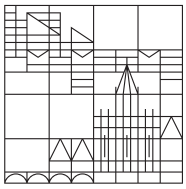
Modeling of Pharmacokinetics and Pharmacodynamics with Application to Cancer and Arthritis

Dissertation zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften

vorgelegt von
Gilbert Koch

an der

Universität Konstanz



Mathematisch-Naturwissenschaftliche Sektion
Fachbereich Mathematik und Statistik

Tag der mündlichen Prüfung: 25. Mai 2012
1. Referent: Prof. Dr. Johannes Schropp
2. Referent: Prof. Dr. Michael Junk

Acknowledgements

First and foremost, I am deeply grateful to my supervisor Prof. Dr. Johannes Schropp for his guidance and support throughout my entire studies. His ideas and advice were essential in finshing this project.

Special thanks go to Dr. Antje Walz for starting the cooperation between Nycomed (A Takeda Company) and the University of Konstanz, and also for her patience in introducing me to biological and pharmacological principles. In particular, I would like to thank Dr. Gezim Lahu from Nycomed for constant financial support of this project. Also special thanks go to Dr. Thomas Wagner (Nycomed) for excellent team work. Finally, I thank Dr. Christine Plater-Zyberk from Micromet and several staff members from Nycomed for their valuable input during the last years.

Abstract

Mathematical modeling of pharmacokinetics / pharmacodynamics (PKPD) is an important and growing field in drug development. In this work we develop preclinical PKPD models based on fundamental biological and pharmacological principles.

Equipped with a PKPD model, different dosing schedules could be simulated and therefore, a valuable contribution to first in human dose selection could be achieved.

We consider different mathematical model figures and discuss the properties and biological basis. Such tools serve as modules for a final PKPD model. We apply ordinary and delay differential equations and especially focus on modeling of delays and lifespans in populations. We show a fundamental relationship between transit compartments and lifespan models. Moreover, we investigate the weighted least squares estimator and derive statistical characteristics of model parameter.

We present a PKPD model to describe tumor growth and anticancer effects for mono- and combination therapy. Further, we construct a PKPD model for arthritis development and antibody effects.

Summarizing, we develop (semi)-mechanistic mathematical PKPD models based on pharmacological assumptions and apply our models to measured data from preclinical phase.

Zusammenfassung

Ein wichtiges und wachsendes Gebiet in der Medikamentenentwicklung ist die mathematische Modellierung der Pharmakokinetik / Pharmakodynamik (PKPD). In der vorliegenden Arbeit entwickeln wir präklinische PKPD Modelle basierend auf grundlegenden biologischen und pharmakologischen Prinzipien.

Mit einem PKPD Modell können verschiedenste Dosierungen simuliert werden und somit ein wertvoller Betrag bei der Suche nach einer Dosis für den Menschen geleistet werden.

Wir betrachten verschiedene mathematische Modelltypen die als Bausteine für ein finales PKPD Modell dienen und diskutieren deren Eigenschaften sowie die biologische Basis. Es werden gewöhnliche und verzögerte Differentialgleichungen verwendet mit einer speziellen Ausrichtung auf die Modellierung von Verzögerungen sowie der Lebensdauer von Objekten in Populationen. Die Arbeit beinhaltet ein grundlegendes Ergebnis über die Beziehung zwischen Transit Kompartimenten und Modellen mit Lebensdauern. Des weiteren zeigen wir Eigenschaften des gewichteten Kleinsten-Quadrate-Schätzers und leiten statistische Kenngrößen für Modellparameter her.

Wir präsentieren ein PKPD Modell für das Wachstum von Tumoren und den Effekt von Krebsmedikamenten in der Mono- und Kombinationstherapie. Außerdem konstruieren wir ein PKPD Modell für die Entwicklung von Arthritis und für den Effekt eines Antikörpers auf die Krankheit.

Zusammengefasst werden in dieser Arbeit (semi-)mechanistische mathematische PKPD Modelle basierend auf pharmakologischen Annahmen entwickelt und auf präklinische Daten angewendet.

Contents

Chapter 1

Introduction

The development of drugs is time-consuming and costly. A study [DHG03] from 2003 reports costs of approximately US$ 800 million to bring a drug to the market. It is estimated that around 90 percent of compounds (drug candidates) will fail during the drug development process. Hence, the drug-producing industry is in search of new tools to support drug development. It is stated by the U.S. Food and Drug Administration (FDA) that computational modeling and simulation is a useful tool to improve the efficiency in developing safe and effective drugs, see [GM01].

The complete process of drug development consists of a preclinical and clinical part. In preclinics, different compounds are tested for an effect in animals. The clinical part is divided into three phases. In phase I, the drug is tested in healthy humans for physiological compatibility. In phase II, the pharmacological / therapeutic effect is investigated. In phase III, the drug is tested in thousands of patients. In Figure 1.1 the time course of drug development is schematically visualized.

