 \ (k \text{ odd}) \\ x(\psi_k) e^{\int_0^t -\mu_{i,0} e^{-\mu_i^* \int_0^\tau \mathbb{1}_{C_i}(\xi) d\xi} d\tau}, \text{ if } \mathbb{1}_{C_i}(t) = 1 \ (k \text{ even}). \end{cases} \tag{4.16}$$

*Proof:* Distinguish the two possibilities of chemotherapy application at the initial time point $\psi_0 := 0$. First assume $\mathbb{1}_{C_i}(\psi_0) = 0$, i.e. examine the case of no application at the initial time point. Then, by definition, the jump discontinuities $\psi_1$ and $\psi_2$ are the time points for which $\mathbb{1}_{C_i}(t) = 0 \ \forall t \in [\psi_0, \psi_1)$ and $\mathbb{1}_{C_i}(t) = 1 \ \forall t \in [\psi_1, \psi_2)$. Further for the next jump discontinuity $\psi_3$ we have again that $\mathbb{1}_{C_i}(t) = 0 \ \forall t \in [\psi_2, \psi_3)$. For the solution of equation (4.12) this yields that

$$x(t) = x(\psi_0)^{e^{-at}} K^{e^{-at}} \text{ for } t \in [\psi_0, \psi_1],$$

$$x(t) = x(\psi_1) e^{\int_0^t -\mu_{i,0} e^{-\mu_i^* \int_0^\tau \mathbb{1}_{C_i}(\xi) d\xi} d\tau} \text{ for } t \in (\psi_1, \psi_2] \text{ and}$$

$$x(t) = x(\psi_2)^{e^{-at}} K^{e^{-at}} \text{ for } t \in (\psi_2, \psi_3]$$

by equation (4.4), Lemma 4.2.1 and Lemma 4.2.2.
By applying this scheme iteratively we have that $\mathbb{1}_{C_i}(t) = 1 \ \forall t \in [\psi_{2k-1}, \psi_{2k})$ and $\mathbb{1}_{C_i}(t) = 0 \ \forall t \in [\psi_{2k}, \psi_{2k+1})$ for $k \in \{1, ..., m\}$ and $m \in \mathbb{N}$. Therefore the overall

solution of equation (4.12) in the case of non-initial chemotherapy and all time points $t \in (\psi_k, \psi_{k+1}]$ for some $k \in \mathbb{N}_0$ is given by

$$x(t) = \begin{cases} x(\psi_k)^{e^{-at}} K^{e^{-at}}, \text{ if } \mathbb{1}_{C_i}(t) = 0 \ (k \text{ even}) \\ x(\psi_k) e^{\int_0^t -\mu_{i,0} e^{-\mu_i^* \int_0^\tau \mathbb{1}_{C_i}(\xi) d\xi} d\tau}, \text{ if } \mathbb{1}_{C_i}(t) = 1 \ (k \text{ odd}). \end{cases}$$

Analogously for assuming $\mathbb{1}_{C_i}(\psi_0) = 1$, i.e. the case of chemotherapy application at the initial time point, the solution of equation (4.12) for time points $t \in (\psi_k, \psi_{k+1}]$ for some $k \in \mathbb{N}_0$ is given by

$$x(t) = \begin{cases} x(\psi_k)^{e^{-at}} K^{e^{-at}}, \text{ if } \mathbb{1}_{C_i}(t) = 0 \ (k \text{ odd}) \\ x(\psi_k) e^{\int_0^t -\mu_{i,0} e^{-\mu_i^* \int_0^\tau \mathbb{1}_{C_i}(\xi) d\xi} d\tau}, \text{ if } \mathbb{1}_{C_i}(t) = 1 \ (k \text{ even}). \end{cases}$$

□

### 4.2.2 Immunotherapy

Since the examined data set features patients treated with anti-PD-1 or anti-PD-L1 antibodies as immuntherapeutic drugs and since these drugs share the property to have half-lives much larger than those of chemotherapeutic drugs, a modelling approach considering pharmacokinetic effects was chosen.

Let, analogously to the setting of chemotherapeutic drug application in equation (4.9), $\mathbb{1}_{I_i}(t)$ denote the characteristic function returning a logical value that indicates the application of an immunotherapeutic drug $I_i$ at time $t \geq 0$. This yields

$$\mathbb{1}_{I_i}(t) = \begin{cases} 0, \text{ if immunotherapy is not applied} \\ 1, \text{ if immunotherapy is applied} \end{cases} \quad \text{at } t \geq 0 \text{ using pharmaceutical } i. \tag{4.17}$$

Since immunotherapy is assumed to be applied over one day the function $\mathbb{1}_{I_i}(t)$ fulfills the property that a jump discontinuity from value 0 to 1 and a jump discontinuity from 1 back to 0 are always exactly one day apart. Further, let $c_i(t)$ denote the concentration of drug $i$ at time $t \geq 0$ in number of drug molecules per body volume. For reasons of simplicity it is assumed that the body volume is constant over time for any individual patient. At given times $t$ this concentration's dynamic depends on the state of the immunotherapy: When the drug $i$ is applied, a dosage $d_i$ measured in milligram per body volume per time increases the drug concentration in the body with respect to the molar mass $M_i$ of drug $i$ and the Avogadro constant $N_A$. The drug's clearance is

influenced by the drug's half-life $t^{1/2}$ and happens at any time point. Therefore, we can formulate a pharmacokinetic equation for the drug concentration as

$$\dot{c}_i(t) = \begin{cases} -\left(\frac{\ln(2)}{t_i^{1/2}}\right) c_i(t) & \text{, if } \mathbb{1}_{I_i}(t) = 0 \\ -\left(\frac{\ln(2)}{t_i^{1/2}}\right) c_i(t) + \frac{N_A}{M_i} d_i & \text{, if } \mathbb{1}_{I_i}(t) = 1 \end{cases} \tag{4.18}$$

with some initial drug concentration $c_i(0) = c_i^{(0)} \geq 0$. The immunotherapeutic drug is applied in cycles of a certain length $l$ measured in time. E.g., this means that after one application of a daily dose, the time to the next application will be $l$ time units measured in days. To simplify model analysis the quasi steady state assumption may be applied to determine the steady state drug concentration that the simulation will periodically meet and approach for smaller application lengths. Let this value be denoted as $c_i^{st}$. The applied mean drug dose over the whole application time $\tilde{d}_i$ can be computed as the application time related mean value:

$$\tilde{d}_i = \frac{d_i}{l}. \tag{4.19}$$

Then the steady state value $c_i^{st}$ can be computed as

$$c_i^{st} = \frac{N_A}{M_i} \tilde{d}_i \frac{t_i^{1/2}}{\ln(2)}. \tag{4.20}$$

The quantitative effects of immunotherapy are not yet clearly resolved [Agr+16; Fen+17; Yoo+18]. Thus, we assume the immunotherapeutic drugs to follow a Hill-Langmuir equation with first order Hill coefficient in its pharmacokinetic efficiency against cancer cells. The parameter $c_i^{50}$ denotes the drug concentration of drug $i$ necessary to show half of this efficiency. The parameter $\chi$ describes the per time number of cancer cells destroyed by the direct application of one single molecule. The number of tumor cells is still denoted by $x(t)$, the pure growth dynamic stays as a Gompertz equation in the type of equation (4.5). The overall growth dynamic of cancer cells under influence of immunotherapy with some initial tumor size $x(0) = x_0$, therefore can be formulated as

$$\frac{dx}{dt} = ax(t) \ln\left(\frac{K}{x(t)}\right) - \chi \frac{c_i(t)x(t)}{c_i^{50} + c_i(t)}. \tag{4.21}$$

The assumption that the drug's concentration is in its approximated steady state value, i.e. $c_i(t) = c_i^{st}\ \forall t \geq 0$, allows to analytically solve for the tumor size at time $t$:

$$x(t) = K e^{\frac{c_i^{st}\chi\left(-e^{-at}+1\right)}{a(c_i^{st}+c_i^{50})}} \left(\frac{x_0}{K}\right)^{e^{-at}} \tag{4.22}$$

With a baseline tumor size value of $x_0$ to start the immunotherapy as reference point this equation allows to estimate the time until which there is potential progress of the disease.
On the other hand, a step-wise solution for the tumor growth dynamics under therapy is possible if $\mathbb{1}_{I_i}(t)$ has finitely many jump discontinuities, analogous to the case for chemotherapy:

**Lemma 4.2.4.** *For some index $i \in \mathbb{N}$ and time $t \in \mathbb{R}_0^+$ let $c_i(t)$ be a function $c_i(t) : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ of the form (4.18) using a function $\mathbb{1}_{I_i}(t) : \mathbb{R}_0^+ \to \{0,1\}$ of the form (4.17) and let the initial condition be $c_i(0) = c_i^{(0)} \geq 0$. Further, let $t_i^{1/2}$, $N_A$, $M_i$, $d_i \geq 0$. If $\mathbb{1}_{I_i}(t)$ is known for all $t$, i.e. the values $\mathbb{1}_{I_i}(t) \in \{0,1\}$ are known for all $t$ and $\{\psi_1, ..., \psi_m\}$ denotes the set of all jump discontinuity time points $\psi_k > 0$ of the function $\mathbb{1}_{I_i}(t)$ with $k \in \{1, ..., m\}$ and $m \in \mathbb{N}$, then $c_i(t)$ can be solved analytically at every time point $t$ between two respective jump discontinuity time points, i.e. $t \in (\psi_k, \psi_{k+1}]$, depending on the initial state of therapy $\mathbb{1}_{I_i}(0)$: For $\mathbb{1}_{I_i}(0) = 0$, i.e. the case of no immunotherapy application at the initial time point, we receive*

$$c_i(t) = \begin{cases} c_i(\psi_k)2^{-t/t_i^{1/2}}, \text{ if } \mathbb{1}_{I_i}(t) = 0 \ (k \text{ even}) \\ c_i(\psi_k)2^{-t/t_i^{1/2}} + \frac{\frac{N_A}{M_i} d_i t_i^{1/2} \left(1 - 2^{-t/t_i^{1/2}}\right)}{\ln(2)}, \text{ if } \mathbb{1}_{I_i}(t) = 1 \ (k \text{ odd}). \end{cases} \tag{4.23}$$

*For $\mathbb{1}_{I_i}(0) = 1$, i.e. the case of immunotherapy application at the initial time point, we receive*

$$c_i(t) = \begin{cases} c_i(\psi_k)2^{-t/t_i^{1/2}}, \text{ if } \mathbb{1}_{I_i}(t) = 0 \ (k \text{ odd}) \\ c_i(\psi_k)2^{-t/t_i^{1/2}} + \frac{\frac{N_A}{M_i} d_i t_i^{1/2} \left(1 - 2^{-t/t_i^{1/2}}\right)}{\ln(2)}, \text{ if } \mathbb{1}_{I_i}(t) = 1 \ (k \text{ even}). \end{cases} \tag{4.24}$$

*Proof:* Distinguish two possibilities of immunotherapy application. Assume that $\mathbb{1}_{I_i}(t) = 0 \ \forall t \in \mathbb{R}$. Then $c_i(t) = -\left(\frac{\ln(2)}{t_i^{1/2}} c_i(t)\right)$ by equation (4.18) with $c_i(0) = c_i^{(0)} \geq 0$ can be solved for

$$c_i(t) = c_i^{(0)} 2^{-t/t_i^{1/2}}. \tag{4.25}$$

In particular, if there was no initial drug concentration, i.e., $c_i^{(0)} = 0$, the concentration over time will stay equal to zero.

Assume now that $\mathbb{1}_{I_i}(t) = 1 \ \forall t \in \mathbb{R}$, i.e., assuming constant immunotherapeutic drug inflow. Then equation (4.18) yields $\dot{c_i}(t) = -\left(\frac{\ln(2)}{t_i^{1/2}} c_i(t) + \frac{N_A}{M_i} d_i\right)$ and we therefore receive for the solution that

$$c_i(t) = c_i^{(0)} 2^{-t/t_i^{1/2}} + \frac{\frac{N_A}{M_i} d_i t_i^{1/2} \left(1 - 2^{-t/t_i^{1/2}}\right)}{\ln(2)}. \tag{4.26}$$

For general $\mathbb{1}_{I_i}(t)$ let $\xi \geq 0$ be the value of $c_i(t^*)$ at a time point $t^*$ where a jump discontinuity happens. For some $\varepsilon > 0$ we receive that from

$$\begin{aligned} \mathbb{1}_{I_i}(t^* + \varepsilon) = 1 \Rightarrow & \mathbb{1}_{I_i}(t^* - \varepsilon) = 0 \text{ (if } t^* > 0) \\ & \text{and } c_i(t^* + t) = \xi 2^{-t/t_i^{1/2}} + \frac{\frac{N_A}{M_i} d_i t_i^{1/2} \left(1 - 2^{-t/t_i^{1/2}}\right)}{\ln(2)}. \end{aligned} \tag{4.27}$$

by equation (4.25). On the other hand we get that (with $0 < \varepsilon < 1$ by equation (4.26)):

$$\begin{aligned} \mathbb{1}_{I_i}(t^* + \varepsilon) = 0 \Rightarrow & \mathbb{1}_{I_i}(t^* - \varepsilon) = 1 \text{ (if } t^* > 0) \\ & \text{and } c_i(t^* + t) = \xi 2^{-t/t_i^{1/2}}, \end{aligned} \tag{4.28}$$

i.e., there is a reverse switch in therapy $k$ earlier in time.

Now distinguish the two possibilities of immunotherapy application at the initial time point $\psi_0 := 0$. First assume $\mathbb{1}_{I_i}(\psi_0) = 0$, i.e. no initial application of immunotherapeutic drugs. By definition we have that the jump discontinuities $\psi_1$ and $\psi_2$ are the two points for which the properties $\mathbb{1}_{I_i}(t) = 0 \ \forall t \in [\psi_0, \psi_1)$ and $\mathbb{1}_{I_i}(t) = 1 \ \forall t \in [\psi_1, \psi_2)$ hold. For the then following jump discontinuity $\psi_3$ we have again $\mathbb{1}_{I_i}(t) = 0 \ \forall t \in [\psi_2, \psi_3)$. With equations (4.27) and (4.28) this yields that

$$\begin{aligned} & c_i(t) = c_i(\psi_0) 2^{-t/t_i^{1/2}} \text{ for } t \in [\psi_0, \psi_1], \\ & c_i(t) = c_i(\psi_1) 2^{-t/t_i^{1/2}} + \frac{\frac{N_A}{M_i} d_i t_i^{1/2} \left(1 - 2^{-t/t_i^{1/2}}\right)}{\ln(2)} \text{ for } t \in (\psi_1, \psi_2] \text{ and} \\ & c_i(t) = c_i(\psi_2) 2^{-t/t_i^{1/2}} \text{ for } t \in (\psi_2, \psi_3] \end{aligned} \tag{4.29}$$

By applying this scheme iteratively we have that $\mathbb{1}_{I_i}(t) = 1 \ \forall t \in [\psi_{2k-1}, \psi_{2k})$ and $\mathbb{1}_{I_i}(t) = 0 \ \forall t \in [\psi_{2k}, \psi_{2k+1})$ for $k \in \{1, ..., m\}$ and $m \in \mathbb{N}$. Therefore the overall solution of equation (4.18) in the case of non-initial immunotherapy and all time points $t \in (\psi_k, \psi_{k+1}]$ for some $k \in \mathbb{N}_0$ is given by

$$c_i(t) = \begin{cases} c_i(\psi_k) 2^{-t/t_i^{1/2}}, \text{ if } \mathbb{1}_{I_i}(t) = 0 \ (k \text{ even}) \\ c_i(\psi_k) 2^{-t/t_i^{1/2}} + \frac{\frac{N_A}{M_i} d_i t_i^{1/2} \left(1 - 2^{-t/t_i^{1/2}}\right)}{\ln(2)}, \text{ if } \mathbb{1}_{I_i}(t) = 1 \ (k \text{ odd}). \end{cases}$$

Analogously for assuming $\mathbb{1}_{I_i}(\psi_0) = 1$, i.e. the case of immunotherapy application at the initial time point, the solution of equation (4.18) for time points $t \in (\psi_k, \psi_{k+1}]$ for some $k \in \mathbb{N}_0$ is given by

$$c_i(t) = \begin{cases} c_i(\psi_k) 2^{-t/t_i^{1/2}}, \text{ if } \mathbb{1}_{I_i}(t) = 0 \ (k \text{ odd}) \\ c_i(\psi_k) 2^{-t/t_i^{1/2}} + \frac{\frac{N_A}{M_i} d_i t_i^{1/2} \left(1 - 2^{-t/t_i^{1/2}}\right)}{\ln(2)}, \text{ if } \mathbb{1}_{I_i}(t) = 1 \ (k \text{ even}). \end{cases}$$

□

## 4.3 Metastatic development

The previously mentioned landmark publication by Iwata et al. [IKS00] introduced a novel interpretation of solutions of the von Foerster transport equation interpreted as density distribution in the context of metastatic cancerous diseases that will be summarized in the following. Adjustments in terms of modeling to this approach are presented thereafter.

### 4.3.1 Growth and size distribution of multiple metastatic tumors

Let, again, $x(t)$ denote the tumor size of a primary tumor at time $t \geq 0$ of the Gompertz growth form of equation (4.5). The metastases are as well assumed to grow in size according to the Gompertz growth dynamics. The growth for a metastasis of size $s(t)$ at time $t \geq 0$ is given by the solution of the initial value problem

$$\frac{ds}{dt} := g(s) := a_m \ln\left(K/s(t)\right)s(t) \text{ with } s(0) = s_0 = 1. \tag{4.30}$$

Analogously to the growth equation of the primary tumor, the tumor carrying capacity is chosen as value $K$. Thus, again, $s(t) \xrightarrow{t\to\infty} K$. The metastatic growth rate $a_m > 0$ is chosen equal to the growth rate of the primary tumor's growth rate originally in [IKS00], i.e., $a_m = a$. This is, however, not necessary in general.
Let $\varrho(s,t)$ be defined as the colony size distribution of metastatic tumors with cell number $s \geq 1$ at time $t \geq 0$. The number of metastatic tumors with size between $s$ and $s + \Delta s$ at some time $t$ can be expressed as $\varrho(s,t)\Delta x$. Since the approach follows the 'one renegade cell' theory again, the initial density distribution reads $\varrho(s,0) = 0\ \forall s$. This means that no metastatic cells but one primary tumor cell are present at time $t = 0$.
The growing primary tumor now starts seeding new single metastatic cells at the so-called 'colonization rate' $\beta(x)$. Any of these cells in turn grows at rate $g(s)$ and in turn forms new single metastatic cells at rate $\beta(s)$. This means that the formation of new metastatic single cells reads $\beta(x(t))$ from the primary tumor and $\int_1^\infty \beta(s)\varrho(s,t)ds$ from the metastatic colony size distribution.
The growth and seeding dynamics of the colony size distribution can therefore be postulated as

$$\frac{\partial \varrho(s,t)}{\partial t} + \frac{\partial g(s)\varrho(s,t)}{\partial s} = 0 \tag{4.31}$$

with initial and boundary conditions

$$\varrho(s,0) = 0, \tag{4.32}$$

$$g(1)\varrho(1,t) = \int_1^\infty \beta(s)\varrho(s,t)ds + \beta(x(t)). \tag{4.33}$$

For the colonization rate $\beta(x)$, [IKS00] suggest the form of

$$\beta(x) = mx^\alpha, \tag{4.34}$$

using some colonization coefficient $m > 0$ and the fractal dimension $\alpha > 0$ of blood vessels infiltrating the tumor. The colonization rate summarizes the necessary steps of the metastatic cascade for a successful formation of a metastasis [GM06], where the parameter $m$ can be seen as the per day per cell probability for a tumor cell to overcome every single step of this metastatic cascade [Bil+19]. By equation (4.34) it is assumed that the rate of metastasizing is proportional to the number of tumor cells in contact with blood vessels. Here, $\alpha$ expresses how the blood vessels geometrically distribute in or on a tumor: for $\alpha = 0$ the colonization rate is a constant pool (e.g. stem cells) while $\alpha = 1$ indicates that every tumor cell has equal probability to metastasize. Any values $0 < \alpha < 1$ can be interpreted as the geometric disposition of cells that potentially could metastasize [Bil+19]. The unit for $\beta(x)$ is the number of formed metastases of size one per unit time.
In its original formulation the model assumes, by construction, that all metastases are assumed to grow and seed new metastases at the same rates as the primary tumor, including the tumor carrying capacity [Bet+12].
From the model formulation we can easily estimate the total metastatic tumor burden, denoted by $N_0(t)$, which is the sum of the metastatic mass described by the colony size distribution $\varrho$ at time $t \geq 0$:

$$N_0(t) = \int_0^\infty \varrho(s,t)ds. \tag{4.35}$$

Further, the colony size distribution of visible metastatic tumors can be computed. For this, a certain visibility size threshold $s_{vis}$ has to be defined, such that all metastases of size $s > s_{vis}$ can be detected clinically. Then, the total detectable metastatic tumor burden is defined by $N_{vis}(t)$ and reads

$$N_{vis}(t) = \int_{s_{vis}}^\infty \varrho(s,t)ds. \tag{4.36}$$

If a growth rate is used that features growth limitation towards a carrying capacity, the upper integral boundary of the boundary condition (4.33) and the two equations (4.35) and (4.36) can be set as this carrying capacity $K$ instead of $\infty$.