An experiment in drug development consists of two parts. The pharmacokinetics (PK) describes the time course of drugs. The pharmacodynamics (PD) is the study of the pharmacological effect of drugs. It is believed *"that by better understanding of the relationship between PK and PD one can shed light on situations where one or the other needs to be optimized in drug discovery and development"*, see [VDGG09].

In this work we develop mathematical pharmacokinetic / pharmacodynamic models based on preclinical experiments. Such models are used to describe measurements, to categorize the pharmacological effect of different compounds, to simulate different dosing schedules (e.g. for first-in-human dose selection) and also to understand underlying mechanisms of disease and drug response. Hence, mathematics has an important impact on drug development and it is commonly believed that the role of mathematical modeling will further increase, see [KD03]. However, it is self-evident that PKPD models have to be based on fundamental biological and pharmacological principles and therefore, the development of such models is in general performed in an interdisciplinary collaboration.

In Chapter 2 we give a brief introduction to pharmacological terms and differential equations. In Chapter 3 typical pharmacokinetic compartment models are introduced. In the next Chapter 4 we derive several models based on biological and pharmacological principles and present theoretical mathematical results. Chapter 5 deals with statistical properties of model parameter estimates. Finally, we develop a PKPD model for tumor growth for mono- and combination therapy in Chapter 6 and also derive a PKPD model for arthritis development with antibody effects in Chapter 7.

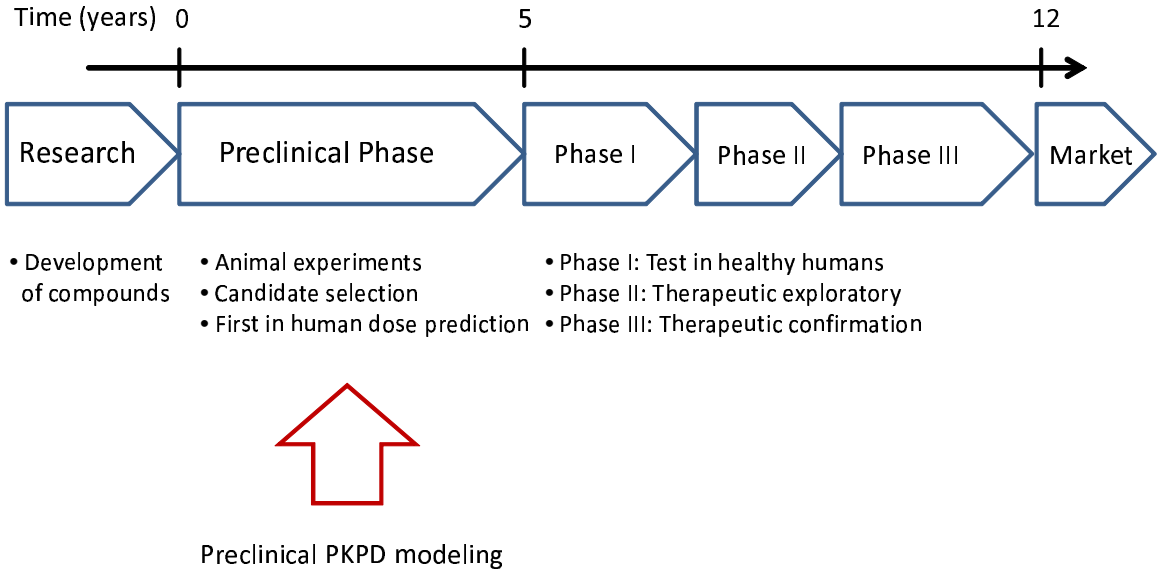


Figure 1.1: *Schematic overview of drug development.*

Several parts of this work are already published in the Journal of Pharmacokinetics and Pharmacodynamics, see [KWLS09], [KWPZ+12] and [KS12]. The presented work was mainly developed as part of the collaboration *Numerical simulation of drug designing experiments (Project no. 735/06)* between the University of Konstanz and Nycomed (A Takeda Company).

Chapter 2

Introduction to Drug and Disease Modeling

In this chapter we give a brief introduction to principles of drug and disease modeling. We will define necessary pharmacological terms and also present typical assumptions from drug development. Further, a general mathematical structure of our models is presented.

2.1 Typical pharmacological assumptions and necessary terms

An experiment in the preclinical phase consists of two parts.

The first part deals with the time course of the drug concentration in blood. The interest is on the distribution of the drug in the body. In this part one does not consider the disease or the effect of the drug on the disease. Roughly spoken, one observes what the body does to the drug. This part is called pharmacokinetics.

The second part observes the development of the disease and the pharmacological effect of the drug on the disease, also called drug response. Again roughly spoken, this time one observes what the drug does to the body. This part is called pharmacodynamics.

Combining pharmacokinetics (PK) and pharmacodynamics (PD) gives an overall picture of the drug response. In PKPD it is assumed that the drug concentration is the driving force of the pharmacological effect on the disease.

A PKPD experiment consists of pharmacokinetic and pharmacodynamic measurements performed in a population of individuals. Typically, the PK data is sparse because blood samples at each measurement time point have to be taken from the individuals. In PD the disease development is described by appropriate readouts. For example, in our experiments, the cancer development was described by the weight of the tumor and, in the arthritis experiments, visual scores describing inflammation and bone destruction were applied. Roughly, one could say that our PD measurements are performed "from the

outside".

To get a realistic overview of the effect of the drug, different doses should be administered in an experiment. The PD data describing the disease with an administered drug is called perturbed. Also a placebo is administered to describe the disease development with no effect of the drug, called unperturbed data. We call the data from one dosing schedule (also including placebo administration) a dosing group. Normally, a dosing group consists of ten animals in our experiments.

When building a PKPD model, the first step is to describe and to fit the PK of a drug. In Chapter 3 we present the modeling of PK with typical linear differential equations. The second and difficult step is to model the disease development. Here it is necessary to understand the mechanism of the disease. We will present in Chapter 4 appropriate model figures based on fundamental biological and pharmacological principles. The final step is to include the PK into the disease development model in order to describe the pharmacological effect. It is obvious that realistic modeling is only possible in close interdisciplinary collaboration.

In this work we focus on so-called (semi-)mechanistic mathematical models. Such models primarily describe the underlying biological situation by first principles and as a result the available data. We are not considering models that just characterize the data without biological assumptions.

We say that a mathematical PKPD model is predictive, if it describes all available dosing groups from one experiment simultaneously by a single model parameter set. The only parameter which varies over the different dosing groups is *dose* or more precisely, the dosing schedule. With a predictive model, simulations for different dosing schedules could be performed. Also for inter-specific scaling of physiological parameters a predictive PKPD model is necessary.

2.2 Mathematical structure of a PKPD model

In this work the general form of a PKPD model is a non-autonomous delay differential equation

$$x'(t) = f\left(t, x(t), x(t-T)\right) \tag{2.1}$$

with the initial function

$$x(s) = \phi(s) \;\text{ for }\; -T \leq s \leq 0\,. \tag{2.2}$$

The parameter $T > 0$ is called a delay. At the moment we suppress additional PKPD model parameter in formulation (2.1)-(2.2).

Note the difference to ordinary differential equations, where $x(t-T)$ does not appear on the right hand side of (2.1) and the initial condition (2.2) is an initial value at $t_0 = 0$.

The main aim of this work is the design of the right hand side in (2.1) to describe PKPD experiments based on fundamental biological and pharmacological principles. Chapter 3 and 4 deals with the derivation of (semi-)mechanistic models. In Chapter 6 and 7, concrete PKPD experiments are modeled.

In PKPD experiments delays are often observed. For example, the effect of a drug is delayed or high concentrations of messengers in the body cause a delayed development of a disease. One major aim of this work is to capture such phenomena by delay differential equations (DDE). In Figure 2.1, we schematically present the standard PKPD approach and indicate possible delays.