### 4.3.2 Metastatic model with systemic therapy

The first data set features a patient that underwent systemic chemotherapy and systemic immunotherapy. To being able to analyze this data set with a modeling approach, the

model formulation has to consider these therapeutic effects. For systemic treatment it can be assumed that metastases are influenced at the same rate for the same drug $i$ [SKS21]. Distinguishing again for chemo- and immunotherapy, the growth rate $g(s)$ can now be chosen analogously to equations (4.12) and (4.21). The full model therefore consists of the following equations:

$$\frac{\partial \varrho(s,t)}{\partial t} + \frac{\partial g(s)\varrho(s,t)}{\partial s} = 0 \tag{4.37}$$

with initial and boundary conditions

$$\varrho(s,0) = 0, \tag{4.38}$$

$$g(1)\varrho(1,t) = \int_1^\infty \beta(s)\varrho(s,t)ds + \beta(x(t)). \tag{4.39}$$

$$\beta(x) = mx^\alpha, \tag{4.40}$$

and the adjusted definition of the growth rate $g(s)$, such that

$$\begin{aligned} g(s) := \frac{ds}{dt} &= \begin{cases} a\ln(K/s(t))s(t) - \chi\frac{c_j(t)s(t)}{c_j^{50}+c_j(t)} & \text{, if } \mathbb{1}_{C_i}(t) = 0 \\ -\mu_i(t)s(t) - \chi\frac{c_j(t)s(t)}{c_j^{50}+c_j(t)} & \text{, if } \mathbb{1}_{C_i}(t) = 1. \end{cases} \\ \text{with } \dot{c}_j(t) &= \begin{cases} -\left(\frac{\ln(2)}{t_j^{1/2}}\right)c_j(t) & \text{, if } \mathbb{1}_{I_j}(t) = 0 \\ -\left(\frac{\ln(2)}{t_j^{1/2}}\right)c_j(t) + \frac{N_A}{M_j}d_j & \text{, if } \mathbb{1}_{I_j}(t) = 1 \end{cases} \\ \text{and } \frac{d\mu_i}{dt} &= \begin{cases} 0 & \text{, if } \mathbb{1}_{C_i}(t) = 0 \\ -\mu_i^*\mu_i(t) & \text{, if } \mathbb{1}_{C_i}(t) = 1 \end{cases} \end{aligned} \tag{4.41}$$

having $c_j(0) = c_j^{(0)} \geq 0$, $\mu_i(0) = \mu_{i,0} \in [0,1]$, $\mu_i^* \geq 0$ and the already known characteristic functions indicating therapy status, i.e., $\mathbb{1}_{C_i}(t)$ and $\mathbb{1}_{I_j}(t)$ known from equations (4.9) and (4.17). Note that $g(s)$ applies completely analogous for the primary tumor's derivative $\frac{dx}{dt}$, since the primary tumor is as well assumed to be affected by the systemic therapy. Since the growth rate $g$ limits the growth to a maximum tumor size of $K$ again the integral's upper bound of equation (4.39) may be replaced by $K$.

### 4.3.3 Simplified metastatic model without secondary metastasis dissemination

The second data set includes patients that underwent surgery of their primary tumor, directed radiotherapy, whole-brain radiotherapy, systemic chemotherapy and systemic

immunotherapy. However, the aim with this data set was to consider only data up to the time point of first metastatic relapse (i.e., the first time that metastases have been diagnosed). Since the patients solely received systemic treatments after this diagnosis only surgery on the primary tumor had to be considered for the modeling approach [Ben+22]. The growth dynamics ordinary differential equation for a primary tumor of size $x(t)$ at time $t \geq 0$ starting with the initial condition $x(0) = 1$ yield using equation (4.2) that we have $e^{r/a} = K$.
The full dynamics considering surgery at some time $T_S > 0$ reads (analogously to equation (4.5) and using equation (4.6))

$$\begin{cases} \frac{dx}{dt} = (r - a\ln(x(t))x(t) & \text{, for } 0 < t \leq T_S \\ x(t) = 0 & \text{, for } t > T_S. \end{cases} \tag{4.42}$$

The primary tumor dynamics can be solved analytically for

$$x(t) = \begin{cases} e^{\frac{r}{a}(1-e^{-at})} & \text{, for } 0 \leq t \leq T_S \\ 0 & \text{, for } t > T_S. \end{cases} \tag{4.43}$$

Dynamics for the metastases are defined analogously without the possibility of surgery. The function $s(t)$ returns the size of a metastasis at time $t > 0$ with initial condition $s(0) = 1$. Then the growth equation reads

$$\frac{ds}{dt} = g_m(s) = (r_m - a_m \ln(s(t)))s(t) \tag{4.44}$$

and can be solved as

$$s(t) = e^{\frac{r}{a}(1-e^{-at})}. \tag{4.45}$$

The parameters $r_m$ and $a_m$ are again the initial specific growth rate and the exponential decay factor of the specific growth rate, but now for the metastases. Assuming the same tumor carrying capacity $K$ we have that $K = e^{r_m/a_m} = e^{r/a}$. Thus, the model assumes that metastases and primary tumor potentially show different growth dynamics.
The density of metastases is again formulated in the following transport equation:

$$\frac{\partial \varrho(s,t)}{\partial t} + \frac{\partial g_m(s)\varrho(s,t)}{\partial s} = 0 \tag{4.46}$$

with initial and boundary conditions

$$\varrho(s,0) = 0, \tag{4.47}$$

$$g(1)\varrho(1,t) = \beta(x(t)), \tag{4.48}$$

$$\beta(x) = mx^{\alpha}. \tag{4.49}$$

The new boundary condition (4.48) now describes that new metastases of cell size one introduced into the density distribution are exclusively formed directly by the primary tumor. Further, it is still assumed that at $t = 0$, the time at which the primary tumor consists of exactly one single cell, no metastases are present. The modeling approach focuses on primary metastasising only, since there is evidence that secondary metastasising is of minor clinical importance [Bet+12].

# 5 Simulation and sensitivity analysis

## 5.1 Tumor growth model

### 5.1.1 Implementation

Since equations (4.4) and (4.6) offer an analytical solution for the tumor growth dynamics, the full simulation is easily implemented for some given initial tumor size of one cell, i.e. $x(0) = 1$, and capacity $K = 10^{12}$ measured in cells [Kle09]. This capacity is equivalent to a volume of 1000ml or one kilogram of tumor mass [SMS95]. An example trajectory formulated in Matlab is presented in the next subsection.

### 5.1.2 Simulation

The simulation resembling to the three patients from data set one is shown in Figure 5.1 having the growth rate fixed to $a = 7 \cdot 10^{-3}$ with unit $[1/\text{day}]$. This value is close to the estimated values for the corresponding first data set, as presented later.

### 5.1.3 Sensitivity analysis

Since the initial condition $x(0) = 1$ and the tumor carrying capacity $K = 10^{12}$ are fixed, the sensitivity analysis for the untreated tumor growth ordinary differential equation is determined via the OAT approach, i.e., the derivative in direction of the free tumor growth rate parameter $a$. We have that the sensitivity $S_a(t)$ in direction of $a$ is determined with equation (4.6) via

$$S_a(t) = \frac{d(x(t))}{da} = te^{-at}K^{1-e^{-at}} \ln(K). \tag{5.1}$$

The elasticity is defined as $E_a(t) = \frac{a}{x(t)} S_a$, therefore

$$E_a(t) = ate^{-at} \ln(K). \tag{5.2}$$

Since all parameters are strictly positive we have that $S_a(t)$ and $E_a(t)$ are both strictly positive for $t > 0$. This means that an increase of the free parameter $a$ results in an

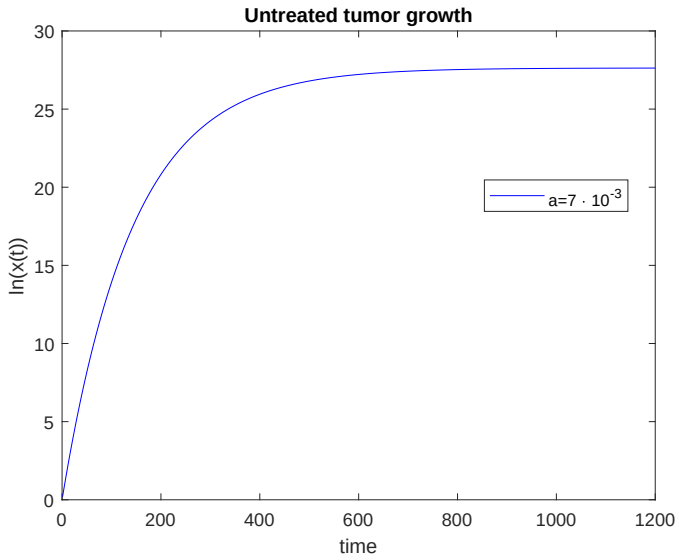


Figure 5.1: Simulation of the untreated primary tumor size $x(t)$ for growth rate $a = 7 \cdot 10^{-3}$ and tumor carrying capacity $K = 10^{12}$.

increase of the tumor size $x(t)$ at any time point $t$. However, we see that this effect vanishes for large $t$:

$$\begin{aligned} S_a(t) &\xrightarrow{t\to\infty} 0, \\ E_a(t) &\xrightarrow{t\to\infty} 0 \end{aligned} \tag{5.3}$$

i.e., for large $t$, the tumor size will still converge to the tumor carrying capacity $K$ independent of the choice of the growth rate parameter $a$. The plots of these functions $S_a(t)$ and $E_a(t)$ for the value $a = 7 \cdot 10^{-3}$ are shown in figure 5.2.

## 5.2 Tumor growth model with therapeutic effects

### 5.2.1 Implementation

For clinical cases, the functions (4.9) and (4.17) are known and are never equal to one at the same time. This allows to solve the sequence of ordinary differential equation initial value problems with the equations (4.15) and (4.16) in the case of chemotherapy. The concentration of immunotherapeutic drug $i$ at time $t$ measured in [molecules/volume], $c_i(t)$, can again be solved analytically by equations (4.23) and (4.24), depending on its initial value. From this, a concrete analytical formulation of equation (4.21) is available.

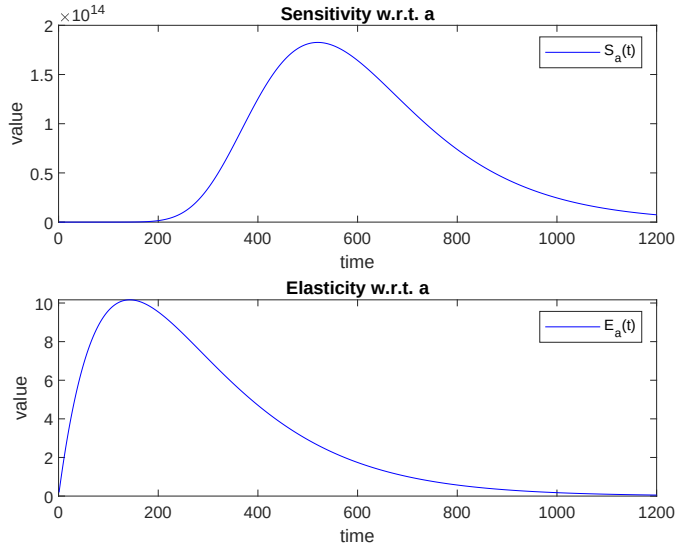


Figure 5.2: One-at-a-time analysis as sensitivity analysis on the primary tumor size without treatment for value $a = 7 \cdot 10^{-3}$.

This formulation can be estimated numerically using Matlab's pre-implemented solver routine `ode45` based on a Runge-Kutta (4,5) formula [DP80; SR97].

### 5.2.2 Simulation

To explore effects of a chemo- or immunotherapeutic treatment, the equations (4.13) and (4.21) were implemented on a primary tumor of initial size of $10^9$ cells. Two simulations for chemotherapy with different strength of the refractory effect towards the chemotherapeutic treatment were introduced ($\mu^* = 0.13$, blue solid line and $\mu^* = 0.85$, blue dashed line). To compare the influence of the refractory effect, another simulation was introduced with $\mu^* = 1$ (blue dash-dotted line), i.e. the chemotherapeutic efficacy would not decrease over application time. In all three cases, an application of two cycles of therapy of one week duration was considered, followed by a break of one week. The simulations are shown in Figure 5.3, it is clearly distinguishable how fast chemotherapeutic treatment becomes obsolete under a faster refractory effect.

To consider treatment in a close-to-reality setting, an example treatment protocol starting with a first-line chemotherapy three weeks after primary diagnosis applied for four weeks was initiated. After another treatment break of one week the immunotherapeutic treatment was implemented. A simulation result for this setting can be found in Figure 5.4, comparing the untreated tumor size (blue dashed line) to the tumor

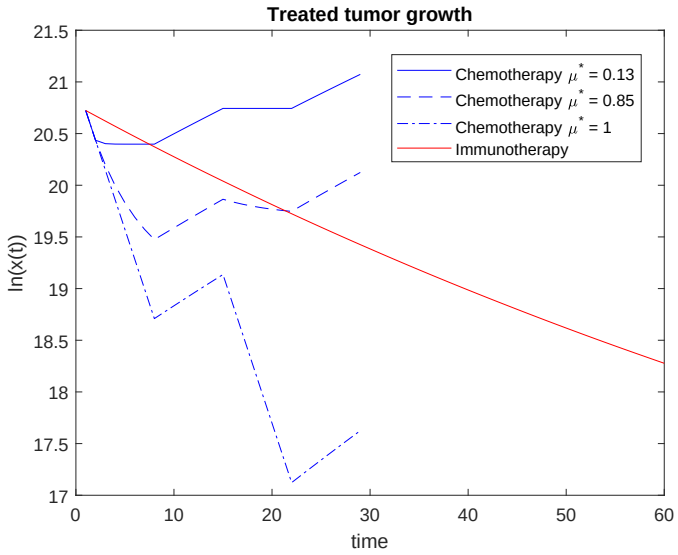


Figure 5.3: Simulation of the treated primary tumor growth under chemotherapy (blue lines) and immunotherapy (red line). The parameters values chosen for the simulations are for growth rate $a = 7 \cdot 10^{-3}$, tumor carrying capacity $K = 10^{12}$, initial chemotherapy efficacy $\mu_0 = 0.25$, immunotherapeutic efficacy $\chi = 0.1$ and drug-specific concentration for half-maximal response $c_i^{50} = 1.01 \cdot 10^{16}$ on a tumor of initial size $10^9$ cells. Refractory effects of chemotherapy were implemented with $\mu^* = 0.13$ (blue solid line) and with $\mu^* = 0.85$ (blue dashed line). Neglecting the refractory effect was introduced as well, having $\mu^* = 1$ (blue dash-dotted line). The drug implemented was Nivolumab, therefore the drug-specific parameters were chosen as dosage $d_i = 0.240$, application interval $l = 14$, molar mass $M_i = 1.46 \cdot 10^5$ and drug-specific half-life $t_i^{1/2} = 26.7$.

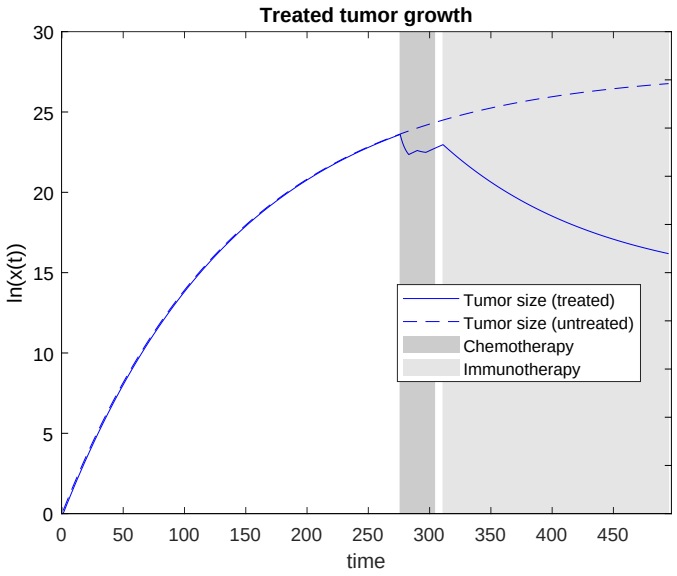


Figure 5.4: Simulation of the treated primary tumor growth for growth rate $a = 7 \cdot 10^{-3}$, tumor carrying capacity $K = 10^{12}$, initial chemotherapy efficacy $\mu_0 = 0.25$, refractory effect on chemotherapy $\mu^* = 0.85$, immunotherapeutic efficacy $\chi = 0.1$ and drug-specific concentration for half-maximal response $c_i^{50} = 1.01 \cdot 10^{16}$. The drug implemented was Nivolumab, therefore the drug-specific parameters were chosen as dosage $d_i = 0.480$, application interval $l = 28$, molar mass $M_i = 1.46 \cdot 10^5$ and drug-specific half-life $t_i^{1/2} = 26.7$. The tumor size under treatment application (blue solid line) is compared to the untreated tumor size (blue dashed line).

size under treatment (blue solid line). The different treatment intervals are shown as differently shaded time blocks. Parameter values chosen for this example simulation were $a = 7 \cdot 10^{-3}$, $K = 10^{12}$, $\mu_0 = 0.25$, $\mu^* = 0.85$, $\chi = 0.1$ and $c_i^{50} = 1.01 \cdot 10^{16}$ with unit [molecules/volume]. The drug implemented was Nivolumab, therefore the drug-specific parameters were chosen as $d_i = 0.480$ grams, $l = 28$ days, $M_i = 1.46 \cdot 10^5$ with unit [Da] and $t_i^{1/2} = 26.7$ days. Again, values are consistent with the findings for the first data set.

### 5.2.3 Sensitivity analysis

For the sensitivity analysis of the tumor growth ordinary differential equation under therapy, a clinically relevant tumor is examined. Assume that at primary diagnosis, i.e., at time $t = 0$, the primary tumor has a size of $x(0) = x_0 \gg 1$. Consider now two

different scenarios for the derivative-based OAT analysis:

**Scenario 1:** A chemotherapy is introduced right away. Let two parameters characterising this therapy in terms of equations (4.10) and (4.12) be defined as $\mu_0 > 0$ and $\mu^* > 0$. Evaluating the sensitivities on the tumor size at the end of an active chemotherapy of four cycles (28 days) can then be evaluated as the respective derivative of the tumor size $x(t) = x_0 e^{\frac{\mu_0 e^{-\mu^* t}}{\mu^*}}$ by Lemma 4.2.2 and read

$$\begin{aligned} S_{\mu_0} &= \frac{d(x(t))}{d\mu_0} = x_0 e^{\frac{e^{-\mu^* t}}{\mu^*}}, \\ S_{\mu^*} &= \frac{d(x(t))}{d\mu^*} = x_0 e^{\frac{\mu_0 e^{-\mu^* t}(\mu^* t+1)}{(\mu^*)^2}}, \\ S_{x_0} &= \frac{d(x(t))}{dx_0} = e^{\frac{\mu_0 e^{-\mu^* t}}{\mu^*}}. \end{aligned} \tag{5.4}$$

The corresponding elasticities read

$$\begin{aligned} E_{\mu_0} &= \frac{\mu_0}{x(t)} S_{\mu_0} = \frac{\mu_0 e^{-\frac{\mu_0}{\mu^*} e^{-\mu^* t} - \mu^* t}}{\mu^*}, \\ E_{\mu^*} &= \frac{\mu^*}{x(t)} S_{\mu^*} = -\frac{\mu_0(\mu^* t+1) e^{-\frac{\mu_0}{\mu^*} e^{-\mu^* t} - \mu^* t}}{\mu^*}, \\ E_{x_0} &= \frac{x_0}{x(t)} S_{x_0} = \frac{\mu_0 e^{-\frac{\mu_0}{\mu^*} e^{-\mu^* t} - \mu^* t}}{\mu^*}. \end{aligned} \tag{5.5}$$

All parameters are strictly positive, therefore we have that $S_{\mu_0}$, $E_{\mu_0}$, $S_{x_0}$ and $E_{x_0}$ are strictly positive for $t > 0$. This means that an increase of the free parameter $\mu_0$ or the initial tumor size prior to therapy $x_0$ results in an increase of the tumor size $x(t)$ at any time point $t$. For the chemotherapy efficacy change factor $\mu^*$ this is slightly different: as $S_{\mu^*}$ and $E_{\mu^*}$ are both strictly negative for all positive parameter values we have that the tumor size $x(t)$ decreases over time $t$. We can again observe that these effects vanish for large $t$:

$$\begin{aligned} E_{\mu_0} &\xrightarrow{t\to\infty} 0, \\ E_{\mu^*} &\xrightarrow{t\to\infty} 0, \\ E_{x_0} &\xrightarrow{t\to\infty} 0, \end{aligned} \tag{5.6}$$

i.e., for large $t$, the tumor size will vanish. This is however not really interesting for the application, since the chemotherapy can only be administered in a limited time interval

due to adverse drug effects. Still it is clear that the effects on tumor size will vanish more if chemotherapy is applied longer. Plots shown in figure 5.5 for values indicate very clearly at which time points the perturbed parameters affect the tumor size.

**Scenario 2:** The immunotherapy is introduced from the very beginning. Assume that the drug concentration's steady state is present in the patient's body and equation (4.22) can be applied. This yields that we can calculate the size of the primary tumor under active immunotherapeutic treatment as

$$x(t) = Ke^{\frac{c^{st}\chi\left(-e^{-at}+1\right)}{a(c^{st}+c^{50})}}\left(\frac{x_0}{K}\right)^{e^{-at}}. \tag{5.7}$$

Let the tumor carrying capacity $K = 10^{12}$ be fixed and let the initial tumor size be defined by $x(0) = x_0 > 1$. Further assume all parameters $a, c^{st}, c^{50}$ and $\chi$ to be strictly positive. Their sensitivities read

$$\begin{aligned}
S_a &= K\left(\frac{x_0}{K}\right)^{e^{-at}} e^{\frac{\chi c^{st}(1-e^{-at})}{a(c^{50}+c^{st})}}\left(\frac{\chi t c^{st} e^{-at}}{a(c^{50}+c^{st})} - \frac{\chi c^{st}(1-e^{-at})}{a^2(c^{50}+c^{st})}\right),\\
&\quad - Kt\ln\left(\frac{x_0}{K}\right)\left(\frac{x_0}{K}\right)^{e^{-at}} e^{\frac{\chi c^{st}(1-e^{-at})}{a(c^{50}+c^{st})}-at},\\
S_{c^{st}} &= K\left(\frac{x_0}{K}\right)^{e^{-at}} e^{\frac{\chi c^{st}(1-e^{-at})}{a(c^{50}+c^{st})}}\left(\frac{\chi(1-e^{-at})}{a(c^{50}+c^{st})} - \frac{\chi c^{st}(1-e^{-at})}{a(c^{50}+c^{st})^2}\right),\\
S_{c^{50}} &= -\frac{\chi K c^{st}(1-e^{-at})\left(\frac{x_0}{K}\right)^{e^{-at}} e^{\frac{\chi c^{st}(1-e^{-at})}{a(c^{50}+c^{st})}}}{a(c^{50}+c^{st})^2},\\
S_{\chi} &= \frac{K c^{st}(1-e^{-at})\left(\frac{x_0}{K}\right)^{e^{-at}} e^{\frac{\chi c^{st}(1-e^{-at})}{a(c^{50}+c^{st})}}}{a(c^{50}+c^{st})},\\
S_{x_0} &= \left(\frac{x_0}{K}\right)^{e^{-at}-1} e^{\frac{\chi c^{st}(1-e^{-at})}{a(c^{50}+c^{st})}-at},
\end{aligned} \tag{5.8}$$

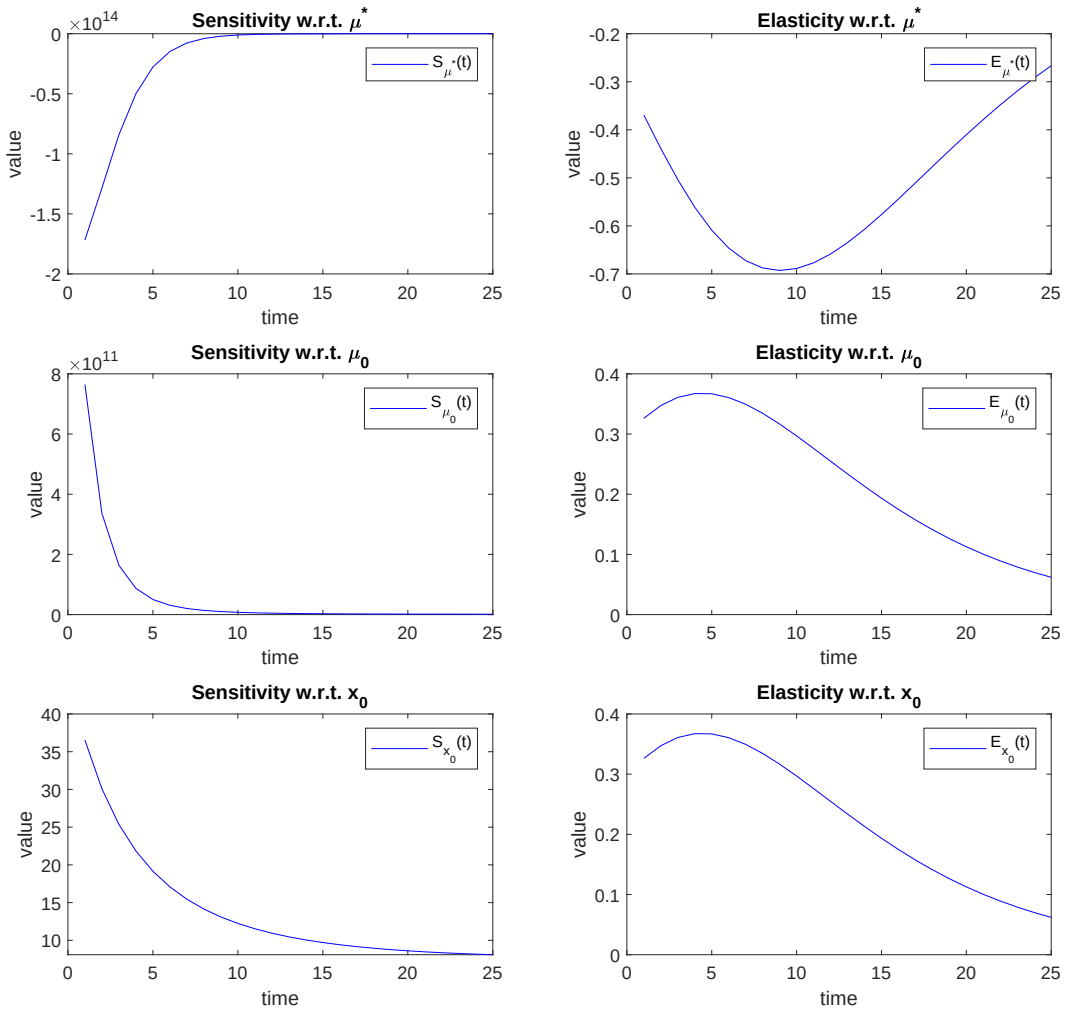


Figure 5.5: One-at-a-time analysis as sensitivity analysis on the primary tumor size under chemotherapeutic treatment.

and their corresponding elasticities are

$$
\begin{aligned}
E_a &= \frac{a}{x(t)} S_a = K \left(\frac{x_0}{K}\right)^{e^{-at}} e^{\frac{\chi c^{st}(1-e^{-at})}{a(c^{50}+c^{st})}} \left(\frac{\chi t c^{st} e^{-at}}{a(c^{50}+c^{st})} - \frac{\chi c^{st}(1-e^{-at})}{a^2(c^{50}+c^{st})}\right) \\
&\quad - Kt \ln\left(\frac{x_0}{K}\right) \left(\frac{x_0}{K}\right)^{e^{-at}} e^{\frac{\chi c^{st}(1-e^{-at})}{a(c^{50}+c^{st})} - at} \frac{a}{Ke^{\frac{c^{st}\chi\left(-e^{-at}+1\right)}{a(c^{st}+c^{50})}} \left(\frac{x_0}{K}\right)^{e^{-at}}}, \\
E_{c^{st}} &= \frac{c^{st}}{x(t)} S_{c^{st}} = c^{st} \left(\frac{\chi(1-e^{-at})}{a(c^{50}+c^{st})} - \frac{\chi c^{st}(1-e^{-at})}{a(c^{50}+c^{st})^2}\right), \\
E_{c^{50}} &= \frac{c^{50}}{x(t)} S_{c^{50}} = -\frac{c^{50}\chi c^{st}(1-e^{-at})}{a(c^{50}+c^{st})^2}, \\
E_{\chi} &= \frac{\chi}{x(t)} S_{\chi} = \frac{\chi c^{st}(1-e^{-at})}{a(c^{50}+c^{st})}, \\
E_{x_0} &= \frac{x_0}{x(t)} S_{x_0} = e^{-at}.
\end{aligned}
\tag{5.9}
$$

For the long time behavior we receive

$$
\begin{aligned}
S_a &\xrightarrow{t\to\infty} \frac{\chi K c^{st} e^{\frac{\chi c^{st}}{a(c^{50}+c^{st})}}}{a^2(c^{50}+c^{st})}, \\
S_{c^{st}} &\xrightarrow{t\to\infty} K\chi e^{\frac{\chi c^{st}}{a(c^{50}+c^{st})}} \left(\frac{1}{a(c^{50}+c^{st})} - \frac{c^{st}}{a(c^{50}+c^{st})^2}\right), \\
S_{c^{50}} &\xrightarrow{t\to\infty} -\frac{\chi K c^{st} e^{\frac{\chi c^{st}}{a(c^{50}+c^{st})}}}{a(c^{50}+c^{st})^2}, \\
S_{\chi} &\xrightarrow{t\to\infty} -\frac{K c^{st} e^{\frac{\chi c^{st}}{a(c^{50}+c^{st})}}}{a(c^{50}+c^{st})}, \\
S_{x_0} &\xrightarrow{t\to\infty} 0.
\end{aligned}
\tag{5.10}
$$

These investigations clearly indicate that all changes in parameters except the initial tumor size at the start of immunotherapy in the OAT setting have lasting effects over time. For increased tumor growth rate $a$ and increased drug concentration steady state $c^{st}$ holding all other parameters equal the tumor size increases in the long run. On the other hand, increasing values of the parameters $c^{50}$ and $\chi$ yield negative values of sensitivities, therefore decreasing tumor sizes in the long run. Only the sensitivity for $x_0$ vanishes over time, which means that the initial size of the tumor does not play a major role on the tumor size in the long run in the context of the OAT analysis. An example plot for these sensitivities and elasticities with values fitted from data is given in figure 5.6. From the equations however, it is difficult to identify the ranking of the

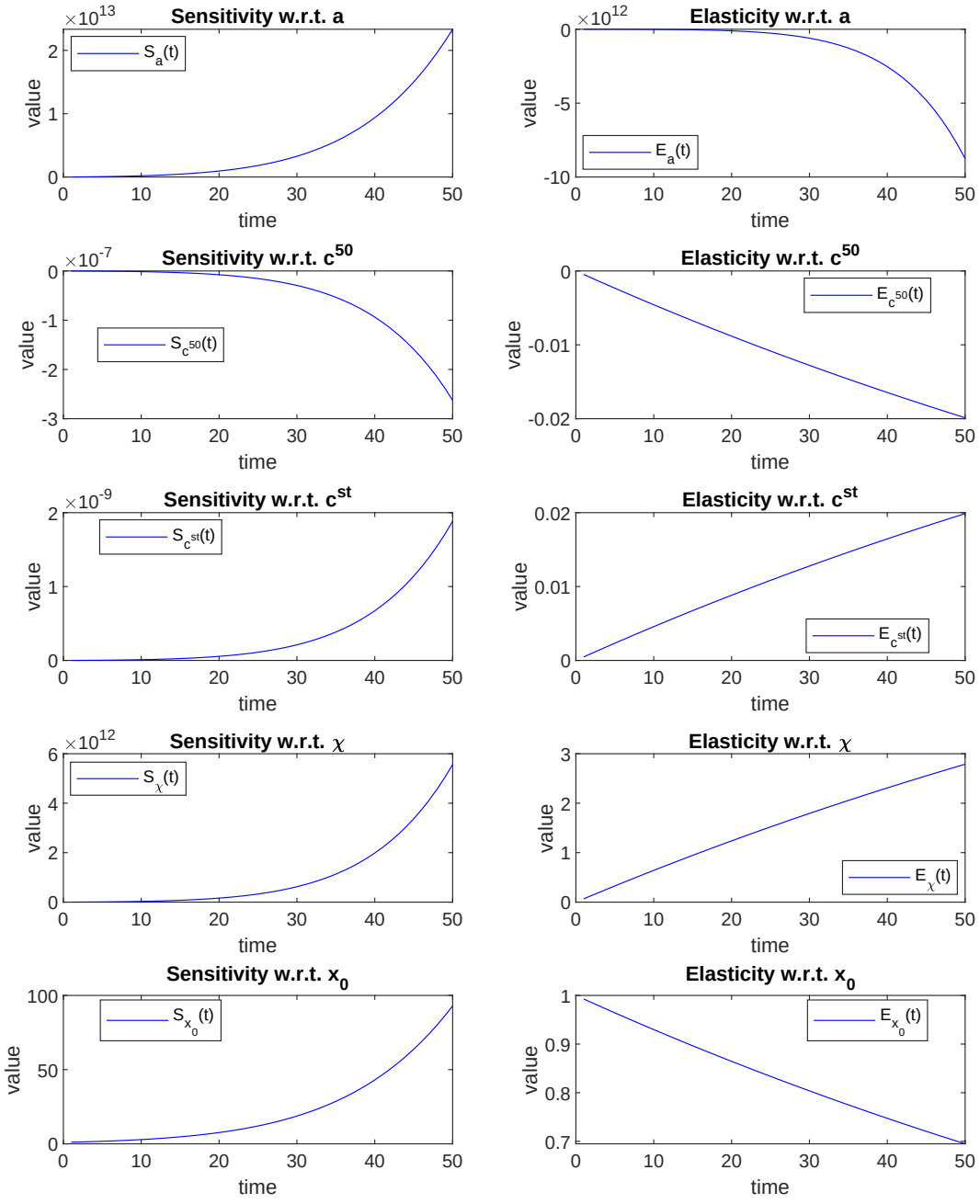


Figure 5.6: One-at-a-time analysis as sensitivity analysis on the primary tumor size under immunotherapeutic treatment.

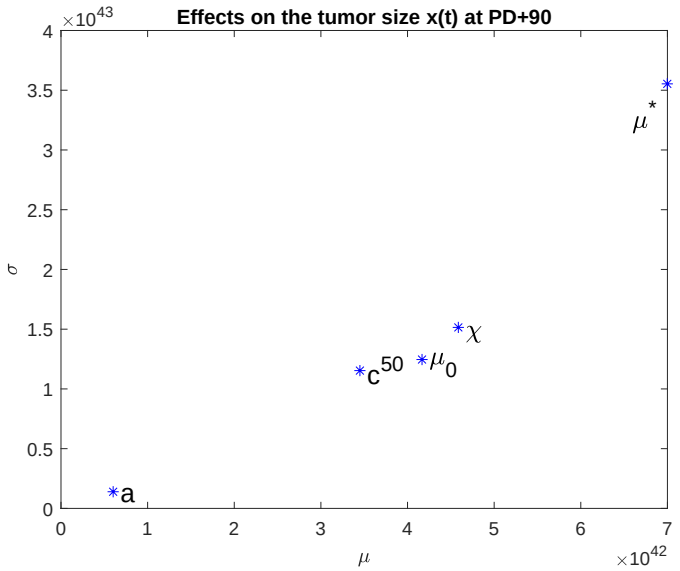


Figure 5.7: Elementary effects method as sensitivity analysis on the primary tumor size under treatment.

sensitivity of parameters.

Previously the concept of the elementary effects method was introduced to address this problem. To consider treatment in this framework, an example standard treatment was defined as a first-line chemotherapy three weeks after primary diagnosis that is applied for four weeks. The immunotherapeutic treatment follows after another week without treatment. This concept was implemented for biologically reasonable parameter intervals with 2000 randomly sampled trajectories checking the tumor size 90 days after primary diagnosis (24 days after initialising the immunotherapeutic treatment). Corresponding results can be found in figure 5.7 and 5.8. Clearly, the initial chemotherapy efficacy parameter $\mu^*$ has the highest influence on the primary tumor size in this protocol, followed by the immunotherapy drug-specific efficacy parameter $\chi$. To shed light on interaction effects, the Sobol' indices were implemented on the same treatment regimen and on the primary tumor size 90 days after primary diagnosis for a sample size of 50 000. The parameter intervals were chosen the same as for the elementary effects method. However, the Sobol' indices allow for sensitivity estimations over the whole parameter space, not just relatively to a certain reference parameter combination as the two previous approaches. The results are shown in the table below.

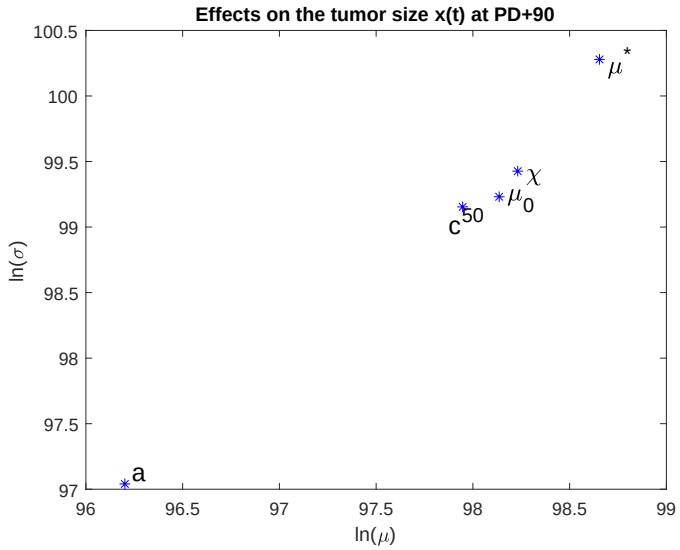


Figure 5.8: Elementary effects method as sensitivity analysis on the primary tumor size under treatment, log-scaled.

| Index | $c^{50}$ | $\chi$ | $a$ | $\mu^*$ | $\mu_0$ |
|---|---|---|---|---|---|
| $S_i$ | 0.0048 | 0.6801 | 0.0702 | 0.0053 | 0.0052 |
| $S_{tot}$ | 0.0076 | 0.9328 | 0.3091 | 0.0163 | 0.0105 |

These values clearly show that interaction effects, i.e., combined changes of multiple parameters at the same time, play a huge role in the outcome of the primary tumor size. The highest effects are measurable for parameters $\chi$, $a$ and $\mu^*$. Interestingly, the drug concentration needed for an efficacy of 50% and the initial chemotherapeutic efficacy $\mu_0$ are only of minor importance when sensitivity on the tumor size is considered.
Comparison of the elementary effects and Sobol' methods clearly shows their respective advantages and disadvantages. While the elementary effects method is based on trajectory sampling and focuses on equidistant changes in each direction for any parameter but only one parameter at a time, the Sobol' method explores the complete parameter space also considering the simultaneous perturbation of potentially multiple parameters. This results in a more concrete estimation for observable effects that might occur in reality and give insight in the changing importance of apparently less sensitive parameters for OAT perturbation towards huge influence of higher order sensitivity measures. An example for such a parameter is the tumor growth rate $a$ in this setting.

## 5.3 Metastatic development

### 5.3.1 Implementation

For the classic model without systemic treatment, the authors offer an analytical solution [IKS00]. However, the formulation does not allow for easy and explicit calculation of the metastatic density at any time point $t$. A discretisation to compute the partial differential equation's solution along their characteristic curves can be established to calculate the density distribution numerically. These characteristic curves of the system (4.31) - (4.34) correspond to exactly the growth rates $g(x,t)$. These can be determined analytically (see equation (4.4)).
Let $T_{max}$ be the maximum observation time of some patient, i.e. the time interval of interest for that patient is $[0, T_{max}]$ and let $k = 1$ be a constant time step discretisation on the time interval $[0, T_{max}]$ such that $t_j = jk$ for $j \in \{0, ..., n\}$ with $t_0 = 0$ and $t_n = T_{max}$. Choose the size discretisation $x_j = x(t_j)$ such that $x_0 = 1 < ... < x_n = K$. Like this, the points $(x_k, t_l)$ and $(x_{k+1}, t_{l+1})$ belong to the same characteristic for any $k, l \in [0, n-1]$. Then the classic model with secondary metastasising can be numerically evaluated with the following explicit scheme making use of Riemann sums and density discretisations $\varrho(x(i), t(j)) \approx u_{x_i,t_j}$ (see [BC01]):

$$\begin{aligned} u_{x_0,0} &= 1 \\ u_{x_i,0} &= 0 \qquad i = 1, ..., n \\ u_{x_0,t_j} &= \frac{1}{1 - mx_0^\alpha} \sum_{i=1}^{n-1} m x_i^\alpha u_{x_i,t_j} \qquad j = 1, ..., n \\ u_{x_i,t_j} &= u_{x_{i-1},t_{j-1}} \qquad i, j = 1, ..., n \end{aligned} \tag{5.11}$$

For the model neglecting secondary metastasising, the approach is very similar, but only considering freshly seeded metastases from the primary tumor, i.e.

$$\begin{aligned} u_{x_0,0} &= 1 \\ u_{x_i,0} &= 0 \qquad i = 1, ..., n \\ u_{x_0,t_j} &= \frac{1}{1 - mx_0^\alpha} m x_j^\alpha u_{x_j,t_j} \qquad j = 1, ..., n \\ u_{x_i,t_j} &= u_{x_{i-1},t_{j-1}} \qquad i, j = 1, ..., n \end{aligned} \tag{5.12}$$

### 5.3.2 Simulation

Having the two different model approaches with and without secondary metastasizing implemented, the corresponding parameters were chosen as the growth rate $a = 7 \cdot 10^{-3}$,

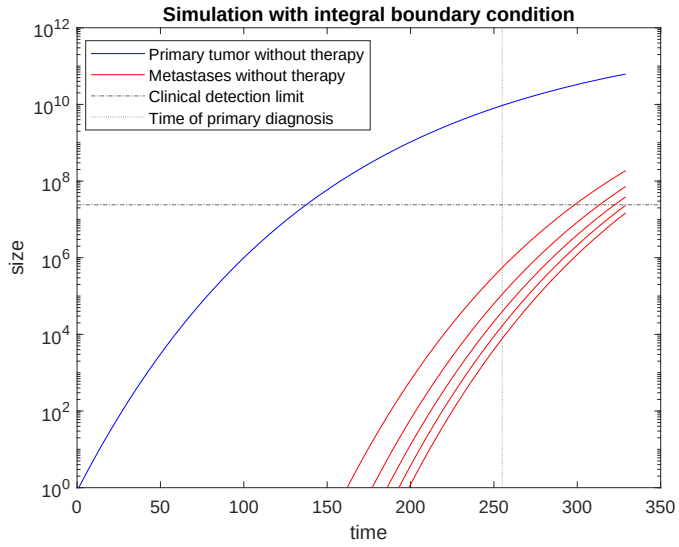


Figure 5.9: Simulation of the primary tumor (blue solid line) and the largest five metastases (red solid line) for the model considering no treatment and secondary metastasising, log-scaled. Critical size for a tumor to be clinically detectable is indicated with a dash-dotted line on the y-axis, the day of primary diagnosis is indicated with a dotted line on the x-axis. The chosen model parameters are $a = 7 \cdot 10^{-3}$, $K = 10^{12}$ and $m = 2 \cdot 10^{-7}$.

tumor carrying capacity $K = 10^{12}$ and the metastatic colonization rate $m = 2 \cdot 10^{-7}$. Simulations for the approach with secondary metastasizing are the full time course of the primary tumor (blue line) and the largest five metastases (red lines), Figure 5.9. The corresponding simulation for the model formulation considering primary metastatic seeding only is shown in Figure 5.10 with analogous colors and indications. The difference in these two approaches is negligible at this reference.
However, the simulation of the metastatic density at the day of primary diagnosis for both modeling approaches gives insight in the different model outcomes. Figures 5.11 and 5.12 show the numerical metastatic density distribution approximations at the day of primary diagnosis. The diamonds indicate integer values, thus full countable metastases. The largest diamond is the primary tumor. Clearly, the difference of these two approximations is at the density distribution's side of very small metastases: the model considering secondary metastasizing has a full additional metastasis of very small size only. The density distribution for larger metastases is approximately the very same, supporting the hypothesis that secondary metastases potentially only play a minor role in clinical treatment [Bet+12].

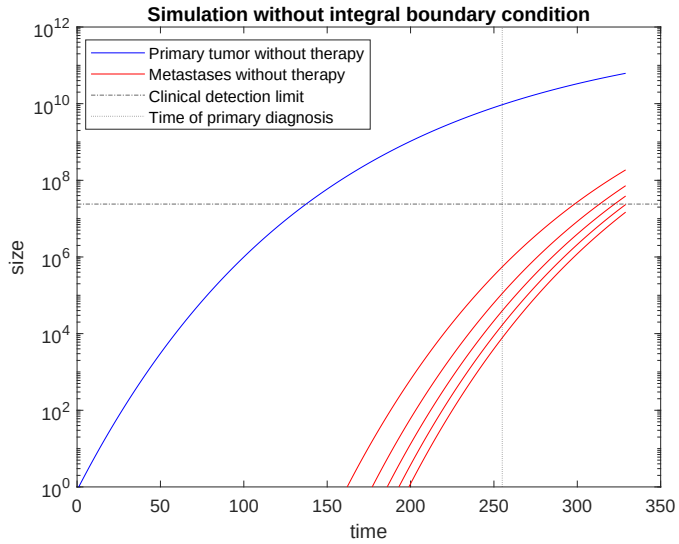


Figure 5.10: Simulation of the primary tumor (blue solid line) and the largest five metastases (red solid line) for the model considering no treatment and only primary metastasising, log-scaled. Critical size for a tumor to be clinically detectable is indicated with a dash-dotted line on the y-axis, the day of primary diagnosis is indicated with a dotted line on the x-axis. The chosen model parameters are $a = 7 \cdot 10^{-3}$, $K = 10^{12}$ and $m = 2 \cdot 10^{-7}$.