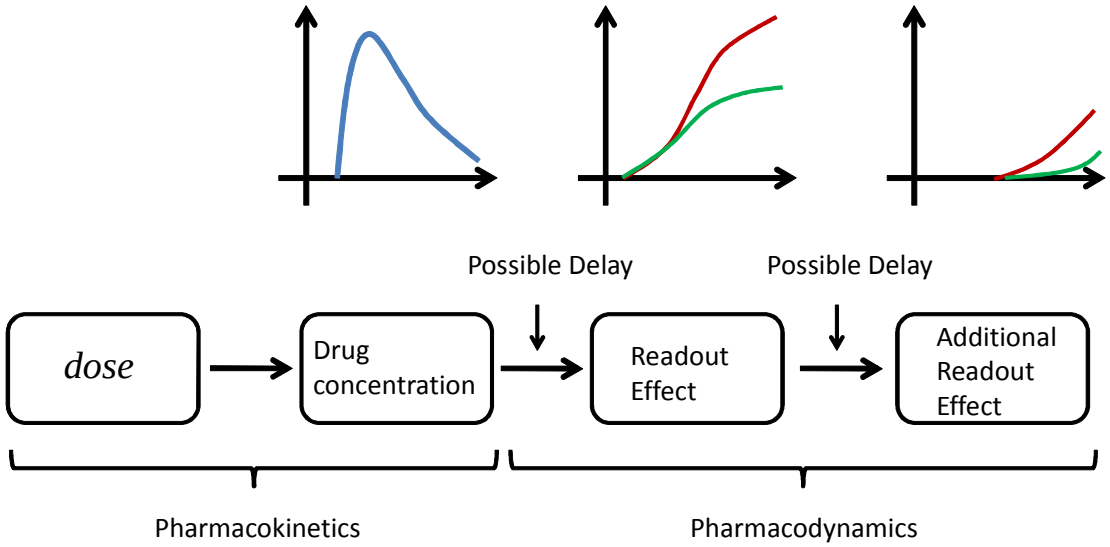


Figure 2.1: *Schematic overview of PKPD principles. Dose drives the drug concentration visualized by the blue curve. The drug concentration (measured in blood) cause an effect (perturbed data, green curve) on the disease (unperturbed data, red curve) with a possible delay. Further also a strongly delayed second response is plotted.*

Until now mainly ordinary differential equations (ODE) are used to build PKPD models in industry as well as in academics. However, we remark that the application of DDEs is of course not new in PKPD. Already in 1982, Steimer et al [SPGB82] presented a model for pharmacokinetics based on DDEs. But somehow DDEs were neglected in PKPD modeling in the last decades and delays were produced by cascades of ODEs. Quite recently, the work of Krzyzanski, Jusko and coworkers (see e.g. [KRJ99], [PRKC+05], [KWJ06] etc.) about lifespan modeling in populations brought DDEs up again to the PKPD community.

A typical existence and uniqueness result for delay differential equations is of the following form.

Theorem 2.2.1
Let $f(t,u,v)$ *be continuous on* $Q = \{(t,u,v) \,|\, 0 \leq t \leq t_{end}\,, u,v \in \mathbb{R}^n\}$ *and satisfy a Lipschitz-condition regarding to* u *and* v*. Let the initial function* $\phi(s)$ *be continuous for* $-T \leq s \leq 0$ *with* $T > 0$*. Then*

$$x'(t) = f(t, x(t), x(t-T))\,, \quad x(s) = \phi(s) \text{ for } -T \leq s \leq 0 \tag{2.3}$$

has a unique solution for $[0, t_{end}]$*.*

A proof based on the contraction mapping principle could be found in [El'73].

Delay differential equations could be rewritten as a system of ordinary differential equations by the method of steps, see e.g. [Dri77]. In Section 4.5 we present a rough categorization of typical PKPD models in DDE form and apply the method of steps to rewrite the models as ODEs.

Chapter 3

Pharmacokinetic Modeling

3.1 Introduction

The pharmacokinetics (PK) describes the behavior of an administered drug in the body over time. The effect of the drug on the disease is not subject of pharmacokinetics. In detail, the PK characterizes the absorption, distribution, metabolism and excretion (called ADME concept, see e.g. [GW06]) of a drug.

The German pediatrist F. H. Dost is deemed to be the founder of the term pharmacokinetics. In his famous books *"Der Blutspiegel"* from 1953 [Dos53] and *"Grundlagen der Pharmakokinetik"* from 1968 [Dos68], he presented a broad overview and analysis of drug behavior in time. For example, he applied linear one-compartment models to describe different drugs and derived several physiological characteristics. However, already in 1937 the Swedish physiologist T. Teorell published first compartment models representing the circulatory system, see [Teo37a], [Teo37b].

The aim of this chapter is to introduce the concept of pharmacokinetic modeling and to motivate the typical structure of PK compartment models. We mainly focus on two-compartment models based on linear differential equations with either intravenous injection or oral absorption of a drug. In this approach, one compartment describes the blood and the other is identified with tissue or more general, with the part of the body which is not heavily supplied with blood. Note that for drug concentration measurements blood samples have to be taken from the patients and therefore, the amount of data is often sparse. Further, measurements in other parts of the body than blood is in the majority of cases impossible. It turned out in practice that two compartments are sufficient to appropriatly describe the time course for most drugs.

This chapter is structured as follows. Firstly, we present in Section 3.2 a general motivation and also typical pharmacological properties and assumptions. In the next step we calculate the explicit solution of the blood compartment by the Laplace transform. We present different parameterizations and important secondary parameter of the two-compartment model to characterize the drug from a physiological point of view. In the next Section 3.3 we present the concept of multiple dosing. Here we focus on n-dimensional

PK compartment models. Finally, we present an example of pharmacokinetic data from our experiments described by a two-compartment model.

3.2 Two-compartment pharmacokinetic models

3.2.1 Assumptions and model building

A two-compartment model consists of two physiological meaningful parts (see e.g. [Kwo01]):

- The first (central) compartment x_1 is identified with the blood and organs heavily supplied with blood like liver or kidney.
- The second (peripheral) compartment x_2 describes for example tissue or more generally, the part of the body which is not heavily supplied with blood.