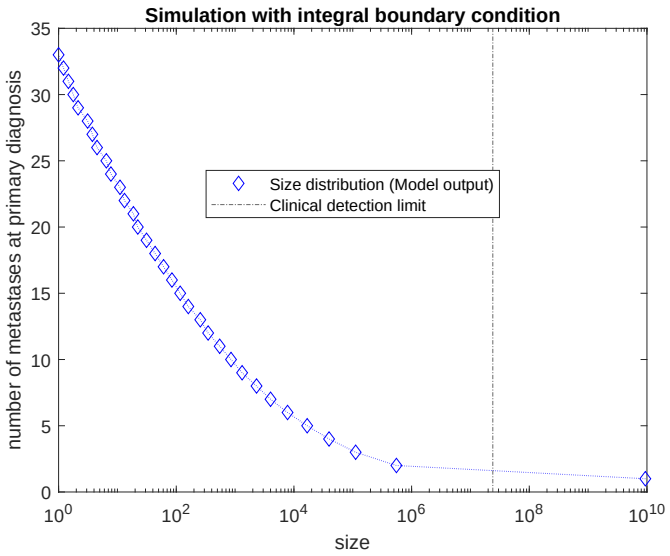


Figure 5.11: Simulation of the metastatic density distribution for the model considering no treatment and secondary metastasising, log-scaled, evaluated at primary diagnosis. Blue diamonds indicate the integer values, thus 'full' metastases. The largest tumor shown here corresponds to the primary tumor. Critical size for a tumor to be clinically detectable is indicated with a dash-dotted line. The chosen model parameters are $a = 7 \cdot 10^{-3}$, $K = 10^{12}$ and $m = 2 \cdot 10^{-7}$.

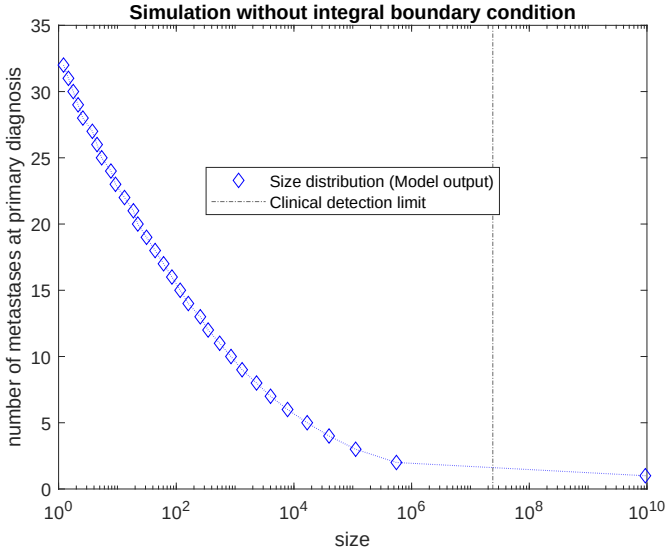


Figure 5.12: Simulation of the metastatic density distribution for the model considering no treatment and only primary metastasising, log-scaled, evaluated at primary diagnosis. Blue diamonds indicate the integer values, thus 'full' metastases. The largest tumor shown here corresponds to the primary tumor. Critical size for a tumor to be clinically detectable is indicated with a dash-dotted line. The chosen model parameters are $a = 7 \cdot 10^{-3}$, $K = 10^{12}$ and $m = 2 \cdot 10^{-7}$.

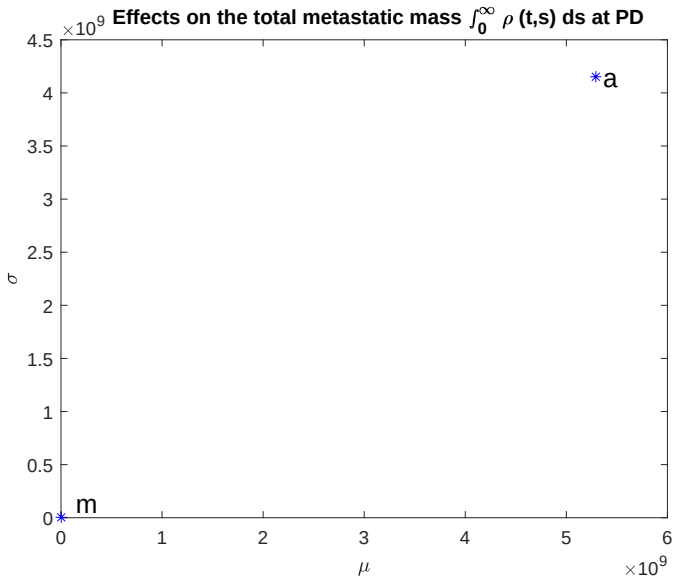


Figure 5.13: Elementary effects method as sensitivity analysis on the metastatic mass with secondary metastasising and without treatment at the day of primary diagnosis (PD).

### 5.3.3 Sensitivity analysis

For the sensitivity analysis of the transport equation, the investigation with derivative-based method might turn out analytically difficult to solve and also neglects combination effects of the parameter perturbations. Therefore for the models including secondary metastasising of equations (4.31) - (4.33) and excluding secondary metastasising of equations (4.46) - (4.49) the elementary effects method and the Sobol' indices estimation are applied directly. For both models, the influence of parameter perturbations on the total metastatic mass and the number of metastases at the day of primary diagnosis was examined, respectively. Figures 5.13 and 5.14 show the comparison of influence of perturbations of the two parameters $a$ and $m$ on the metastatic mass at the day of primary diagnosis for included secondary metastasizing and figures 5.15 and 5.16 without secondary metastasizing. On the other hand, figures 5.17 and 5.18 as well as figures 5.19 and 5.20 show the influence of the same parameter perturbations on parameters $a$ and $m$ on the number of metastases at the day of primary diagnosis for possible secondary metastasizing and without secondary metastasizing, respectively.

The Sobol' indices were implemented as introduced before and sampled for 15 000 pairs of parameter values from biologically reasonable domains. The advantage of the

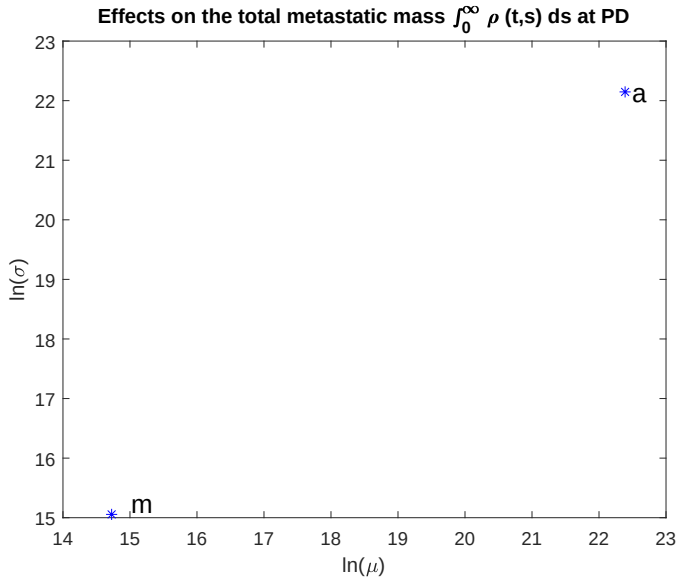


Figure 5.14: Elementary effects method as sensitivity analysis on the metastatic mass with secondary metastasising and without treatment at the day of primary diagnosis (PD), log-scaled.

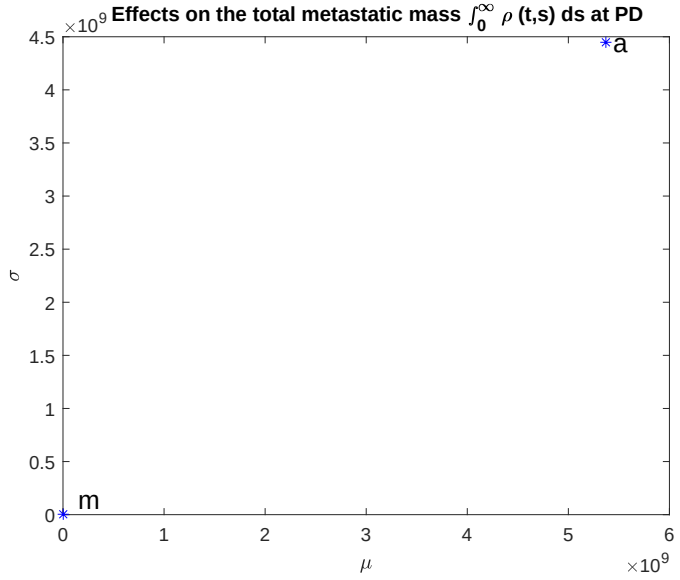


Figure 5.15: Elementary effects method as sensitivity analysis on the metastatic mass neglecting secondary metastasising without treatment at the day of primary diagnosis (PD).

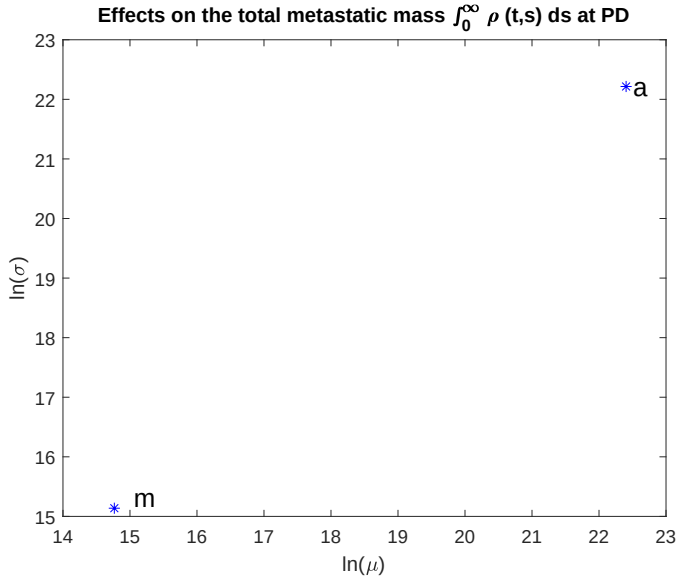


Figure 5.16: Elementary effects method as sensitivity analysis on the metastatic mass neglecting secondary metastasising without treatment at the day of primary diagnosis (PD), log-scaled.

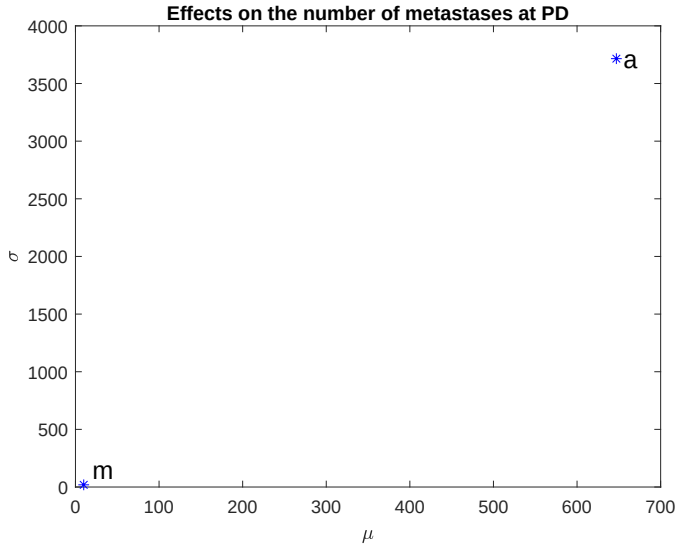


Figure 5.17: Elementary effects method as sensitivity analysis on the number of metastases with secondary metastasising and without treatment at the day of primary diagnosis (PD).

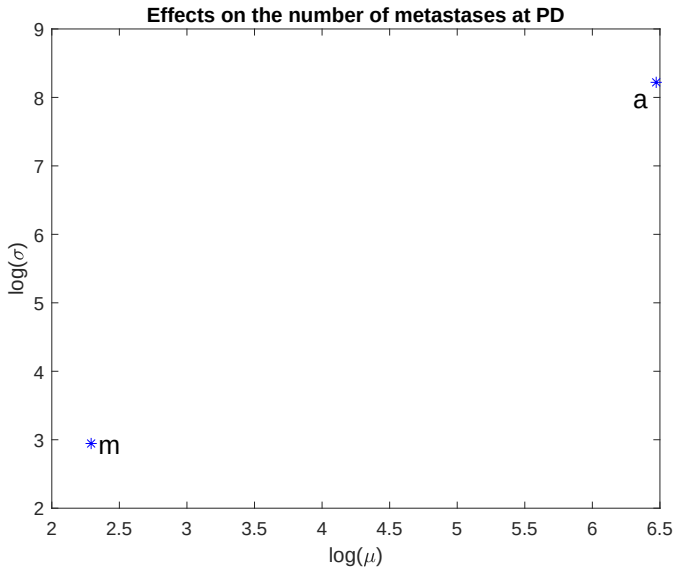


Figure 5.18: Elementary effects method as sensitivity analysis on the number of metastases with secondary metastasising and without treatment at the day of primary diagnosis (PD), log-scaled.

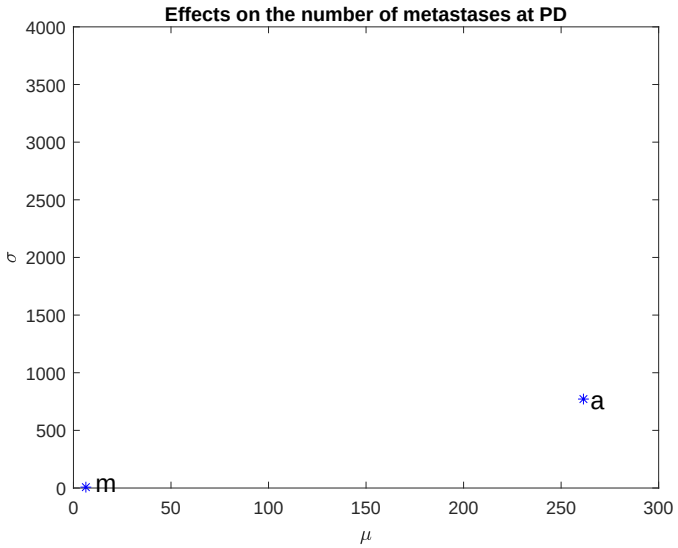


Figure 5.19: Elementary effects method as sensitivity analysis on the number of metastases neglecting secondary metastasising without treatment at the day of primary diagnosis (PD).

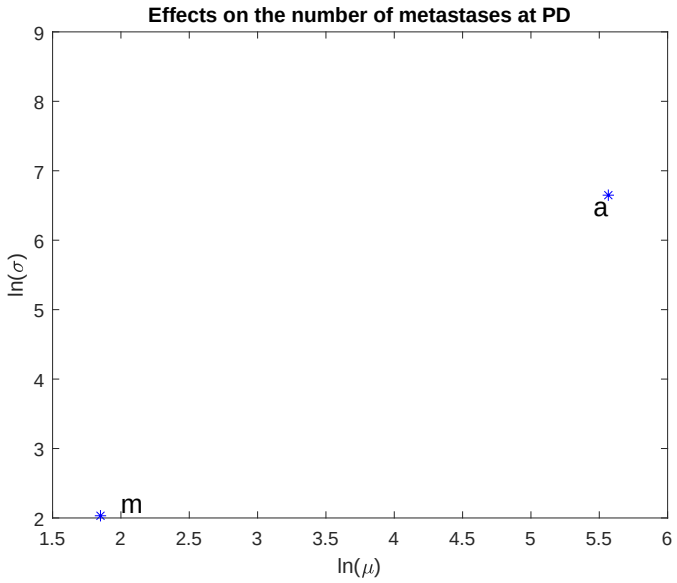


Figure 5.20: Elementary effects method as sensitivity analysis on the number of metastases neglecting secondary metastasising without treatment at the day of primary diagnosis (PD), log-scaled.

estimated indices over the elementary effects method from above is that these indices also consider combinations of multiple parameter perturbations at the same time. For the number of metastases calculated from the model including secondary metastasizing at the time of primary diagnosis the indices were estimated as

| Index | $m$ | $a$ |
|---|---|---|
| $S_i$ | 0.0285 | 0.7435 |
| $S_{tot}$ | 0.3532 | 0.9740 |

and estimated as

| Index | $m$ | $a$ |
|---|---|---|
| $S_i$ | 0.0387 | 0.7128 |
| $S_{tot}$ | 0.2255 | 0.9617 |

for the model without secondary metastasizing. We see that for both models the tumor growth parameter $a$ has the larger sensitivity on the number of metastases compared to the colonization coefficient $m$. Further, this indicates that the colonization coefficient parameter is of higher sensitivity in parallel combined perturbation with the tumor growth parameter than of the colonization coefficient parameter alone.
For the total metastatic mass the sensitivities of the model with secondary metastasizing was estimated as

| Index | $m$ | $a$ |
|---|---|---|
| $S_i$ | 0.0086 | 0.9981 |
| $S_{tot}$ | 0.0289 | 1.0062 |

compared to the sensitivity of the model without secondary metastasizing with respect to the total metastatic mass

| Index | $m$ | $a$ |
|---|---|---|
| $S_i$ | 0.0058 | 0.9839 |
| $S_{tot}$ | 0.0110 | 1.0037 |

Again, the parameter $a$ has the massive increase in total effect sensitivity on the total metastatic mass and the combined effect of perturbing $a$ and $m$ at the same time yields higher outcome differences than a single perturbation of the parameter $m$. This insight is gained exclusively from the Sobol' indices approach, whereas the ranking of the two parameters with respect to their sensitivity was also observed in the estimations of the elementary effects method before.

## 5.4 Metastatic model with systemic therapy

### 5.4.1 Implementation

Since there is no analytical expression to fully solve the metastatic model of equations (4.37) - (4.41) including treatment, a discretisation was performed to implement the model in Matlab (this subsection has been published in the appendix of [SKS21]).
The discretisation follows the computation of partial differential equation solutions along their characteristics. The characteristics of the partial differential equation system (4.37) - (4.41) are exactly the growth rates $g(x,t)$. They can be solved stepwise when knowing the therapy regimen (which is obviously the case for patients with a clinical history).

Let $T_{max}$ be the maximum observation time of some patient, i.e. the time interval of interest for that patient is $[0, T_{max}]$. Assume to know the time of primary diagnosis $t_{PD} \in [0, T_{max}]$ and to know the starting times of the $j'th$ treatment interval of chemo- and immunotherapy with drug $i$, respectively called $t_{C,j,i,s}, t_{I,j,i,s}$ for $j = 1, ..., n$. Further, assume to also know the end times of those treatments denoted by $t_{C,j,i,e}$ and $t_{I,j,i,e}$. Assume further that patients do not have different chemo- or immunotherapy being applied at the same time, we have that $t_{\cdot,j,i,s} < t_{\cdot,j,i,e}$ and $t_{\cdot,j,i,e} < t_{\cdot,j+1,i,s}$ for all $j$.
By construction of the model equations (4.9) and (4.17) we have that for $t \in \left[t_{C,j,i,s}, t_{C,j,i,e}\right]$ it is $\mathbb{1}_{C_i}(t) = 1$ (and zero otherwise) and that for $t \in \left[t_{I,j,i,s}, t_{I,j,i,e}\right]$ it is $\mathbb{1}_{I_i}(t) = 1$ (and zero otherwise).

For the untreated time intervals (i.e. for all $t$ at which $\mathbb{1}_{C_i}(t) = 0$ and $\mathbb{1}_{I_i}(t) = 1$ for all $i$) we can solve the growth rate $g(x,t)$ analytically (cf. equations (4.12) and (4.21) with lemma 4.2.4).
For time intervals of chemotherapeutic treatment (i.e. for all $t$ at which some $\mathbb{1}_{C_i}(t) = 1$) first the initial value problem (4.10) is solved to have an explicit expression for the function $\mu_i(t)$, which is equation (4.11) by lemma 4.2.1. Then this function is used to solve equation (4.12), the corresponding analytical solutions (4.15) and (4.16) are given in Theorem 4.2.3.
For time intervals in the immunotherapeutic treatment applications the initial value problem (4.18) is solved. The dose is just applied in times $t$ where immunotherapy is active, i.e. $\mathbb{1}_{I_j}(t) = 1$, so this can be done step-wise until every next switch of the function $\mathbb{1}_{I_j}(t)$ (before the very first application, i.e. for $t \in [0, t_{I,1,i,s}]$, the function $c_i(t)$ simply equals zero since no drug has been applied yet).