The compartments are connected among each other in both directions and therefore, a distribution between the central and the peripheral compartment takes place.

Main assumption in pharmacokinetics:

- The drug is completely eliminated (metabolism and excretion) from the body through the blood compartment. In most cases, the metabolism takes place in the liver and the excretion via the kidneys.

We consider two different types of drug administration (absorption):

- The drug is directly administered by an intravenous bolus injection (i.v.) into the blood. It is assumed that the drug is immediately completely distributed in the blood.
- The drug is orally (p.o.) administered by a tablet. Hence, absorption through the stomach takes place. Therefore, the distribution is not immediate and further, only a part of the amount of drug will reach the blood circulation (called bioavailability).

A schematic overview of the two-compartment model is presented in Figure 3.1. To

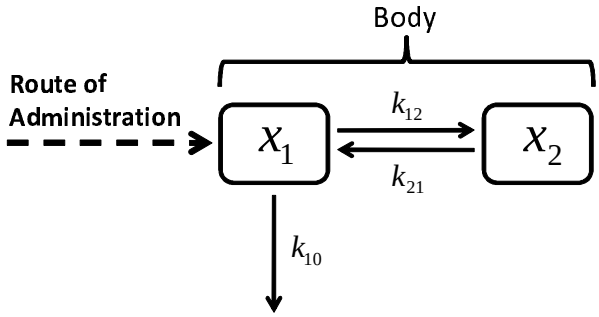


Figure 3.1: *General scheme of the two-compartment model.*

i.v. administration	p.o. administration
$k_{10}, k_{12}, k_{21} > 0$ and $k_{31} \equiv 0$	$k_{10}, k_{12}, k_{21}, k_{31} > 0$
$x^{iv}(0) = (x_1^0, 0, 0)$	$x^{po}(0) = (0, 0, x_3^0)$

Table 3.1: *Different settings for (3.1)-(3.3) or (3.4) to present either i.v. or p.o. administration.*

shorten the notation, we consider the i.v. and p.o. administration at ones. The general form of a two-compartment model describing either i.v. or p.o. drug administration reads

$$x_1'(t) = -k_{10}x_1(t) - k_{12}x_1(t) + k_{21}x_2(t) + k_{31}x_3(t)\,, \quad x_1(0) = x_1^0 \geq 0 \tag{3.1}$$
$$x_2'(t) = k_{12}x_1(t) - k_{21}x_2(t)\,, \quad x_2(0) = 0 \tag{3.2}$$
$$x_3'(t) = -k_{31}x_3(t)\,, \quad x_3(0) = f \cdot x_3^0 \geq 0 \tag{3.3}$$

where $0 < f \leq 1$ is a fraction parameter regulating the amount of drug which effectively reaches the blood in case of p.o. administration (bioavailability). We set without loss of generality $f \equiv 1$ for our mathematical consideration. (3.1) describes the blood compartment, (3.2) the peripheral compartment and (3.3) the absorption in case of p.o. administration. The parameter k_{10} describes the elimination from the body. The rates k_{12} and k_{21} stand for the distribution between central and peripheral compartment and k_{31} is the absorption rate in case of p.o.. Note that the third absorption compartment for p.o. does not count for the nomenclature of the model.

In matrix notation, (3.1)-(3.3) reads with $x \in \mathbb{R}^3$

$$x'(t) = \underbrace{\begin{pmatrix} -k_{10} - k_{12} & k_{21} & k_{31} \\ k_{12} & -k_{21} & 0 \\ 0 & 0 & -k_{31} \end{pmatrix}}_{=:A} \cdot x(t)\,, \qquad x(0) = \begin{pmatrix} x_1^0 \\ 0 \\ f \cdot x_3^0 \end{pmatrix}\,. \tag{3.4}$$

In Table 3.1, the different settings for (3.1)-(3.3) or (3.4) to realize either i.v. or p.o. administration are presented.

Remark 3.2.1
Note that the eigenvalues of the submatrix

$$B = \begin{pmatrix} -k_{10} - k_{12} & k_{21} \\ k_{12} & -k_{12} \end{pmatrix}$$

are real because $T^{-1}BT = C$ is symmetric with

$$T = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{k_{12}k_{21}} \end{pmatrix}\,.$$

Although (3.4) is a linear homogeneous differential equation the representation (3.4) is unhandy in application. In a fitting process the blood compartment $x_1(t)$ has to be evaluated in each iteration at the different measurement time points. If a gradient based optimization method is used, then the gradient of $x_1(t)$ has to be calculated. Further in case of multiple dosing, the representation (3.4) is not adequate. Finally, in a full PKPD model the PK has to be calculated in a tremendous number. Hence, the need for the analytical solution of x_1 is evident.

3.2.2 Analytical solution

We calculate the analytical solution of the blood compartment $x_1(t)$ of (3.4) by the Laplace transform $\mathcal{L}$.

The Laplace transform (see [Wid66] or [Doe76]) is an integral transform where the linear operator $\mathcal{L}\{f(t)\}$ transforms a function $f(t)$ with $t \in \mathbb{R}_{\geq 0}$ from the time domain to a function $F(s)$ with $s \in \mathbb{C}$ in a so-called image domain. The advantage of this transformation is that differentiation and integration in the time domain corresponds to simple algebraic operations in the image domain, for more details see Appendix A.

Applying the Laplace transform to (3.4) gives

$$\begin{aligned}
\mathcal{L}\{x'(t)\} = \mathcal{L}\{Ax(t)\} \quad &\Longleftrightarrow \quad sX(s) - x(0) = AX(s) \\
&\Longleftrightarrow \quad (sI - A)X(s) = x(0) \\
&\Longleftrightarrow \quad \underbrace{\begin{pmatrix} s + k_{12} + k_{10} & -k_{21} & -k_{31} \\ -k_{12} & s + k_{21} & 0 \\ 0 & 0 & s + k_{31} \end{pmatrix}}_{=:L(s)} \underbrace{\begin{pmatrix} X_1(s) \\ X_2(s) \\ X_3(s) \end{pmatrix}}_{=:X(s)} = \underbrace{\begin{pmatrix} x_1^0 \\ 0 \\ x_3^0 \end{pmatrix}}_{=:b} \\
&\Longleftrightarrow \quad L(s) \cdot X(s) = b\,. \qquad (3.5)
\end{aligned}$$

We solve the system of equations (3.5) by Cramer's rule. The determinant of $L(s)$ reads

$$\begin{aligned}
\det(L(s)) = \det(sI - A) &= (s + k_{31})\left[(s + k_{12} + k_{10})(s + k_{21}) - k_{21}k_{12}\right] \\
&= (s + k_{31})\left[s^2 + s(k_{21} + k_{12} + k_{10}) + k_{10}k_{21}\right] \\
&= (s + k_{31})(s + \alpha)(s + \beta)
\end{aligned}$$

where

$$\alpha, \beta = \frac{1}{2}\left(k_{12} + k_{21} + k_{10} \pm \sqrt{(k_{12} + k_{21} + k_{10})^2 - 4k_{21}k_{10}}\right)$$

and hence

$$\alpha\beta = k_{21}k_{10} \quad \text{and} \quad \alpha + \beta = k_{12} + k_{21} + k_{10}\,. \qquad (3.6)$$

By Remark 3.2.1 together with (3.6) we have $\alpha, \beta \in \mathbb{R}_{>0}$ and therefore, $\det(L(s)) = (s + k_{31})(s + \alpha)(s + \beta) > 0$ for all $s \geq 0$.

To calculate the solution of the central compartment x_1 we substitute the vector b into the first column of the matrix $L(s)$ and denote the resulting matrix by $L_1(s)$. The quotient from Cramer's rule reads

$$X_1(s) = \frac{\det(L_1(s))}{\det(L(s))}\,.$$

Now we exemplarily consider the i.v. case. Here the absorption compartment does not exists and we obtain $\det(L^{iv}(s)) = (s+\alpha)(s+\beta)$ and $\det(L_1^{iv}(s)) = x_1^0(s+k_{21})$. Therefore, the Laplace back transform is

$$\begin{aligned}\mathcal{L}^{-1}\{X_1(s)\} &= \mathcal{L}^{-1}\left\{\frac{x_1^0(s+k_{21})}{(s+\alpha)(s+\beta)}\right\} \\ &= x_1^0\mathcal{L}^{-1}\left\{\frac{s}{(s+\alpha)(s+\beta)}\right\} + x_1^0 k_{21}\mathcal{L}^{-1}\left\{\frac{1}{(s+\alpha)(s+\beta)}\right\}.\end{aligned}$$

Because the order of the enumerator polynomial is smaller than the order of the denominator polynomial in each term and the denominator polynomial has distinct roots, we could apply Heaviside's theorem (Appendix A). The derivative of the denominator polynomial $q(s) = (s+\alpha)(s+\beta)$ reads $q'(s) = 2s + \alpha + \beta$ and therefore, $q'(-\alpha) = -\alpha + \beta$ and $q'(-\beta) = -\beta + \alpha$. Hence, we obtain the solution for the first compartment in (3.4)

$$\begin{aligned}\mathcal{L}^{-1}\{X_1(s)\} = x_1^{iv}(t) = x_1^0&\left(\frac{-\alpha}{-\alpha+\beta}\exp(-\alpha t) + \frac{-\beta}{-\beta+\alpha}\exp(-\beta t)\right) \\ &+ x_1^0 k_{21}\left(\frac{1}{-\alpha+\beta}\exp(-\alpha t) + \frac{1}{-\beta+\alpha}\exp(-\beta t)\right)\end{aligned}$$

and finally

$$x_1^{iv}(t) = \frac{x_1^0(k_{21}-\alpha)}{\beta-\alpha}\exp(-\alpha t) + \frac{x_1^0(k_{21}-\beta)}{\alpha-\beta}\exp(-\beta t)\,. \tag{3.7}$$