Let now $k = 1$ be a constant time step discretisation on the time interval $[0, T_{max}]$ such that $t_j = jk$ for $j \in \{0, ..., n\}$ with $t_0 = 0$ and $t_n = T_{max}$. For times $t \leq t_{\cdot,1,i,s}$ that represent the untreated setting, we can restrict the sizes to the interval $[1, x(t_{\cdot,1,i,s})]$. We therefore choose the size discretisation $x_j = x(t_j)$ such that $x_0 = 1 < ... < x_r = x(t_{\cdot,1,i,s})$ for $r \leq n$. Like this, the points $(x_k, t_l)$ and $(x_{k+1}, t_{l+1})$ belong to the same characteristic for any $k, l \in [0, r-1]$. Then we can approximate the model for untreated tumor growth with the following explicit scheme making use of Riemann sums and density discretisations $\varrho(x(i), t(j)) \approx u_{x_i,t_j}$ (see therapy-free model implementation with secondary metastasizing and [BC01]):

$$\begin{aligned} u_{x_0,0} &= 1 \\ u_{x_i,0} &= 0 \qquad i = 1, ..., r \\ u_{x_0,t_j} &= \frac{1}{1 - m x_0^\alpha} \sum_{i=1}^{n-1} m x_i^\alpha u_{x_i,t_j} \qquad j = 1, ..., r \\ u_{x_i,t_j} &= u_{x_{i-1},t_{j-1}} \qquad i, j = 1, ..., r \end{aligned} \tag{5.13}$$

For times in active treatment, i.e. $t \in \{r, t_e\}$ with $t_e \leq T_{max}$, we know that the tumor sizes are affected. We start to count from the switch with index $j$ and recalculate the new size first. The computation of the partial differential equation's boundary condition is analogous to before. So for any therapy switch at time point $r$ we have for

chemotherapy with $\mu_{d,r+j} = \mu_d(r+j)$ that

$$\begin{aligned} x_{i,1} &= x_i\left(1-\mu_{d,r+1}\right) \qquad i = 0,...,r \\ x_{i,j} &= x_{i-1,j-1}\left(1-\mu_{d,r+j}\right) \qquad i = 0,...,r; j = 2,...,t_e - r \\ u_{x_{i,1},r+1} &= u_{x_i,r} \qquad i = 0,...,r \\ u_{1,(j+r)} &= \frac{1}{1-m}\sum_{i=1}^{n-1} m u_{x_{i,j-1},(j-1+r)}\mathbb{1}_{x_{i,j-1}\geq 1} \qquad j = 2,...,t_e - r \end{aligned} \tag{5.14}$$

and for immunotherapy with $c_{d,r+1} = c_d(r+j)$ as the step wise numerical solution of equation (4.18) or using the analytical solution of lemma 4.2.4 that

$$\begin{aligned} x_{i,1} &= x_i\left(1-c_{d,r+1}\right) \qquad i = 0,...,r \\ x_{i,j} &= x_{i-1,j-1}\left(1-c_{d,r+j}\right) \qquad i = 0,...,r; j = 2,...,t_e - r \\ u_{x_{i,1},r+1} &= u_{x_i,r} \qquad i = 0,...,r \\ u_{1,(j+r)} &= \frac{1}{1-m}\sum_{i=1}^{n-1} m u_{x_{i,j-1},(j-1+r)}\mathbb{1}_{x_{i,j-1}\geq 1} \qquad j = 2,...,t_e - r \end{aligned} \tag{5.15}$$

### 5.4.2 Simulation

The implementation presented in the previous subsection was programmed in Matlab. Considering some treatment close to reality, the example treatment protocol was reintroduced: starting with a first-line chemotherapy three weeks after primary diagnosis, which is applied for four weeks. A treatment break of one week is followed by the immunotherapeutic treatment. The parameters used for this simulations were the growth rate $a = 7 \cdot 10^{-3}$, tumor carrying capacity $K = 10^{12}$, initial chemotherapy efficacy $\mu_0 = 0.25$, refractory effect on chemotherapy $\mu^* = 0.85$, immunotherapeutic efficacy $\chi = 0.1$ and drug-specific concentration for half-maximal response $c_i^{50} = 1.01 \cdot 10^{16}$. The drug implemented was Nivolumab, therefore the drug-specific parameters were again chosen as dosage $d_i = 0.480$, application interval length $l = 28$, molar mass $M_i = 1.46 \cdot 10^5$ and drug-specific half-life $t_i^{1/2} = 26.7$.
The full time course of primary tumor (blue line) and metastases (red lines) are shown in Figure 5.21 with chemo- and immunotherapeutic treatment application indicated with shaded areas. It can clearly be seen that the metastases are quite close to be detected clinically, still none of them crossed the visibility threshold during the whole time course. The primary tumor shrinks during immunotherapeutic treatment and eventually its size decreases below the clinical detection threshold. Evaluating the metastatic density distributions gives an idea of the full metastatic threat. Figures 5.22 and 5.23 are the corresponding numerical approximations of the metastatic density

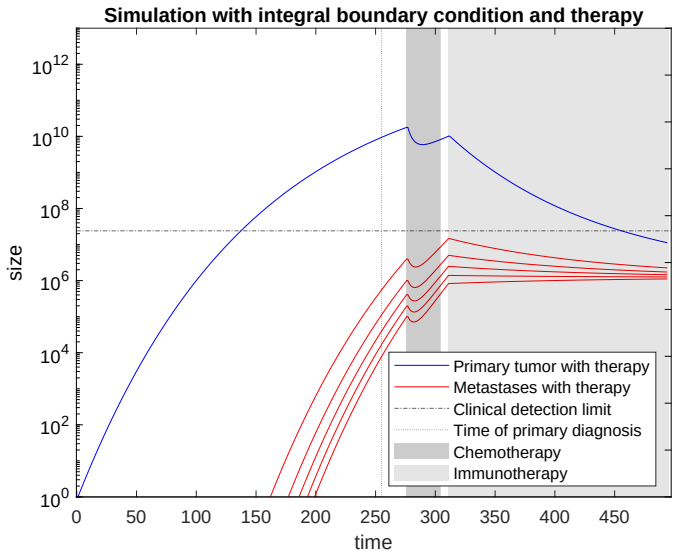


Figure 5.21: Full time course of the primary tumor (blue solid line) and metastases (red lines) for the metastatic density distribution model with therapy. The parameters used for this simulation are $a = 7 \cdot 10^{-3}$, $K = 10^{12}$, $\mu_0 = 0.25$, $\mu^* = 0.85$, $\chi = 0.1$ and $c_i^{50} = 1.01 \cdot 10^{16}$. The drug implemented was Nivolumab, therefore the drug-specific parameters were again chosen as $d_i = 0.480$, $l = 28$, $M_i = 1.46 \cdot 10^5$ and $t_i^{1/2} = 26.7$. Critical size for a tumor to be clinically detectable is indicated with a dash-dotted line on the y-axis, the day of primary diagnosis is indicated with a dotted line on the x-axis.

distribution at primary diagnosis and 90 days after primary diagnosis, respectively. Note that in both, the diamonds correspond to integer metastatic values again, showing fully countable metastases. However, the largest diamond is the primary tumor.

### 5.4.3 Sensitivity analysis

For the model with equations (4.37) - (4.41), the derivative OAT approach is again not leading to useful insights. Since we already observed that for both number of metastases and total metastatic mass in the untreated tumor setting combined parameter perturbations play an important role, for the general therapy setting the elementary effects method and the estimation of the Sobol' indices was introduced and applied once more.
To have comparable outcome estimations, the example standard treatment from the previous ordinary differential equation analysis was reused: a first-line chemotherapy

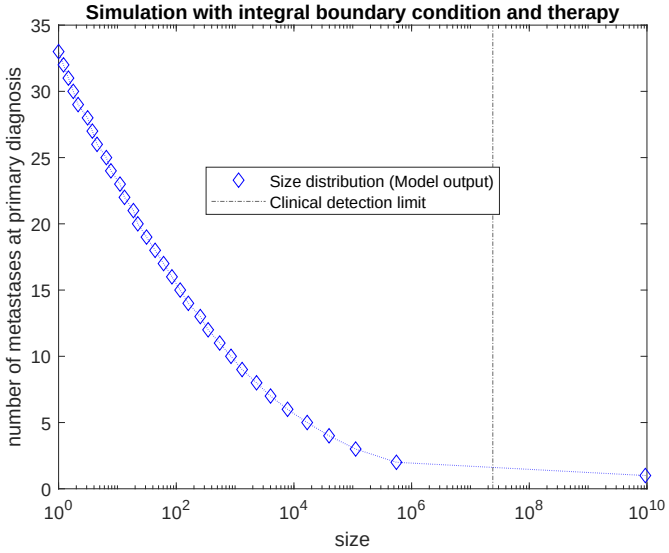


Figure 5.22: Metastatic density distribution corresponding to the previously mentioned simulation, evaluated at the day of primary diagnosis.

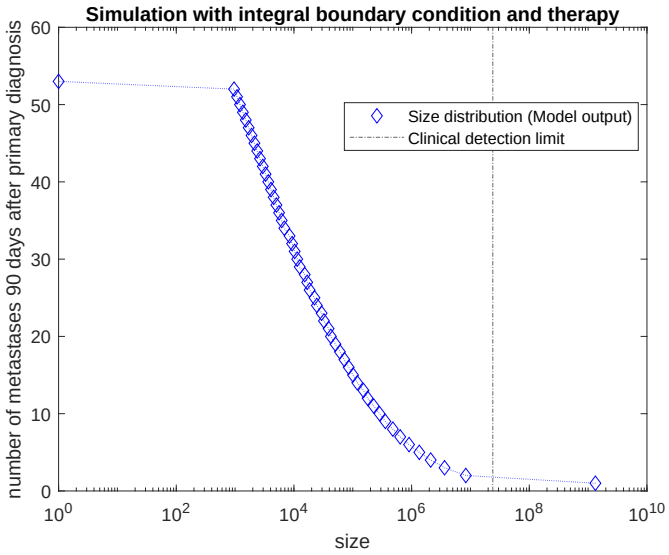


Figure 5.23: Metastatic density distribution corresponding to the previously mentioned simulation, evaluated 90 days after primary diagnosis.

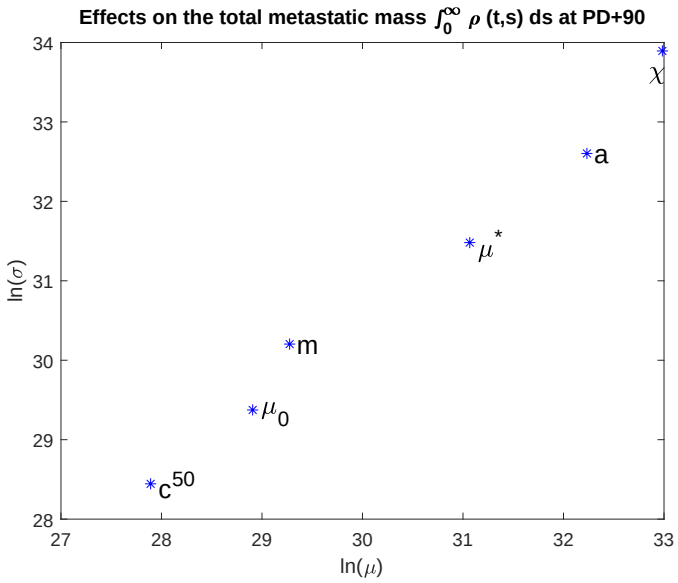


Figure 5.24: Elementary effects method as sensitivity analysis on the metastatic mass under example treatment at day 90 after primary diagnosis.

is started three weeks after primary diagnosis and is applied for four weeks. The immunotherapeutic treatment follows after another week without treatment. The implementation was performed for biologically reasonable parameter intervals with 100 000 parameter pairs and checking the number of metastases as well as the total metastatic mass 90 days after primary diagnosis (24 days after initialising the immunotherapeutic treatment). The results for the elementary effects methods are shown in figure 5.24 and 5.25, respectively. The ranking clearly shows that in both endpoints for sensitivities on the total metastatic mass and the number of metastases, the parameter $\chi$, the immunotherapeutic drug effect, has the highest influence on the respective covariate. This is followed by the tumor growth rate $a$. On the other hand we can identify that perturbations of $\mu^*$ have relatively higher influence on the total metastatic mass and relatively lower influence on the number of metastases.

The Sobol' indices for the total metastatic mass 90 days after primary diagnosis were estimated as

| Index | $m$ | $c^{50}$ | $\chi$ | $a$ | $\mu^*$ | $\mu_0$ |
|---|---|---|---|---|---|---|
| $S_i$ | 0.0035 | 0.0023 | 0.1989 | 0.0164 | 0.0010 | 0.0010 |
| $S_{tot}$ | 0.3974 | 0.0300 | 0.9820 | 0.6322 | 0.0302 | 0.0285 |

while the corresponding Sobol' indices for the number of metastases 90 days after

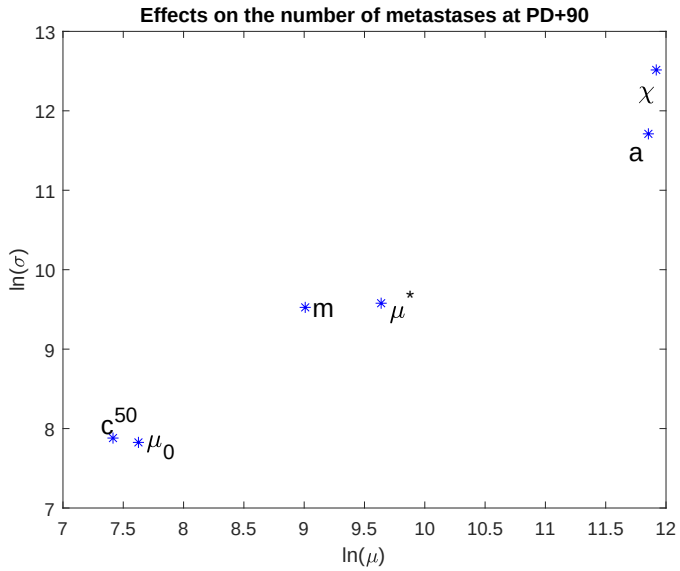


Figure 5.25: Elementary effects method as sensitivity analysis on the number of metastases under example treatment at day 90 after primary diagnosis.

primary diagnosis were calculated as

| Index | $m$ | $c^{50}$ | $\chi$ | $a$ | $\mu^*$ | $\mu_0$ |
|---|---|---|---|---|---|---|
| $S_i$ | 0.0162 | $-0.0024$ | 0.1622 | 0.0065 | $-0.0023$ | $-0.0022$ |
| $S_{tot}$ | 0.5630 | 0.0498 | 0.9620 | 0.6581 | 0.0456 | 0.0456 |

The results of the elementary effects method do reproduce in these estimations. The largest sensitivities are attributed to the two parameters $\chi$ and $a$. However, the effects of the parameter $\mu^*$ are a lot smaller. Interestingly, the total order sensitivities unveil that the parameter $m$ has huge interaction effects with the other parameters. This means that multiple parameter perturbations including perturbations of the parameter $m$ are a lot more sensitive than the parameters $c^{50}$, $\mu_0$ and in particular $\mu^*$ alone.

# 6 Simulation for clinical cases

For the two available data sets of treated patients with NSCLC, the GlobalSearch environment [Ugr+07] of Matlab was performed on a least squares minimization function with respect to the number and size of clinically detectable and measured metastases towards the simulated values.

## 6.1 Metastatic model with systemic therapy

This whole section is based on methods and results previously published in [SKS21] and proves that the modeling framework is capable of describing clinical treatment protocols and measurement series. For this step and due to the long time course of the patients' clinical history, secondary metastasising was considered.
The parameters fixed for evaluation were the tumor carrying capacity $K = 10^{12}$ and $\alpha = 2/3$ [IKS00]. For the drug-specific parameters, the values presented in section 2.3 of this thesis were used.

### 6.1.1 Parameter values

The whole measurement series were implemented to be evaluated with the minimization problem. The resulting parameters are shown in table 6.1 and are in biologically reasonable ranges. The cell cycle length of a tumor can be estimated by $\frac{\ln 2}{a}$ [Bil+19], yielding values between 93 and 101 hours for these three patients. These values are reasonably close to observed values of about two to four days [Kuf+03].

### 6.1.2 Simulations

The previously presented implementation scheme was programmed in Matlab to run the model on the respective clinical treatment history with the fitted and fixed parameter values. Results were published in [SKS21] in detail, the simulation for the total tumor burden, i.e. the integral of the density distribution $\int_0^\infty \varrho(s,t)ds$ over time, see equation (4.35), is shown in Figure 6.1 for the three individual patients of the first data set along with the corresponding concentrations of the immunotherapeutic drug that was applied at some point of the patients' clinical history. The comparison of clinical data with the

Table 6.1: Patient-specific model parameters for the model of metastatic development with respect to systemic treatment evaluated on clinical data.

| | Patient | KE-01 | KE-02 | KE-03 |
|---|---|---|---|---|
| | Explanation [unit] | | | |
| $K$ | Environmental carrying capacity [cells] [Kle09] | $10^{12}$ | $10^{12}$ | $10^{12}$ |
| $a$ | Growth rate [1/day] | $7.284 \cdot 10^{-3}$ | $6.877 \cdot 10^{-3}$ | $6.984 \cdot 10^{-3}$ |
| $m$ | Colonization coefficient [1/cell 1/day] | $1.635 \cdot 10^{-7}$ | $1.984 \cdot 10^{-7}$ | $2.738 \cdot 10^{-7}$ |
| $\alpha$ | Fractal dimension [-] [IKS00] | $2/3$ | $2/3$ | $2/3$ |
| | Chemotherapeutic drug in use | *CisPT/Pemet. (1L)* | - | *CisPT/Pemet. (1L)* |
| $\mu_0$ | Initial chemotherapy efficiency [-] | 0.237 | - | 0.081 |
| $\mu^*$ | Refractory effect for chemotherapy application [-] | 0.132 | - | 0.132 |
| | Immunotherapeutic drug in use | *Pembr. (2L)* | *Nivol. (1L)* | *Pembr. (2L)* |
| $\chi$ | Immunotherapeutic effect under application [1/day] | 0.067 | 0.499 | 0.088 |
| $c_i^{50}$ | Drug concentration of immunotherapeutic drug for half-maximal response [molecules per volume], cf. eq. (4.20) | $1.012 \cdot 10^{16}$ | $1.010 \cdot 10^{16}$ | $1.009 \cdot 10^{16}$ |
| $c_i^{st}$ | Drug concentration of immunotherapeutic drug in serum steady state [molecules per volume] | $1.41 \cdot 10^{18}$ | $2.77 \cdot 10^{18}$ | $1.41 \cdot 10^{18}$ |

*CisPT/Pemet.* = Combination therapy of Cisplatinum and Pemetrexed. *Pembr.* = Pembrolizumab. *Nivol.* = Nivolumab. *(1L)* = First-line therapy. *(2L)* = Second-line therapy.

metastatic density distribution for one of these patients at two different time points under treatment is shown in Figure 6.2. Simulations for the other two patients can be found on their respective comparison to clinical measurements in [SKS21].

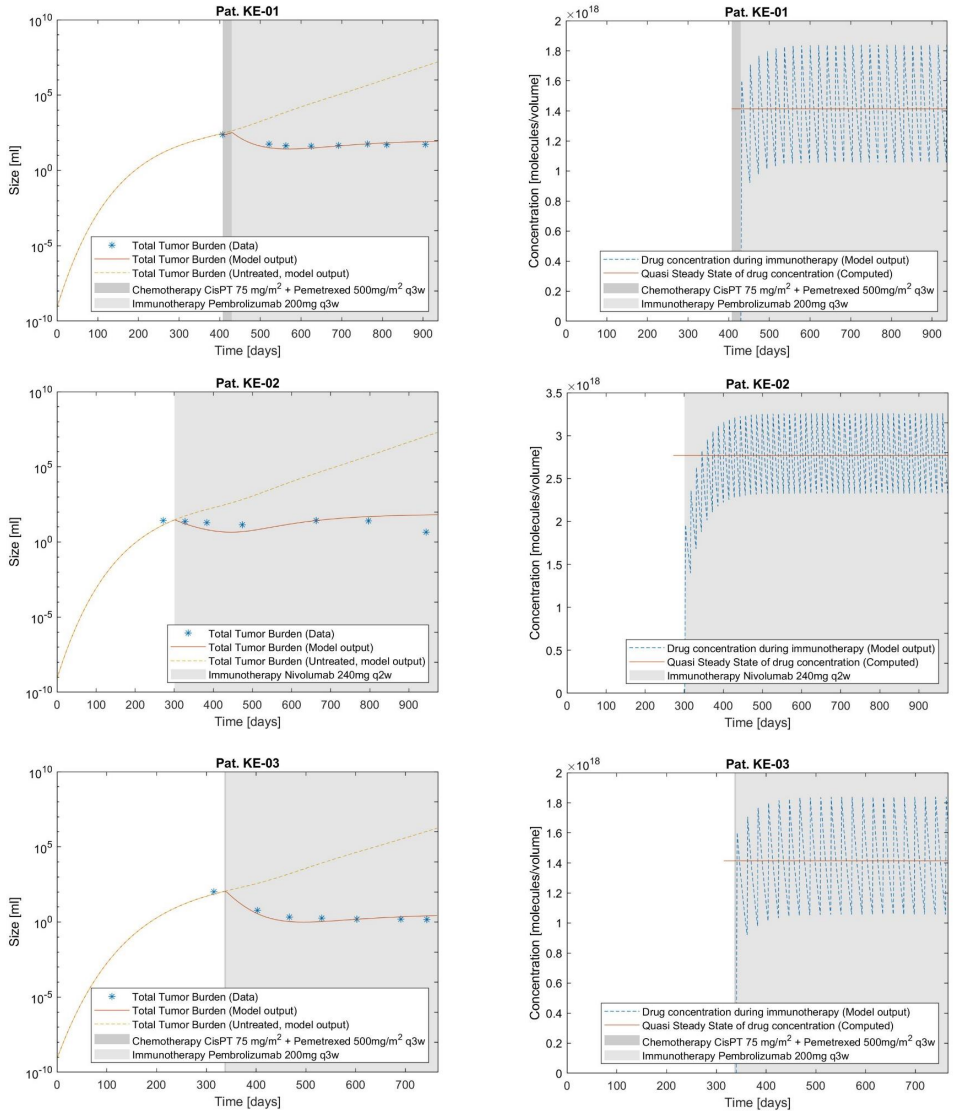


Figure 6.1: Model simulations over the entire clinical treatment history for three individual patients, using parameters shown in table 6.1. The graphs on the left shown the total tumor burden for all three patients over time (solid line) compared to a hypothetical untreated metastatic disease with the same growth parameters (dashed line) and compared to the clinical measurements (asterisks). The right graphs show the drug concentration of the immunotherapeutic drug over time (dashed line), compared to the calculated quasi-steady-state that the concentration circulates around. The shaded areas show the times where chemo- and immunotherapy were applied.

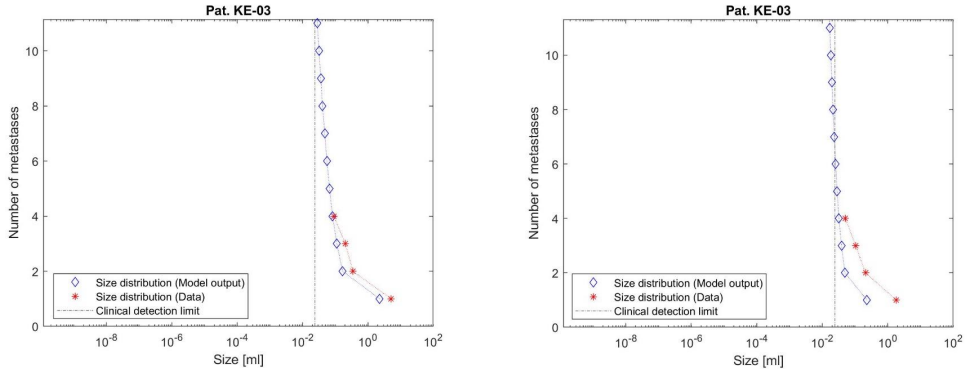


Figure 6.2: Model simulations for patient KE-03, using parameters shown in table 6.1. The simulated metastatic density distribution (blue diamonds) is plotted in comparison with the clinical measurements (red asterisks) and the clinical visibility threshold (dash-dotted line). The left graph shows the simulation and the measurements at day 88 after primary diagnosis, the right graph the same simulation and corresponding measurements at day 152 after primary diagnosis.

## 6.2 Metastatic model without treatment

This whole section is based on methods and results previously published in [Ben+22]. To assess prognostic value of the framework, no therapy was considered in the modeling approach. Further, the hypothesis was followed that secondary seeded metastases play a minor role in clinical outcome [Bet+12]. The data was used up to the time point where brain metastases were firstly detected clinically. Let this time point be denoted by $t_{BM}$. Then we have that the time of primary diagnosis, i.e., the time where the primary tumor was discovered first, denoted by $t_{PD}$, fulfills $t_{PD} \leq t_{BM}$. Assume that the primary tumor and the corresponding metastases show different growth parameters, i.e., $r, r_m$ and $a, a_m$ are generally different, respectively.
To reduce the number of free parameters for the optimization problem, the approach by [Bil+19] was followed. As usual, the tumor carrying capacity for both primary tumor and metastases was fixed to

$$K = e^{r/a} = e^{r_m/a_m} = 10^{12}. \tag{6.1}$$

By this property the parameter $a$,$a_m$ can be calculated directly from a determined parameter $r$,$r_m$.
A meta-analysis of different studies that examine doubling times of primary tumors at the time of primary diagnosis yields different doubling time values for different tumor

histologies. The meta-analysis is presented in the supplementary material to [Bil+19] and found mean values of 91, 104 and 201 days for undifferentiated carcinoma, SCC and ADC, respectively. These values can be used in equation (4.8) to determine the parameter values $r$, $r_m$ and therefore also $a$, $a_m$. For this, the size of the primary tumor at primary diagnosis has to be known. This was the case for 28 of the 31 patients of the second data set.
The parameter $\alpha$ was found to lead to biologically meaningful simulations exclusively for values below 0.3. Further manual exploration suggested that the value $\alpha = 0.1$ potentially describes best the clinical data for this data set. The two free parameters remaining for the fitting and optimization procedure are the metastatic growth rate $r_m = a_m \ln(K)$ and the metastatic colonization coefficient $m$.

### 6.2.1 Density smoothing

As the objective of this part of the model evaluation was to form predictions on the metastatic density distribution, especially on the sizes of non-detectable metastases, and since only data up to time $T_{BM}$ was used to fit the whole density distribution on, a huge standard error for parameter estimation was expected. To reduce this variance, a window of possible outcomes for the number of metastases for each of the individual patients was introduced. Three cases were established to smoothen the shape of the density distribution that may be distinguished as follows, using $v$ as the number of clinically detected metastases:

1. no additional brain metastasis, i.e. an artificial metastasis with size of the clinical detection limit introduced at the value of the clinically detected metastases $v$. This translates as the case that the clinicians have successfully detected all existing and detectable metastases.

2. one additional brain metastasis at the size of the clinical detection limit. This artificial metastasis with amount $v + 1$ is the case where the clinicians were unable to detect a metastasis that is close to the detection limit and can be interpreted as a possible 'worst case' scenario.

3. an artificial metastasis of size of the clinical detection limit introduced with an amount $v + \frac{1}{2}$. This case was used to generate parameters that were used in the statistical evaluation later in this section.

All parameters extracted from each of these three cases were highly correlated with the other respective parameters of the other two cases. The parameter estimations of the third case were used as reference for the following subsections.

### 6.2.2 Parameter values

All parameter values fitted by this procedure lie in reasonable biological ranges for all patients of the second data set. The description of the fitted parameter values is given as follows

| parameter | mean | median | std | min | max |
|---|---|---|---|---|---|
| $r_m$ | $4.93 \cdot 10^{-2}$ | $4.03 \cdot 10^{-2}$ | $3.18 \cdot 10^{-2}$ | $1.22 \cdot 10^{-2}$ | $1.34 \cdot 10^{-1}$ |
| $m$ | $3.57 \cdot 10^{-4}$ | $2.33 \cdot 10^{-4}$ | $2.77 \cdot 10^{-4}$ | $8.22 \cdot 10^{-5}$ | $1.22 \cdot 10^{-3}$ |

while the other parameters are either fixed to $\alpha = 0.1$ or calculated from equations (4.8) and (6.1), respectively:

| parameter | mean | median | std | min | max |
|---|---|---|---|---|---|
| $r$ | $4.75 \cdot 10^{-2}$ | $2.92 \cdot 10^{-2}$ | $4.91 \cdot 10^{-2}$ | $1.38 \cdot 10^{-2}$ | $2.51 \cdot 10^{-1}$ |
| $a$ | $1.72 \cdot 10^{-3}$ | $1.06 \cdot 10^{-3}$ | $1.78 \cdot 10^{-3}$ | $4.99 \cdot 10^{-4}$ | $9.08 \cdot 10^{-3}$ |
| $a_m$ | $1.79 \cdot 10^{-3}$ | $1.46 \cdot 10^{-3}$ | $1.15 \cdot 10^{-3}$ | $4.40 \cdot 10^{-4}$ | $4.84 \cdot 10^{-3}$ |

### 6.2.3 Simulations

Figure 6.3 shows an example patient for the second data set, subplot A is the clinical history of the patient. The resulting simulations of case 3 yield the metastatic density distribution over time until time $t_{BM}$. This density distribution can be evaluated at times $t_{PD}$ and $t_{BM}$ to identify number and size of all metastases that exist at the respective time point - both clinically detectable and clinically undetectable ones. An example calculation and density distribution plot for these two time points can be found in Figure 6.3, subplot B and C, respectively. The full time course of primary tumor and metastases can be calculated by this model formulation until time $t_{BM}$, results of this are shown in subplot D. Subplot E finally presents the calculated metastatic sizes (black) compared to the sizes of clinically detected metastases at $t_{BM}$ (gray).
For all the other patients of the second data set where the method was applicable, i.e., the primary tumor size at time of primary diagnosis was known, the data and simulations are fully presented in the very exact same structure in the appendix of [Ben+22].
Since the whole clinical history for these patients is known, we were able to compare the prediction of the metastatic density distribution at primary diagnosis to the actual number of different clinically detected metastases during the whole treatment course.

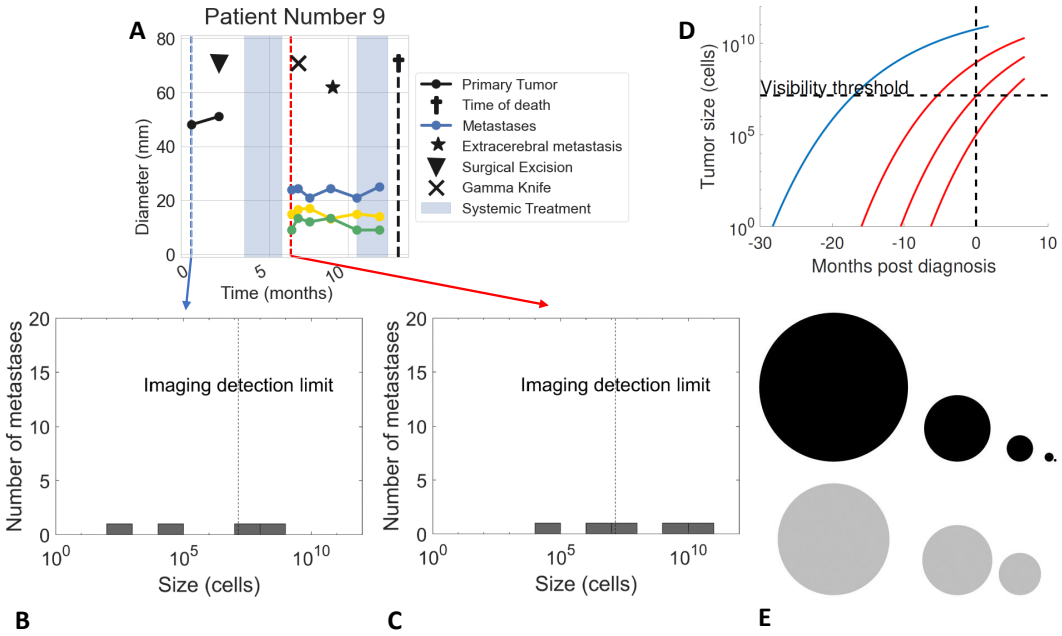


Figure 6.3: Model simulations for patient 9 of the second data set, using parameters from the fitting procedure presented before. The clinical history is shown in subplot A. Subplot B and C present the model evaluations of the metastatic density distribution at primary diagnosis (blue dashed line) and time $t_{BM}$ (red dashed line), respectively. Subplot D shows the calculated time course of the primary tumor (blue solid line) and the metastases (red solid lines). Subplot E compares the calculated sizes of metastases (black) to the measured sizes of clinically detected metastases (gray) at time $t_{BM}$. Taken from [Ben+22].

The accuracy to correctly predict a range of number of metastases (using case 1 and 2) that included the observed real number of metastases during the whole treatment course had an astonishing accuracy of up to 82%.

### 6.2.4 Statistical evaluation

To identify the prediction possibilities of the fitted parameters $r_m$ and $m$, the dichotomized Kaplan-Meier curves were examined for both, overall survival (OS) and progression-free survival (PFS). For the growth rate of the metastases, parameter $r_m$, significant differences in the survival curves were observed for both OS (p=0.0026) and PFS (p=0.0108) performing a log-rank test. For $m$, the split at the median value was not found to be significant, but a split for the 80% quantile resulted in significant differences in the survival curves for again both OS (p=0.0356) and PFS (p=0.0254). This could potentially show that the computationally determined parameters $r_m$ and $m$ can indeed be interpreted as computational biomarkers, as observed in earlier publications [Ben+21]. Parameter $r_m$ is suitable as a general biomarker, whereas parameter $m$ is eventually more useful to identify patients at high-risk. The evaluation of survival curves is presented in Figure 6.4.

Univariate and multivariate Cox regression were performed on the known clinical parameters for both OS and PFS using the *lifelines* python package. To assess the clinical prediction benefit of the computational biomarkers $r_m$ and $m$, the univariate and multivariate Cox regression models were first evaluated without, then with the two parameters. The corresponding results of the Cox regression analyses are shown in tables 6.2 and 6.3 for univariate analysis and in tables 6.4 and 6.5 for multivariate analysis.

It is noteworthy that the predictive model without the two computational biomarkers has no significant covariate in the univariate analysis. However, the predictive model with the two computational biomarkers has both of them significant for OS (p=0.0229 for $r_m$ and p=0.0011 for $m$). The estimated hazard ratios seem biologically reasonable, in particular for the univariate analysis the covariates $r_m$ and $m$ have two of the three largest positive values for both OS and PFS.
For the multivariate analysis, the two computational biomarkers again have very large values for hazard ratio and a significant p-value for $r_m$.

The covariates with a p-value below 0.2 in the univariate analysis were chosen for another multivariate analysis with and without the computational biomarkers, respectively. These Cox regression models are referred to as the 'reduced multivariate Cox

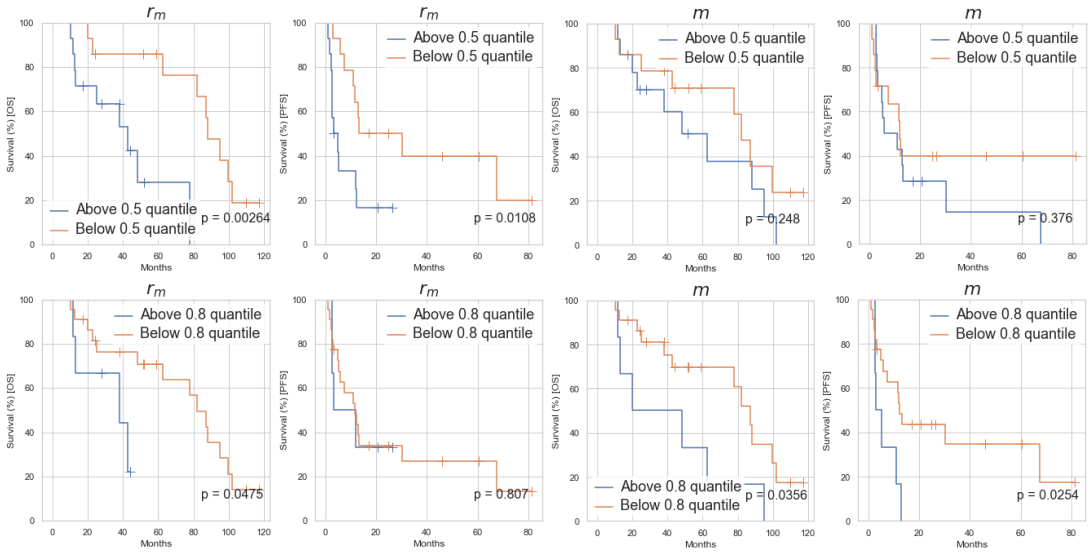


Figure 6.4: Kaplan-Meier survival curves for the computational parameters $r_m$ and $m$. The first row shows the survival curves above and below the population's mean value of the respective parameter, whereas the second row shows the split for the 80% quantile. The first and the third column represent the overall survival, second and fourth column focus on the progression-free survival. The corresponding p-value from a corresponding log-rank test for dichotomized groups is shown in the right lower corner. Taken from [Ben+22].

Table 6.2: Results of the univariate Cox regression analysis for OS and PFS without the computational biomarkers $r_m$ and $m$.

| | HR (OS) | p-value (OS) | 95% CI (OS) | HR (PFS) | p-value (PFS) | 95% CI (PFS) |
|---|---|---|---|---|---|---|
| Covariate | | | | | | |
| Sex | 1.14 | 0.619 | 0.683-1.90 | 0.92 | 0.711 | 0.596-1.42 |
| Age at diagnosis | 1.04 | 0.893 | 0.604-1.78 | 0.68 | 0.101 | 0.434-1.08 |
| Number of bm at relapse | 1.37 | 0.150 | 0.892-2.11 | 1.42 | 0.104 | 0.931-2.15 |
| Stage at diagnosis | 1.17 | 0.552 | 0.695-1.98 | 1.01 | 0.963 | 0.639-1.60 |
| Histology of pt | 0.68 | 0.134 | 0.413-1.13 | 0.78 | 0.272 | 0.504-1.21 |
| Size of pt at diagnosis | 1.13 | 0.604 | 0.704-1.83 | 0.90 | 0.648 | 0.574-1.41 |
| Size of bm at relapse | 1.30 | 0.189 | 0.877-1.94 | 1.63 | 0.020 | 1.08-2.46 |

HR = Hazard Ratio, CI = Confidence Interval, OS = overall survival, PFS = progression-free survival, bm = brain metastases, pt = primary tumor

models'.
To examine whether the two computational biomarkers provide clinical prediction benefit, the multivariate Cox regression models with and without the computational parameters for all covariates and in the reduced multivariate models were trained in a three-fold cross-validation learning set. For the corresponding test set, Harrel's concordance index was calculated. The whole procedure was repeated one hundred times, calculating the overall c-index as the mean value of the one hundred estimated ones.
For PFS, the c-indices of the full models were estimated as 0.560 (95% CI 0.545 to 0.575) and 0.595 (95% CI 0.582 to 0.608) without and with the computational parameters, respectively. The slight improvement in predictive value disappears when only considering covariates with p<0.2: the corresponding c-indices were calculated as 0.708 (95% CI 0.697 to 0.718) and 0.703 (95% CI 0.693 to 0.713) for the two approaches.
For OS, the c-index for the two full models without and with the computational biomarkers were 0.585 (95% CI 0.569 to 0.602) and 0.713 (95% CI 0.700 to 0.726), respec-

Table 6.3: Results of the univariate Cox regression analysis for OS and PFS with the computational biomarkers $r_m$ and $m$.

| | HR (OS) | p-value (OS) | 95% CI (OS) | HR (PFS) | p-value (PFS) | 95% CI (PFS) |
|---|---|---|---|---|---|---|
| Covariate | | | | | | |
| Sex | 1.10 | 0.602 | 0.768-1.58 | 0.98 | 0.936 | 0.663-1.46 |
| Age at diagnosis | 1.03 | 0.900 | 0.679-1.55 | 0.74 | 0.167 | 0.479-1.14 |
| Number of bm at relapse | 1.72 | 0.0056 | 1.17-2.53 | 1.25 | 0.255 | 0.851-1.84 |
| Stage at diagnosis | 0.87 | 0.435 | 0.604-1.24 | 1.10 | 0.674 | 0.716-1.67 |
| Histology of pt | 0.79 | 0.223 | 0.545-1.15 | 0.72 | 0.117 | 0.470-1.09 |
| Parameter $r_m$ | 1.65 | 0.0229 | 1.07-2.53 | 1.25 | 0.237 | 0.862-1.83 |
| Parameter $m$ | 1.95 | 0.0011 | 1.31-2.91 | 1.54 | 0.073 | 0.961-2.46 |
| Size of pt at diagnosis | 1.05 | 0.781 | 0.755-1.45 | 0.90 | 0.648 | 0.574-1.41 |
| Size of bm at relapse | 0.99 | 0.967 | 0.677-1.45 | 1.52 | 0.040 | 1.02-2.26 |

HR = Hazard Ratio, CI = Confidence Interval, OS = overall survival, PFS = progression-free survival, bm = brain metastases, pt = primary tumor

tively. This increase in predictive value is even observable for the two corresponding reduced models for covariates with p<0.2 only: the corresponding indices were 0.647 (95% CI 0.631 to 0.664) and 0.789 (95% CI 0.779 to 0.800), respectively. This can be interpreted as an improvement in predictive power by 22%, compared to the Cox proportional hazards model based on usual covariate routinely acquired in the clinics and clearly shows the astonishing potential of the modeling approach and parameter interpretation in prediction context, supporting earlier discoveries [Ben+21].

Table 6.4: Results of the multivariate Cox regression analysis for OS and PFS without the computational biomarkers $r_m$ and $m$.

| | HR (OS) | p-value (OS) | 95% CI (OS) | HR (PFS) | p-value (PFS) | 95% CI (PFS) |
|---|---|---|---|---|---|---|
| Covariate | | | | | | |
| Sex | 1.13 | 0.662 | 0.649-1.98 | 1.00 | 0.999 | 0.630-1.59 |
| Age at diagnosis | 1.51 | 0.248 | 0.752-3.02 | 0.75 | 0.355 | 0.402-1.39 |
| Number of bm at relapse | 1.44 | 0.175 | 0.851-2.43 | 1.35 | 0.196 | 0.856-2.14 |
| Stage at diagnosis | 1.00 | 0.999 | 0.904-1.11 | 0.95 | 0.840 | 0.561-1.60 |
| Histology of pt | 0.65 | 0.214 | 0.326-1.28 | 1.00 | 0.987 | 0.569-1.77 |
| Size of pt at diagnosis | 1.24 | 0.387 | 0.762-2.01 | 1.00 | 0.999 | 0.979-1.02 |
| Size of bm at relapse | 1.39 | 0.113 | 0.925-2.10 | 1.65 | 0.0215 | 1.08-2.52 |

HR = Hazard Ratio, CI = Confidence Interval, OS = overall survival, PFS = progression-free survival, bm = brain metastases, pt = primary tumor

Table 6.5: Results of the multivariate Cox regression analysis for OS and PFS with the computational biomarkers $r_m$ and $m$.

| | HR (OS) | p-value (OS) | 95% CI (OS) | HR (PFS) | p-value (PFS) | 95% CI (PFS) |
|---|---|---|---|---|---|---|
| Covariate | | | | | | |
| Sex | 1.10 | 0.769 | 0.548-2.07 | 1.01 | 0.979 | 0.637-1.59 |
| Age at diagnosis | 1.99 | 0.095 | 0.887-4.47 | 0.61 | 0.174 | 0.300-1.24 |
| Number of bm at relapse | 1.32 | 0.506 | 0.586-2.96 | 1.05 | 0.894 | 0.52-2.11 |
| Stage at diagnosis | 1.48 | 0.266 | 0.740-2.97 | 1.17 | 0.585 | 0.667-2.05 |
| Histology of pt | 1.06 | 0.879 | 0.475-2.39 | 1.56 | 0.258 | 0.721-3.39 |
| Parameter $r_m$ | 3.40 | 0.0039 | 1.48-7.81 | 1.52 | 0.219 | 0.780-2.96 |
| Parameter $m$ | 1.92 | 0.136 | 0.815-4.52 | 2.04 | 0.087 | 0.902-4.59 |
| Size of pt at diagnosis | 0.86 | 0.647 | 0.443-1.66 | 0.75 | 0.463 | 0.344-1.63 |
| Size of bm at relapse | 1.46 | 0.113 | 0.914-2.34 | 1.81 | 0.0143 | 1.130-2.91 |

HR = Hazard Ratio, CI = Confidence Interval, OS = overall survival, PFS = progression-free survival, bm = brain metastases, pt = primary tumor

# 7 Conclusion

In this thesis, different modeling approaches were formulated to be calibrated to clinical data. These included ordinary and partial differential equation systems and statistical models. The models were analysed in terms of parameter sensitivity, accuracy to explain clinical case data and prognostic value of estimated model parameters.
First, the biological background of lung cancer, the metastatic process and treatment possibilities was introduced and the two evaluated data sets were described. Chapter 3 provided all mathematical tools to construct the deterministic models as well as to investigate their parameters' sensitivities and to perform statistical survival analysis in a Cox proportional hazards model. Chapter 4 set up the models step by step, starting with regular untreated tumoral growth in an ordinary differential equation formulation. This formulation was adjusted to also consider different treatment approaches. The two approaches were integrated into partial differential equations in an untreated and a treated setting to consider metastatic growth and development. Further, a slightly simplified model adjusting for the hypothesis to neglect secondary metastasation in early disease trend was formalised. This model was then used to assess the prognostic possibilities of the whole modeling framework in terms of evaluated parameters' interpretation. For each of the single models an implementation was shown, exploratory simulations were formed and different sensitivity analyses were evaluated to shed light into the model and parameter behaviors. Chapter 6 then successfully simulated the respective models of interest to the corresponding clinical histories of two data sets. The parameter values were extracted individually and for the simplified modeling approach, these were used for a statistical analysis. The concordance indices were improved from 0.647 for fully clinical covariate Cox proportional hazards models to 0.789 for Cox proportional hazards models constructed of the very same clinical parameters plus two computational biomarkers that were generated by the mathematical modeling approach. The increase in predictive power is therefore 22%. This evaluation clearly proves the advantage of considering the model's parameters in clinical routine care to potentially establish even better clinical decisions.
The different modeling approaches can be used to quantitatively determine the metastatic distribution in a patient's body, estimating what has been so far impossible to measure - the invisible metastatic threat early in clinical routine care, the greatest unknown for clinicians, influencing treatment decisions and thus treatment success.

This work with corresponding publications has proven functioning and usability of the different approaches. Also, statistical evaluations have shown the clinical predictive value these approaches already carry. It remains that the approaches are tested in depth on larger data sets to decipher for which patients with certain risk factors the methods are most accurate and for which covariates they are not. Accordingly, these insights give hints at which dynamics the modeling approaches potentially might have to be adjusted and further allow to even better predict disease progression for individual cancer patients.
Extensions of the framework would then allow to chose the most efficient treatment setting on this individual basis to assist clinical decision finding especially in debatable situations.

# List of Figures

# List of Tables

# Bibliography


[AB98] D. G. Altman and J. M. Bland. "Statistics Notes: Time to event (survival) data." In: *BMJ* 317.7156 (Aug. 1998), pp. 468–469. ISSN: 0959-8138, 1468-5833. DOI: 10.1136/bmj.317.7156.468.

[Agr+16] S. Agrawal, Y. Feng, A. Roy, G. Kollia, and B. Lestini. "Nivolumab dose selection: challenges, opportunities, and lessons learned for cancer immunotherapy." In: *Journal for ImmunoTherapy of Cancer* 4.1 (Dec. 2016), p. 72. ISSN: 2051-1426. DOI: 10.1186/s40425-016-0177-2.