Using the same technique gives for the p.o. case

$$\begin{aligned}x_1^{po}(t) = &\frac{x_3^0 k_{31}(k_{21}-\alpha)}{(k_{31}-\alpha)(\beta-\alpha)}\exp(-\alpha t) + \frac{x_3^0 k_{31}(k_{21}-\beta)}{(k_{31}-\beta)(\alpha-\beta)}\exp(-\beta t) \\ &+ \frac{x_3^0 k_{31}(k_{21}-k_{31})}{(k_{31}-\beta)(k_{31}-\alpha)}\exp(-k_{31} t)\,.\end{aligned} \tag{3.8}$$

3.2.3 Micro/macro parameterization and secondary parameters

In practice, the drug is measured as concentration in blood. Therefore, the volume of distribution $V_1 > 0$ for the central compartment $x_1(t)$ is introduced to obtain the drug concentration

$$c(t) = \frac{x_1(t)}{V_1}\,. \tag{3.9}$$

V_1 is a proportionality factor between the amount of drug and the drug concentration. In this work, $c(t)$ will always denote the drug concentration in blood.

Finally, we obtain the model parameters of the two-compartment model (3.4)

$$\theta_{mic}^{iv} = (k_{10}, k_{12}, k_{21}, V_1) \quad \text{or} \quad \theta_{mic}^{po} = (k_{10}, k_{12}, k_{21}, V_1, k_{31})$$

which are called the micro constant parameterization. We denote the initial value x_1^0 or x_3^0 by *dose* when speaking of concentration terms. Based on (3.7), (3.8) and (3.9) we define

$$A_{iv} := \frac{k_{21} - \alpha}{V_1(\beta - \alpha)}, \quad B_{iv} := \frac{k_{21} - \beta}{V_1(\alpha - \beta)}$$

as well as

$$A_{po} := \frac{k_{31}}{(k_{31} - \alpha)} A_{iv}, \quad B_{po} := \frac{k_{31}}{(k_{31} - \beta)} B_{iv}.$$

The parameter

$$\theta_{mac}^{iv} = (A_{iv}, B_{iv}, \alpha, \beta) \quad \text{or} \quad \theta_{mac}^{po} = (A_{po}, B_{po}, \alpha, \beta, k_{31})$$

are called macro constant parameterization. The i.v. and p.o. model (3.7) and (3.8) in concentration terms then reads

$$c^{iv}(t) = dose \cdot A_{iv} \cdot \exp(-\alpha t) + dose \cdot B_{iv} \cdot \exp(-\beta t)$$

and

$$c^{po}(t) = dose \cdot A_{po} \cdot \exp(-\alpha t) + dose \cdot B_{po} \cdot \exp(-\beta t) - dose \cdot (A_{po} + B_{po}) \cdot \exp(-k_{31} t)$$

because of

$$\frac{k_{31}}{k_{31} - \alpha} \frac{k_{21} - \alpha}{(\beta - \alpha)(k_{31} - \alpha)} + \frac{k_{31}}{k_{31} - \beta} \frac{k_{21} - \beta}{(\alpha - \beta)(k_{31} - \beta)} = -\frac{k_{31}(k_{21} - k_{31})}{(\alpha - k_{31})(\beta - k_{31})}.$$

Remark 3.2.2
Following relationships between micro and macro parameterization are valid:

$$k_{21} = \frac{A_{iv}\beta + B_{iv}\alpha}{A_{iv} + B_{iv}}, \quad k_{10} = \frac{\alpha\beta}{k_{21}} = \frac{A_{iv} + B_{iv}}{\frac{A_{iv}}{\alpha} + \frac{B_{iv}}{\beta}}, \quad V_1 = \frac{dose}{doseA_{iv} + doseB_{iv}}$$

$$and \quad k_{12} = \frac{A_{iv}B_{iv}(\beta - \alpha)^2}{(A_{iv} + B_{iv})(A_{iv}\beta + B_{iv}\alpha)}.$$

An important pharmacokinetic secondary parameter is the integral of the concentration $c(t)$. This value is called the area under the curve (AUC).