[Aka78] A. Akanuma. "Parameter analysis of Gompertzian function growth model in clinical tumors." In: *European Journal of Cancer (1965)* 14.6 (June 1978), pp. 681–688. ISSN: 00142964. DOI: 10.1016/0014-2964(78)90304-3.

[Alm+14] V. Almendro, H. J. Kim, Y.-K. Cheng, M. Gönen, S. Itzkovitz, P. Argani, A. van Oudenaarden, S. Sukumar, F. Michor, and K. Polyak. "Genetic and Phenotypic Diversity in Breast Tumor Metastases." In: *Cancer Research* 74.5 (Mar. 2014), pp. 1338–1348. ISSN: 0008-5472, 1538-7445. DOI: 10.1158/0008-5472.CAN-13-2357-T.

[ALM15] P. M. Altrock, L. L. Liu, and F. Michor. "The mathematics of cancer: integrating quantitative models." In: *Nature Reviews Cancer* 15.12 (Dec. 2015), pp. 730–745. ISSN: 1474-175X, 1474-1768. DOI: 10.1038/nrc4029.

[AM04] R. Araujo and D. McElwain. "A history of the study of solid tumour growth: the contribution of mathematical modelling." In: *Bulletin of Mathematical Biology* 66.5 (Sept. 2004), pp. 1039–1091. ISSN: 00928240. DOI: 10.1016/j.bulm.2003.11.002.

[BA04] J. M. Bland and D. G. Altman. "The logrank test." In: *BMJ* 328.7447 (May 2004), p. 1073. ISSN: 0959-8138, 1468-5833. DOI: 10.1136/bmj.328.7447.1073.

[BA98] J. M. Bland and D. G. Altman. "Statistics Notes: Survival probabilities (the Kaplan-Meier method)." In: *BMJ* 317.7172 (Dec. 1998), pp. 1572–1580. ISSN: 0959-8138, 1468-5833. DOI: 10.1136/bmj.317.7172.1572.

[BC01] F. Brauer and C. Castillo-Chávez. *Mathematical Models in Population Biology and Epidemiology*. Ed. by J. E. Marsden, L. Sirovich, and M. Golubitsky. Vol. 40. Texts in Applied Mathematics. New York, NY: Springer New York, 2001. ISBN: 978-1-4419-3182-5. DOI: 10.1007/978-1-4757-3516-1.

[Ben+14] S. Benzekry, C. Lamont, A. Beheshti, A. Tracz, J. M. L. Ebos, L. Hlatky, and P. Hahnfeldt. "Classical Mathematical Models for Description and Prediction of Experimental Tumor Growth." In: *PLoS Computational Biology* 10.8 (Aug. 2014). Ed. by F. Mac Gabhann, e1003800. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1003800.

[Ben+16] S. Benzekry, A. Tracz, M. Mastri, R. Corbelli, D. Barbolosi, and J. M. Ebos. "Modeling Spontaneous Metastasis following Surgery: An *In Vivo-In Silico* Approach." In: *Cancer Research* 76.3 (Feb. 2016), pp. 535–547. ISSN: 0008-5472, 1538-7445. DOI: 10.1158/0008-5472.CAN-15-1389.

[Ben+17] S. Benzekry, C. Lamont, D. Barbolosi, L. Hlatky, and P. Hahnfeldt. "Mathematical Modeling of Tumor–Tumor Distant Interactions Supports a Systemic Control of Tumor Growth." In: *Cancer Research* 77.18 (Sept. 2017), pp. 5183–5193. ISSN: 0008-5472, 1538-7445. DOI: 10.1158/0008-5472.CAN-17-0564.

[Ben+21] S. Benzekry, C. Sentis, C. Coze, L. Tessonnier, and N. André. "Development and Validation of a Prediction Model of Overall Survival in High-Risk Neuroblastoma Using Mechanistic Modeling of Metastasis." In: *JCO Clinical Cancer Informatics* 5 (Jan. 2021), pp. 81–90. ISSN: 2473-4276. DOI: 10.1200/CCI.20.00092.

[Ben+22] S. Benzekry*, P. Schlicke*, P. Tomasini, and E. Simon. "Mechanistic modeling of brain metastases in NSCLC provides computational markers for personalized prediction of outcome." In: *submitted* (2022).

[Ber+17] D. P. Berger, R. Mertelsmann, J. Duyster, and T. F. CCCF, eds. *Das Rote Buch: Hämatologie und internistische Onkologie*. 6., überarbeitete und erweiterte Auflage. ecomed Medizin. Landsberg am Lech: ecomed Medizin, 2017. ISBN: 978-3-609-51221-1.

[Ber57] L. von Bertalanffy. "Quantitative Laws in Metabolism and Growth." In: *The Quarterly Review of Biology* 32.3 (Sept. 1957), pp. 217–231. ISSN: 0033-5770, 1539-7718. DOI: 10.1086/401873.

[Bet+12] A. Bethge, U. Schumacher, A. Wree, and G. Wedemann. "Are Metastases from Metastases Clinical Relevant? Computer Modelling of Cancer Spread in a Case of Hepatocellular Carcinoma." In: *PLoS ONE* 7.4 (Apr. 2012). Ed. by V. Brusic, e35689. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0035689.

[Bil+19] M. Bilous, C. Serdjebi, A. Boyer, P. Tomasini, C. Pouypoudat, D. Barbolosi, F. Barlesi, F. Chomy, and S. Benzekry. "Quantitative mathematical modeling of clinical brain metastasis dynamics in non-small cell lung cancer." In: *Scientific Reports* 9.1 (Dec. 2019), p. 13018. ISSN: 2045-2322. DOI: 10.1038/s41598-019-49407-3.

[Boe+09] M. de Boer, C. H. van Deurzen, J. A. van Dijck, G. F. Borm, P. J. van Diest, E. M. Adang, J. W. Nortier, E. J. Rutgers, C. Seynaeve, M. B. Menke-Pluymers, P. Bult, and V. C. Tjan-Heijnen. "Micrometastases or Isolated Tumor Cells and the Outcome of Breast Cancer." In: *New England Journal of Medicine* 361.7 (Aug. 2009), pp. 653–663. ISSN: 0028-4793, 1533-4406. DOI: 10.1056/NEJMoa0904832.

[Bor+09] T. Borovski, L. Vermeulen, M. R. Sprick, and J. P. Medema. "One renegade cancer stem cell?" In: *Cell Cycle* 8.6 (Mar. 2009), pp. 803–808. ISSN: 1538-4101, 1551-4005. DOI: 10.4161/cc.8.6.7935.

[Bra+18] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal. "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries." In: *CA: A Cancer Journal for Clinicians* 68.6 (Nov. 2018), pp. 394–424. ISSN: 00079235. DOI: 10.3322/caac.21492.

[CCS07] F. Campolongo, J. Cariboni, and A. Saltelli. "An effective screening design for sensitivity analysis of large models." In: *Environmental Modelling & Software* 22.10 (Oct. 2007), pp. 1509–1518. ISSN: 13648152. DOI: 10.1016/j.envsoft.2006.10.004.

[CG86] A. J. Coldman and J. H. Goldie. "A stochastic model for the origin and treatment of tumors containing drug-resistant cells." In: *Bulletin of Mathematical Biology* 48.3-4 (May 1986), pp. 279–292. ISSN: 0092-8240, 1522-9602. DOI: 10.1007/BF02459682.

[Cla+09] L. Claret, P. Girard, P. M. Hoff, E. Van Cutsem, K. P. Zuideveld, K. Jorga, J. Fagerberg, and R. Bruno. "Model-based prediction of phase III overall survival in colorectal cancer on the basis of phase II tumor dynamics." In: *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology* 27.25 (Sept. 2009), pp. 4103–4108. ISSN: 1527-7755. DOI: 10.1200/JCO.2008.21.0807.

[CN12] E. Comen and L. Norton. "Self-Seeding in Cancer." In: *Minimal Residual Disease and Circulating Tumor Cells in Breast Cancer*. Ed. by M. Ignatiadis, C. Sotiriou, and K. Pantel. Vol. 195. Series Title: Recent Results in Cancer

Research. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 13–23. ISBN: 978-3-642-28160-0. DOI: 10.1007/978-3-642-28160-0_2.

[Cox72] D. R. Cox. "Regression Models and Life-Tables." In: *Journal of the Royal Statistical Society: Series B (Methodological)* 34.2 (Jan. 1972), pp. 187–202. ISSN: 00359246. DOI: 10.1111/j.2517-6161.1972.tb00899.x.

[CSC11] F. Campolongo, A. Saltelli, and J. Cariboni. "From screening to quantitative sensitivity analysis. A unified approach." In: *Computer Physics Communications* 182.4 (Apr. 2011), pp. 978–988. ISSN: 00104655. DOI: 10.1016/j.cpc.2010.12.039.

[CW11] C. L. Chaffer and R. A. Weinberg. "A Perspective on Cancer Cell Metastasis." In: *Science* 331.6024 (Mar. 2011), pp. 1559–1564. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1203543.

[Dem80] R. Demicheli. "Growth of testicular neoplasm lung metastases: Tumor-specific relation between two Gompertzian parameters." In: *European Journal of Cancer (1965)* 16.12 (Dec. 1980), pp. 1603–1608. ISSN: 00142964. DOI: 10.1016/0014-2964(80)90034-1.

[Der+15] B. A. Derman, K. F. Mileham, P. D. Bonomi, M. Batus, and M. J. Fidler. "Treatment of advanced squamous cell carcinoma of the lung: a review." In: *Translational Lung Cancer Research* 4.5 (Oct. 2015), pp. 524–532. ISSN: 2218-6751. DOI: 10.3978/j.issn.2218-6751.2015.06.07.

[DG76] R. P. Dickinson and R. J. Gelinas. "Sensitivity analysis of ordinary differential equation systems—A direct method." In: *Journal of Computational Physics* 21.2 (June 1976), pp. 123–143. ISSN: 00219991. DOI: 10.1016/0021-9991(76)90007-3.

[DP80] J. Dormand and P. Prince. "A family of embedded Runge-Kutta formulae." In: *Journal of Computational and Applied Mathematics* 6.1 (Mar. 1980), pp. 19–26. ISSN: 03770427. DOI: 10.1016/0771-050X(80)90013-3.

[Egg12] A. Eggermont. "Can immuno-oncology offer a truly pan-tumour approach to therapy?" In: *Annals of Oncology* 23 (Sept. 2012), pp. viii53–viii57. ISSN: 09237534. DOI: 10.1093/annonc/mds264.

[End+09] H. Enderling, A. R. Anderson, M. A. Chaplain, A. Beheshti, L. Hlatky, and P. Hahnfeldt. "Paradoxical Dependencies of Tumor Dormancy and Progression on Basic Cell Kinetics." In: *Cancer Research* 69.22 (Nov. 2009), pp. 8814–8821. ISSN: 0008-5472, 1538-7445. DOI: 10.1158/0008-5472.CAN-09-2115.

[ES81] B. Efron and C. Stein. "The Jackknife Estimate of Variance." In: *The Annals of Statistics* 9.3 (May 1981). ISSN: 0090-5364. DOI: 10.1214/aos/1176345462.

[Fen+17] Y. Feng, X. Wang, G. Bajaj, S. Agrawal, A. Bello, B. Lestini, F. G. Finckenstein, J.-S. Park, and A. Roy. "Nivolumab Exposure–Response Analyses of Efficacy and Safety in Previously Treated Squamous or Nonsquamous Non–Small Cell Lung Cancer." In: *Clinical Cancer Research* 23.18 (Sept. 2017), pp. 5394–5405. ISSN: 1078-0432, 1557-3265. DOI: 10.1158/1078-0432.CCR-16-2842.

[Fid03] I. J. Fidler. "The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited." In: *Nature Reviews Cancer* 3.6 (June 2003), pp. 453–458. ISSN: 1474-175X, 1474-1768. DOI: 10.1038/nrc1098.

[Fil+11] P. L. Filosso, E. Ruffini, S. Asioli, R. Giobbe, L. Macri, M. C. Bruna, A. Sandri, and A. Oliaro. "Adenosquamous lung carcinomas: A histologic subtype with poor prognosis." In: *Lung Cancer* 74.1 (Oct. 2011), pp. 25–29. ISSN: 01695002. DOI: 10.1016/j.lungcan.2011.01.030.

[Fru+10] S. Fruehauf, A. Radujkovic, J. Topaly, and W. J. Zeller. "Chemotherapie." In: *Praxis der Viszeralchirurgie Onkologische Chirurgie*. Ed. by J. R. Siewert, M. Rothmund, and V. Schumpelick. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 249–264. ISBN: 978-3-642-03808-2. DOI: 10.1007/978-3-642-03808-2_23.

[Ged79] D. Geddes. "The natural history of lung cancer: A review based on rates of tumour growth." In: *British Journal of Diseases of the Chest* 73 (Jan. 1979), pp. 1–17. ISSN: 00070971. DOI: 10.1016/0007-0971(79)90002-0.

[GM03] R. A. Gatenby and P. K. Maini. "Mathematical oncology: Cancer summed up." In: *Nature* 421.6921 (Jan. 2003), pp. 321–321. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/421321a.

[GM06] G. P. Gupta and J. Massagué. "Cancer Metastasis: Building a Framework." In: *Cell* 127.4 (Nov. 2006), pp. 679–695. ISSN: 00928674. DOI: 10.1016/j.cell.2006.11.001.

[Gom25] B. Gompertz. "On the nature of the function expressive of the law of human mortality, and on the new mode of determining the value of life contingencies." In: *a letter to Francis Baily* 115 (1825), pp. 513–585. DOI: https://doi.org/10.1098/rstl.1825.0026.

[Hae+12] H. Haeno, M. Gonen, M. B. Davis, J. M. Herman, C. A. Iacobuzio-Donahue, and F. Michor. "Computational Modeling of Pancreatic Cancer Reveals Kinetics of Metastasis Suggesting Optimum Treatment Strategies." In: *Cell* 148.1-2 (Jan. 2012), pp. 362–375. ISSN: 00928674. DOI: 10.1016/j.cell.2011.11.060.

[Har82] F. E. Harrell. "Evaluating the Yield of Medical Tests." In: *JAMA: The Journal of the American Medical Association* 247.18 (May 1982), p. 2543. ISSN: 0098-7484. DOI: 10.1001/jama.1982.03320430047030.

[He+15] J. He, Y. Hu, M. Hu, and B. Li. "Development of PD-1/PD-L1 Pathway in Tumor Immune Microenvironment and Treatment for Non-Small Cell Lung Cancer." In: *Scientific Reports* 5.1 (Oct. 2015), p. 13110. ISSN: 2045-2322. DOI: 10.1038/srep13110.

[Hec+01] J. L. Hecht, J. L. Pinkus, L. J. Weinstein, and G. S. Pinkus. "The Value of Thyroid Transcription Factor-1 in Cytologic Preparations as a Marker for Metastatic Adenocarcinoma of Lung Origin." In: *American Journal of Clinical Pathology* 116.4 (Oct. 2001), pp. 483–488. ISSN: 0002-9173, 1943-7722. DOI: 10.1309/NL4Y-FHG8-2XBC-F9XH.

[HLM96] F. E. Harrell, K. L. Lee, and D. B. Mark. "Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors." In: *Statistics in Medicine* 15.4 (Feb. 1996), pp. 361–387. ISSN: 02776715, 10970258. DOI: 10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4.

[HM10] H. Haeno and F. Michor. "The evolution of tumor metastases during clonal expansion." In: *Journal of Theoretical Biology* 263.1 (Mar. 2010), pp. 30–44. ISSN: 00225193. DOI: 10.1016/j.jtbi.2009.11.005.

[Hoe48] W. Hoeffding. "A Class of Statistics with Asymptotically Normal Distribution." In: *The Annals of Mathematical Statistics* 19.3 (Sept. 1948), pp. 293–325. ISSN: 0003-4851. DOI: 10.1214/aoms/1177730196.

[HS96] T. Homma and A. Saltelli. "Importance measures in global sensitivity analysis of nonlinear models." In: *Reliability Engineering & System Safety* 52.1 (Apr. 1996), pp. 1–17. ISSN: 09518320. DOI: 10.1016/0951-8320(96)00002-6.

[HTC20] R. Huber, Tumorzentrum München, and Comprehensive Cancer Center München, eds. *Tumoren der Lunge und des Mediastinums: Empfehlungen zur Diagnostik, Therapie und Nachsorge*. 12. überarbeitete Auflage. Manual / Tumorzentrum München an den Medizinischen Fakultäten der Ludwig-Maximilians-Universität und der Technischen Universität. München: Zuckschwerdt Verlag, 2020. ISBN: 978-3-86371-330-0.

[HW11] D. Hanahan and R. A. Weinberg. "Hallmarks of Cancer: The Next Generation." In: *Cell* 144.5 (Mar. 2011), pp. 646–674. ISSN: 00928674. DOI: 10.1016/j.cell.2011.02.013.

[IKS00] K. Iwata, K. Kawasaki, and N. Shigesada. "A Dynamical Model for the Growth and Size Distribution of Multiple Metastatic Tumors." In: *Journal of Theoretical Biology* 203.2 (Mar. 2000), pp. 177–186. ISSN: 00225193. DOI: 10.1006/jtbi.2000.1075.

[IL15] B. Iooss and P. Lemaître. "A Review on Global Sensitivity Analysis Methods." In: *Uncertainty Management in Simulation-Optimization of Complex Systems*. Ed. by G. Dellino and C. Meloni. Vol. 59. Series Title: Operations Research/Computer Science Interfaces Series. Boston, MA: Springer US, 2015, pp. 101–122. ISBN: 978-1-4899-7547-8. DOI: 10.1007/978-1-4899-7547-8_5.

[Ish+92] T. Ishida, S. Kaneko, H. Yokoyama, T. Inoue, K. Sugio, and K. Sugimachi. "Adenosquamous Carcinoma of the Lung: *Clinicopathologic and Immunohistochemical Features*." In: *American Journal of Clinical Pathology* 97.5 (May 1992), pp. 678–685. ISSN: 1943-7722, 0002-9173. DOI: 10.1093/ajcp/97.5.678.

[Iwa+17] Y. Iwai, J. Hamanishi, K. Chamoto, and T. Honjo. "Cancer immunotherapies targeting the PD-1 signaling pathway." In: *Journal of Biomedical Science* 24.1 (Dec. 2017), p. 26. ISSN: 1423-0127. DOI: 10.1186/s12929-017-0329-9.

[Jai13] R. K. Jain. "Normalizing Tumor Microenvironment to Treat Cancer: Bench to Bedside to Biomarkers." In: *Journal of Clinical Oncology* 31.17 (June 2013), pp. 2205–2218. ISSN: 0732-183X, 1527-7755. DOI: 10.1200/JCO.2012.46.3653.

[JRL02] R. W. Johnstone, A. A. Ruefli, and S. W. Lowe. "Apoptosis." In: *Cell* 108.2 (Jan. 2002), pp. 153–164. ISSN: 00928674. DOI: 10.1016/S0092-8674(02)00625-6.

[Kea16] G. M. Keating. "Nivolumab: A Review in Advanced Nonsquamous Non-Small Cell Lung Cancer." In: *Drugs* 76.9 (June 2016), pp. 969–978. ISSN: 0012-6667, 1179-1950. DOI: 10.1007/s40265-016-0589-9.

[Ken+08] S. A. Kenfield, E. K. Wei, M. J. Stampfer, B. A. Rosner, and G. A. Colditz. "Comparison of aspects of smoking among the four histological types of lung cancer." In: *Tobacco Control* 17.3 (Feb. 2008), pp. 198–204. ISSN: 0964-4563. DOI: 10.1136/tc.2007.022582.

[Kle09] C. A. Klein. "Parallel progression of primary tumours and metastases." In: *Nature Reviews Cancer* 9.4 (Apr. 2009), pp. 302–312. ISSN: 1474-175X, 1474-1768. DOI: 10.1038/nrc2627.

[KM58] E. L. Kaplan and P. Meier. "Nonparametric Estimation from Incomplete Observations." In: *Journal of the American Statistical Association* 53.282 (June 1958), pp. 457–481. ISSN: 0162-1459, 1537-274X. DOI: 10.1080/01621459.1958.10501452.

[Koc16] R. Koch-Institut. "Bericht zum Krebsgeschehen in Deutschland 2016." In: (2016). Publisher: RKI-Bib1 (Robert Koch-Institut). DOI: 10.17886/RKIPUBL-2016-014.

[Koy+08] K. Koyanagi, A. J. Bilchik, S. Saha, R. R. Turner, D. Wiese, M. McCarter, P. Shen, L. Deacon, D. Elashoff, and D. S. Hoon. "Prognostic Relevance of Occult Nodal Micrometastases and Circulating Tumor Cells in Colorectal Cancer in a Prospective Multicenter Trial." In: *Clinical Cancer Research* 14.22 (Nov. 2008), pp. 7391–7396. ISSN: 1078-0432, 1557-3265. DOI: 10.1158/1078-0432.CCR-08-0290.

[KT85] H. Krug and G. Taubert. "Practical application of logistic function to the growth of experimental tumors." In: *Archiv Für Geschwulstforschung* 55.4 (1985), pp. 235–244. ISSN: 0003-911X.

[Kuf+03] D. W. Kufe, J. F. Holland, E. Frei, and A. C. Society, eds. *Cancer medicine 6*. 6th ed. Hamilton, Ont. ; Lewiston, NY: BC Decker, 2003. ISBN: 978-1-55009-213-4.

[Lai64] A. K. Laird. "Dynamics of Tumor Growth." In: *British Journal of Cancer* 18.3 (Sept. 1964), pp. 490–502. ISSN: 0007-0920, 1532-1827. DOI: 10.1038/bjc.1964.55.

[Lai65] A. K. Laird. "Dynamics of Tumour Growth: Comparison of Growth Rates and Extrapolation of Growth Curve to One Cell." In: *British Journal of Cancer* 19.2 (June 1965), pp. 278–291. ISSN: 0007-0920, 1532-1827. DOI: 10.1038/bjc.1965.32.

[Lee+16] T. Lee, B. Lee, Y.-L. Choi, J. Han, M.-J. Ahn, and S.-W. Um. "Non-small Cell Lung Cancer with Concomitant *EGFR*, *KRAS*, and *ALK* Mutation: Clinicopathologic Features of 12 Cases." In: *Journal of Pathology and Translational Medicine* 50.3 (May 2016), pp. 197–203. ISSN: 2383-7837, 2383-7845. DOI: 10.4132/jptm.2016.03.09.

[LPW17] A. W. Lambert, D. R. Pattabiraman, and R. A. Weinberg. "Emerging Biological Principles of Metastasis." In: *Cell* 168.4 (Feb. 2017), pp. 670–691. ISSN: 00928674. DOI: 10.1016/j.cell.2016.11.037.

[Luz+98] K. J. Luzzi, I. C. MacDonald, E. E. Schmidt, N. Kerkvliet, V. L. Morris, A. F. Chambers, and A. C. Groom. "Multistep Nature of Metastatic Inefficiency." In: *The American Journal of Pathology* 153.3 (Sept. 1998), pp. 865–873. ISSN: 00029440. DOI: 10.1016/S0002-9440(10)65628-3.

[Lyn+04] T. J. Lynch, D. W. Bell, R. Sordella, S. Gurubhagavatula, R. A. Okimoto, B. W. Brannigan, P. L. Harris, S. M. Haserlat, J. G. Supko, F. G. Haluska, D. N. Louis, D. C. Christiani, J. Settleman, and D. A. Haber. "Activating Mutations in the Epidermal Growth Factor Receptor Underlying Responsiveness of Non–Small-Cell Lung Cancer to Gefitinib." In: *New England Journal of Medicine* 350.21 (May 2004), pp. 2129–2139. ISSN: 0028-4793, 1533-4406. DOI: 10.1056/NEJMoa040938.

[Mak+18] R. H. Mak, G. Hermann, H. J. Aerts, E. H. Baldini, A. B. Chen, D. Kozono, M. S. Rabin, S. J. Swanson, Y.-H. Chen, P. Catalano, B. E. Johnson, and P. A. Jänne. "Outcomes by *EGFR* , *KRAS* , and *ALK* Genotype After Combined Modality Therapy for Locally Advanced Non–Small-Cell Lung Cancer." In: *JCO Precision Oncology* 2 (Nov. 2018), pp. 1–18. ISSN: 2473-4284. DOI: 10.1200/PO.17.00219.

[May32] W. V. Mayneord. "On a Law of Growth of Jensen's Rat Sarcoma." In: *The American Journal of Cancer* 16.4 (July 1932), p. 841. DOI: 10.1158/ajc.1932.841.

[MB93] M. Marusic and Z. Bajzer. "Generalized Two-Parameter Equation of Growth." In: *Journal of Mathematical Analysis and Applications* 179.2 (Nov. 1993), pp. 446–462. ISSN: 0022247X. DOI: 10.1006/jmaa.1993.1361.

[Meh15] T. Mehrling. "Chemotherapy is getting 'smarter'." In: *Future Oncology* 11.4 (Feb. 2015), pp. 549–552. ISSN: 1479-6694, 1744-8301. DOI: 10.2217/fon.14.248.

[Mic+11] F. Michor, J. Liphardt, M. Ferrari, and J. Widom. "What does physics have to do with cancer?" In: *Nature Reviews Cancer* 11.9 (Sept. 2011), pp. 657–670. ISSN: 1474-175X, 1474-1768. DOI: 10.1038/nrc3092.

[Mor91] M. D. Morris. "Factorial Sampling Plans for Preliminary Computational Experiments." In: *Technometrics* 33.2 (May 1991), pp. 161–174. ISSN: 0040-1706, 1537-2723. DOI: 10.1080/00401706.1991.10484804.

[NBM09] D. X. Nguyen, P. D. Bos, and J. Massagué. "Metastasis: from dissemination to organ-specific colonization." In: *Nature Reviews Cancer* 9.4 (Apr. 2009), pp. 274–284. ISSN: 1474-175X, 1474-1768. DOI: 10.1038/nrc2622.

[New+12] P. K. Newton, J. Mason, K. Bethel, L. A. Bazhenova, J. Nieva, and P. Kuhn. "A Stochastic Markov Chain Model to Describe Lung Cancer Growth and Metastasis." In: *PLoS ONE* 7.4 (Apr. 2012). Ed. by B. Ermentrout, e34637. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0034637.

[Nor+76] L. Norton, R. Simon, H. D. Brereton, and A. E. Bogden. "Predicting the course of Gompertzian growth." In: *Nature* 264.5586 (Dec. 1976), pp. 542–545. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/264542a0.

[Nor88] L. Norton. "A Gompertzian model of human breast cancer growth." In: *Cancer Research* 48.24 Pt 1 (Dec. 1988), pp. 7067–7071. ISSN: 0008-5472.

[NS86] L. Norton and R. Simon. "The Norton-Simon hypothesis revisited." In: *Cancer Treatment Reports* 70.1 (Jan. 1986), pp. 163–169. ISSN: 0361-5960.

[Oh+09] Y. Oh, S. Taylor, B. N. Bekele, J. M. Debnam, P. K. Allen, D. Suki, R. Sawaya, R. Komaki, D. J. Stewart, and D. D. Karp. "Number of metastatic sites is a strong predictor of survival in patients with nonsmall cell lung cancer with or without brain metastases." In: *Cancer* 115.13 (July 2009), pp. 2930–2938. ISSN: 0008543X, 10970142. DOI: 10.1002/cncr.24333.

[Par12] D. M. Pardoll. "The blockade of immune checkpoints in cancer immunotherapy." In: *Nature Reviews Cancer* 12.4 (Apr. 2012), pp. 252–264. ISSN: 1474-175X, 1474-1768. DOI: 10.1038/nrc3239.

[PCF99] K. Pantel, R. J. Cote, and O. Fodstad. "Detection and Clinical Importance of Micrometastatic Disease." In: *JNCI Journal of the National Cancer Institute* 91.13 (July 1999), pp. 1113–1124. ISSN: 0027-8874, 1460-2105. DOI: 10.1093/jnci/91.13.1113.

[Pea05] K. Pearson. *On the general theory of skew correlation and non-linear regression*. London: Dulau and co., 1905.

[PK12] A. Pathak and S. Kumar. "Independent regulation of tumor cell migration by matrix stiffness and confinement." In: *Proceedings of the National Academy of Sciences* 109.26 (June 2012), pp. 10334–10339. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1118073109.

[PK15] S. P. Patel and R. Kurzrock. "PD-L1 Expression as a Predictive Biomarker in Cancer Immunotherapy." In: *Molecular Cancer Therapeutics* 14.4 (Apr. 2015), pp. 847–856. ISSN: 1535-7163, 1538-8514. DOI: 10.1158/1535-7163.MCT-14-0983.

[QP13] S. A. Quezada and K. S. Peggs. "Exploiting CTLA-4, PD-1 and PD-L1 to reactivate the host immune response against cancer." In: *British Journal of Cancer* 108.8 (Apr. 2013), pp. 1560–1565. ISSN: 0007-0920, 1532-1827. DOI: 10.1038/bjc.2013.117.

[RCM15] R. Richard, J. Casas, and E. McCauley. "Sensitivity analysis of continuous-time models for ecological and evolutionary theories." In: *Theoretical Ecology* 8.4 (Nov. 2015), pp. 481–490. ISSN: 1874-1738, 1874-1746. DOI: 10.1007/s12080-015-0265-9.

[Rei+17] M. Reiser, F.-P. Kuhn, J. Debus, P. Bartenstein, and H. Holtermann, eds. *Radiologie*. 4., vollständig überarbeitete Auflage. Duale Reihe. Stuttgart: Thieme, 2017. ISBN: 978-3-13-125324-8. DOI: 10.1055/b-004-132212.

[Rib12] A. Ribas. "Tumor Immunotherapy Directed at PD-1." In: *New England Journal of Medicine* 366.26 (June 2012), pp. 2517–2519. ISSN: 0028-4793, 1533-4406. DOI: 10.1056/NEJMe1205943.

[RK15] A. Rounds and J. Kolesar. "Nivolumab for second-line treatment of metastatic squamous non-small-cell lung cancer." In: *American Journal of Health-System Pharmacy* 72.21 (Nov. 2015), pp. 1851–1855. ISSN: 1079-2082, 1535-2900. DOI: 10.2146/ajhp150235.

[RN03] S. J. Russell and P. Norvig. *Artificial intelligence: a modern approach; the intelligent agent book*. 2. ed., internat. ed. Prentice Hall series in artificial intelligence. Upper Saddle River, NJ: Prentice Hall, 2003. ISBN: 978-0-13-080302-3.

[Roz+12] E. N. Rozali, S. V. Hato, B. W. Robinson, R. A. Lake, and W. J. Lesterhuis. "Programmed Death Ligand 2 in Cancer-Induced Immune Suppression." In: *Clinical and Developmental Immunology* 2012 (2012), pp. 1–8. ISSN: 1740-2522, 1740-2530. DOI: 10.1155/2012/656340.

[Sal02] A. Saltelli. "Making best use of model evaluations to compute sensitivity indices." In: *Computer Physics Communications* 145.2 (May 2002), pp. 280–297. ISSN: 00104655. DOI: 10.1016/S0010-4655(02)00280-1.

[Sal04] A. Saltelli, ed. *Sensitivity analysis in practice: a guide to assessing scientific models*. Hoboken, NJ: Wiley, 2004. ISBN: 978-0-470-87093-8.

[Sal08] A. Saltelli, ed. *Global sensitivity analysis: the primer*. OCLC: ocn180852094. Chichester, England ; Hoboken, NJ: John Wiley, 2008. ISBN: 978-0-470-05997-5.

[Sav79] M. A. Savageau. "Allometric morphogenesis of complex systems: Derivation of the basic equations from first principles." In: *Proceedings of the National Academy of Sciences* 76.12 (Dec. 1979), pp. 6023–6025. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.76.12.6023.

[Sch+22] S. Schönfeld, A. Ozkan, L. Scarabosio, M. N. Rylander, and C. Kuttler. "Environmental stress level to model tumor cell growth and survival." In: *arXiv:2201.06985 [cs, math, q-bio]* (Jan. 2022). arXiv: 2201.06985. To appear in *Mathematical Biosciences and Engineering*.

[Sch61] M. Schwartz. "A biomathematical approach to clinical tumor growth." In: *Cancer* 14.6 (Nov. 1961), pp. 1272–1294. ISSN: 0008-543X, 1097-0142. DOI: 10.1002/1097-0142(196111/12)14:6<1272::AID-CNCR2820140618>3.0.CO;2-H.

[SFR16] A. Steven, S. A. Fisher, and B. W. Robinson. "Immunotherapy for lung cancer: Immunotherapy for lung cancer." In: *Respirology* 21.5 (July 2016), pp. 821–833. ISSN: 13237799. DOI: 10.1111/resp.12789.

[Ski86] H. E. Skipper. "Laboratory models: some historical perspective." In: *Cancer Treatment Reports* 70.1 (Jan. 1986), pp. 3–7. ISSN: 0361-5960.

[SKS21] P. Schlicke, C. Kuttler, and C. Schumann. "How mathematical modeling could contribute to the quantification of metastatic tumor burden under therapy: insights in immunotherapeutic treatment of non-small cell lung cancer." In: *Theoretical Biology and Medical Modelling* 18.1 (Dec. 2021), p. 11. ISSN: 1742-4682. DOI: 10.1186/s12976-021-00142-1.

[Sle+90] R. J. Slebos, R. E. Kibbelaar, O. Dalesio, A. Kooistra, J. Stam, C. J. Meijer, S. S. Wagenaar, R. G. Vanderschueren, N. van Zandwijk, W. J. Mooi, J. L. Bos, and S. Rodenhuis. "K- *ras* Oncogene Activation as a Prognostic Marker in Adenocarcinoma of the Lung." In: *New England Journal of Medicine* 323.9 (Aug. 1990), pp. 561–565. ISSN: 0028-4793, 1533-4406. DOI: 10.1056/NEJM199008303230902.

[SLK76] G. M. Saidel, L. A. Liotta, and J. Kleinerman. "System dynamics of a metastatic process from an implanted tumor." In: *Journal of Theoretical Biology* 56.2 (Feb. 1976), pp. 417–434. ISSN: 00225193. DOI: 10.1016/S0022-5193(76)80083-5.

[SMS95] J. S. Spratt, J. S. Meyer, and J. A. Spratt. "Rates of growth of human solid neoplasms: Part I." In: *Journal of Surgical Oncology* 60.2 (Oct. 1995), pp. 137–146. ISSN: 00224790, 10969098. DOI: 10.1002/jso.2930600216.

[SN06] R. Simon and L. Norton. "The Norton–Simon hypothesis: designing more effective and less toxic chemotherapeutic regimens." In: *Nature Clinical Practice Oncology* 3.8 (Aug. 2006), pp. 406–407. ISSN: 1743-4254, 1743-4262. DOI: 10.1038/ncponc0560.

[Sob+07] I. Sobol', S. Tarantola, D. Gatelli, S. Kucherenko, and W. Mauntz. "Estimating the approximation error when fixing unessential factors in global sensitivity analysis." In: *Reliability Engineering & System Safety* 92.7 (July 2007), pp. 957–960. ISSN: 09518320. DOI: 10.1016/j.ress.2006.07.001.

[Sob01] I. Sobol'. "Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates." In: *Mathematics and Computers in Simulation* 55.1-3 (Feb. 2001), pp. 271–280. ISSN: 03784754. DOI: 10.1016/S0378-4754(00)00270-6.

[Sob93] I. Sobol'. "Sensitivity analysis for non-linear mathematical models." In: *Mathematical Modelling and Computational Experiment* 1 (1993), pp. 407–414.

[Sot+10] A. Sottoriva, J. J. Verhoeff, T. Borovski, S. K. McWeeney, L. Naumov, J. P. Medema, P. M. Sloot, and L. Vermeulen. "Cancer Stem Cell Tumor Model Reveals Invasive Morphology and Increased Phenotypical Heterogeneity." In: *Cancer Research* 70.1 (Jan. 2010), pp. 46–56. ISSN: 0008-5472, 1538-7445. DOI: 10.1158/0008-5472.CAN-09-3663.

[SR97] L. F. Shampine and M. W. Reichelt. "The MATLAB ODE Suite." In: *SIAM Journal on Scientific Computing* 18.1 (Jan. 1997), pp. 1–22. ISSN: 1064-8275, 1095-7197. DOI: 10.1137/S1064827594276424.

[SS72] P. W. Sullivan and S. E. Salmon. "Kinetics of tumor growth and regression in IgG multiple myeloma." In: *Journal of Clinical Investigation* 51.7 (July 1972), pp. 1697–1708. ISSN: 0021-9738. DOI: 10.1172/JCI106971.

[SS76] J. S. Spratt and J. A. Spratt. "The prognostic value of measuring the gross linear radial growth of pulmonary metastases and primary pulmonary cancers." In: *The Journal of Thoracic and Cardiovascular Surgery* 71.2 (Feb. 1976), pp. 274–278. ISSN: 0022-5223.

[SSW64] H. E. Skipper, F. M. Schabel, and W. S. Wilcox. "On the criteria and kinetics associated with 'curability' of experimental leukemia." In: *Cancer Chemotherapy Reports* 35 (Feb. 1964), pp. 1–111. ISSN: 0069-0112.

[Ste16] P. S. Steeg. "Targeting metastasis." In: *Nature Reviews Cancer* 16.4 (Apr. 2016), pp. 201–218. ISSN: 1474-175X, 1474-1768. DOI: 10.1038/nrc.2016.25.

[Suz+05] A. Suzuki, N. Shijubo, G. Yamada, S. Ichimiya, M. Satoh, S. Abe, and N. Sato. "Napsin A is useful to distinguish primary lung adenocarcinoma from adenocarcinomas of other organs." In: *Pathology - Research and Practice* 201.8-9 (Oct. 2005), pp. 579–586. ISSN: 03440338. DOI: 10.1016/j.prp.2005.05.010.

[SWZ16] M. Schmid, M. N. Wright, and A. Ziegler. "On the use of Harrell's C for clinical risk prediction via random survival forests." In: *Expert Systems with Applications* 63 (Nov. 2016), pp. 450–459. ISSN: 09574174. DOI: 10.1016/j.eswa.2016.07.018.

[TD15] A. Talkington and R. Durrett. "Estimating Tumor Growth Rates In Vivo." In: *Bulletin of Mathematical Biology* 77.10 (Oct. 2015), pp. 1934–1954. ISSN: 0092-8240, 1522-9602. DOI: 10.1007/s11538-015-0110-8.

[TDP15] S. L. Topalian, C. G. Drake, and D. M. Pardoll. "Immune Checkpoint Blockade: A Common Denominator Approach to Cancer Therapy." In: *Cancer Cell* 27.4 (Apr. 2015), pp. 450–461. ISSN: 15356108. DOI: 10.1016/j.ccell.2015.03.001.

[TG55] R. H. Thomlinson and L. H. Gray. "The Histological Structure of Some Human Lung Cancers and the Possible Implications for Radiotherapy." In: *British Journal of Cancer* 9.4 (Dec. 1955), pp. 539–549. ISSN: 0007-0920, 1532-1827. DOI: 10.1038/bjc.1955.55.

[Tra+11] W. D. Travis, E. Brambilla, M. Noguchi, A. G. Nicholson, K. R. Geisinger, Y. Yatabe, D. G. Beer, C. A. Powell, G. J. Riely, P. E. Van Schil, K. Garg, J. H. Austin, H. Asamura, V. W. Rusch, F. R. Hirsch, G. Scagliotti, T. Mitsudomi, R. M. Huber, Y. Ishikawa, J. Jett, M. Sanchez-Cespedes, J.-P. Sculier, T. Takahashi, M. Tsuboi, J. Vansteenkiste, I. Wistuba, P.-C. Yang, D. Aberle, C. Brambilla, D. Flieder, W. Franklin, A. Gazdar, M. Gould, P. Hasleton, D. Henderson, B. Johnson, D. Johnson, K. Kerr, K. Kuriyama, J. S. Lee, V. A. Miller, I. Petersen, V. Roggli, R. Rosell, N. Saijo, E. Thunnissen, M. Tsao, and D. Yankelewitz. "International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society International Multidisciplinary Classification of Lung Adenocarcinoma." In: *Journal of Thoracic Oncology* 6.2 (Feb. 2011), pp. 244–285. ISSN: 15560864. DOI: 10.1097/JTO.0b013e318206a221.

[Tra+15] W. D. Travis, E. Brambilla, A. G. Nicholson, Y. Yatabe, J. H. Austin, M. B. Beasley, L. R. Chirieac, S. Dacic, E. Duhig, D. B. Flieder, K. Geisinger, F. R. Hirsch, Y. Ishikawa, K. M. Kerr, M. Noguchi, G. Pelosi, C. A. Powell, M. S. Tsao, and I. Wistuba. "The 2015 World Health Organization Classification

of Lung Tumors." In: *Journal of Thoracic Oncology* 10.9 (Sept. 2015), pp. 1243–1260. ISSN: 15560864. DOI: 10.1097/JTO.0000000000000630.

[TV72] R. Tomović and M. Vukobratović. *General sensitivity theory*. Modern analytic and computational methods in science and mathematics no. 35. New York: American Elsevier Pub. Co, 1972. ISBN: 978-0-444-00108-5.

[Ugr+07] Z. Ugray, L. Lasdon, J. Plummer, F. Glover, J. Kelly, and R. Martí. "Scatter Search and Local NLP Solvers: A Multistart Framework for Global Optimization." In: *INFORMS Journal on Computing* 19.3 (Aug. 2007), pp. 328–340. ISSN: 1091-9856, 1526-5528. DOI: 10.1287/ijoc.1060.0175.

[VA82] V. G. Vaidya and F. J. Alexandro. "Evaluation of some mathematical models for tumor growth." In: *International Journal of Bio-Medical Computing* 13.1 (Jan. 1982), pp. 19–35. ISSN: 00207101. DOI: 10.1016/0020-7101(82)90048-4.

[Ver+12] G. Veronesi, P. Maisonneuve, M. Bellomi, C. Rampinelli, I. Durli, R. Bertolotti, and L. Spaggiari. "Estimating Overdiagnosis in Low-Dose Computed Tomography Screening for Lung Cancer: A Cohort Study." In: *Annals of Internal Medicine* 157.11 (Dec. 2012), p. 776. ISSN: 0003-4819. DOI: 10.7326/0003-4819-157-11-201212040-00005.

[Ver45] P. Verhulst. "Recherches mathématiques sur la loi d'accroissement de la population." In: 18 (1845), pp. 1–42.

[Wak+07] H. A. Wakelee, E. T. Chang, S. L. Gomez, T. H. Keegan, D. Feskanich, C. A. Clarke, L. Holmberg, L. C. Yong, L. N. Kolonel, M. K. Gould, and D. W. West. "Lung Cancer Incidence in Never Smokers." In: *Journal of Clinical Oncology* 25.5 (Feb. 2007), pp. 472–478. ISSN: 0732-183X, 1527-7755. DOI: 10.1200/JCO.2006.07.2983.

[Wal93] W. Walter. *Gewöhnliche Differentialgleichungen: eine Einführung*. 5., überarb. u. erg. Aufl. Springer-Lehrbuch. Berlin u.a: Springer, 1993. ISBN: 978-3-540-56294-8.

[Wei06] R. A. Weinberg. *The Biology of Cancer*. 2nd ed. W.W. Norton & Company, June 2006. ISBN: 978-1-136-97738-1. DOI: 10.1201/9780203852569.

[Xu87] X. Xu. "The biological foundation of the Gompertz model." In: *International Journal of Bio-Medical Computing* 20.1-2 (Jan. 1987), pp. 35–39. ISSN: 00207101. DOI: 10.1016/0020-7101(87)90012-2.

[Yoo+18] S. H. Yoo, B. Keam, M. Kim, S. H. Kim, Y. J. Kim, T. M. Kim, D.-W. Kim, J. S. Lee, and D. S. Heo. "Low-dose nivolumab can be effective in non-small cell lung cancer: alternative option for financial toxicity." In: *ESMO Open* 3.5 (2018), e000332. ISSN: 20597029. DOI: 10.1136/esmoopen-2018-000332.

[Zau89] E. Zauderer. *Partial differential equations of applied mathematics*. 2. ed. Pure and applied mathematics. New York: Wiley, 1989. ISBN: 0-471-53298-3.

[ZLB07a] A. Ziegler, S. Lange, and R. Bender. "Überlebenszeitanalyse: Der Log-Rang-Test." In: *DMW - Deutsche Medizinische Wochenschrift* 132.S 01 (2007), e39–e41. ISSN: 0012-0472, 1439-4413. DOI: 10.1055/s-2007-959040.

[ZLB07b] A. Ziegler, S. Lange, and R. Bender. "Überlebenszeitanalyse: Die Cox-Regression." In: *DMW - Deutsche Medizinische Wochenschrift* 132.S 01 (2007), e42–e44. ISSN: 0012-0472, 1439-4413. DOI: 10.1055/s-2007-959039.

[ZLB07c] A. Ziegler, S. Lange, and R. Bender. "Überlebenszeitanalyse: Eigenschaften und Kaplan-Meier Methode." In: *DMW - Deutsche Medizinische Wochenschrift* 132.S 01 (2007), e36–e38. ISSN: 0012-0472, 1439-4413. DOI: 10.1055/s-2007-959038.