Remark 3.2.3
The AUCs read

$$AUC_\infty^{iv} := \int_0^\infty c^{iv}(s)\, ds = \frac{dose}{V_1 k_{10}} \quad and \quad AUC_\infty^{po} := \int_0^\infty c^{po}(s)\, ds = \frac{dose}{V_1 k_{10}}.$$

3.2.4 Physiological parameterization based on the clearance concept and the idea of allometric scaling

The micro / macro parameterization is not physiological interpretable. In this section we present a physiological meaningful parameterization of the two-compartment model based on the concept of clearance. Gabrielsson stated in [GW06] that *"the clearance is defined as the volume of blood that is totally cleared from its content of drug"*. Hence, one defines the clearance

$$Cl := k_{10}V_1 \,.$$

The inter-compartmental clearance (also called inter-compartmental distribution) from the central to the peripheral compartment and vice versa reads

$$Cl_d^{12} = k_{12}V_1 \quad \text{and} \quad Cl_d^{21} = k_{21}V_2$$

where V_2 denotes the volume of distribution of the peripheral compartment. It yields that

$$Cl_d^{12} = Cl_d^{21} =: Cl_d$$

see e.g. [Hil04]. Hence, the physiological parameterization reads

$$\theta_{phy}^{iv} = (Cl, Cl_d, V_1, V_2) \quad \text{or} \quad \theta_{phy}^{po} = (Cl, Cl_d, V_1, V_2, k_{31}) \,.$$

Finally, we give a short comment on allometric (inter-species) scaling of physiological parameters like clearance or volume of distribution. First, to perform a scaling, the underlying mechanism in the different species has to be similar. Second, it is commonly believed that clearance or volume of distribution depend on the body weight w, see [MCM+91]. A typical allometric model for scaling a physiological parameter y is

$$y(w) = a \cdot w^b \tag{3.10}$$

where a, b are allometric parameters, see [MCM+91] or [GW06]. It is suggested that at least 4 to 5 species are necessary to predict from mouse to human. For example, in [MCM+91] different therapeutic proteins were scaled with (3.10). A typical structure is mouse, rat, rabbit, monkey and finally human.

3.3 Multiple dosing for n-dimensional compartment models

The next step to describe the pharmacokinetics of a drug is to handle multiple dosing, that means, a drug is administered several times to the body. Hence, one has also to account for the remaining drug concentration in the body from a previous dosage.

In application, a drug is designed for equidistant administration, for example, every day, every second day, every week and so on. This makes the application of drugs more secure for patients and therefore, increases the success on the market.

In this section we focus on general n-dimensional linear PK compartment models. A n-dimensional compartment model is a linear homogenous differential equation

$$x'(t) = Ax(t)\,, \qquad x(t_0) = x^0 \in \mathbb{R}^n \tag{3.11}$$

with $A \in \mathbb{R}^{n,n}$. The analytical solution is given by the matrix exponential function

$$x(t) = \exp((t - t_0)A)x^0\,.$$

In a first step, we focus on the general situation of arbitrary dosing time points. We denote by $\pi = (\pi_1, ..., \pi_m) \in \mathbb{R}^m_{\geq 0}$ the different dosing time points with $\pi_{k-1} < \pi_k$ for $k \in \{2, ..., m\}$. By $\delta_j \in \mathbb{R}^n$ we denote the doses for every compartment for $j = 1, ..., m$. Now, $x_j \in \mathbb{R}^n$ describes the j-th dosage and not the scalar compartment of the n-dimensional compartment system.

Proposition 3.3.1
The multiple dosing formula for a linear homogenous differential equation (3.11) reads

$$x(t) = \begin{cases} 0 & \text{for } 0 \leq t < \pi_1 \\ x_1(t) & \text{for } \pi_1 \leq t < \pi_2 \\ \vdots & \\ x_m(t) & \text{for } \pi_m \leq t \end{cases} \tag{3.12}$$

with

$$x_j(t) = \sum_{i=1}^{j} \exp((t - \pi_i)A)\delta_i \in \mathbb{R}^n\,. \tag{3.13}$$

Proof: We have

$$\lim_{t \nearrow \pi_j} x(t) = \lim_{t \nearrow \pi_j} x_{j-1}(t) = x_{j-1}(\pi_j) \quad \text{and} \quad \lim_{t \searrow \pi_j} x(t) = \lim_{t \searrow \pi_j} x_j(t) = x_j(\pi_j)\,.$$

Hence, we have to show that

$$\lim_{t \searrow \pi_j} x(t) - \lim_{t \nearrow \pi_j} x(t) = x_j(\pi_j) - x_{j-1}(\pi_j) = \delta_j\,.$$

This follows by

$$\begin{aligned} x_{j-1}(\pi_j) + \delta_j &= \sum_{i=1}^{j-1} \exp((\pi_j - \pi_i)A)\delta_i + \delta_j = \sum_{i=1}^{j-1} \exp((\pi_j - \pi_i)A)\delta_i + \underbrace{\exp((\pi_j - \pi_j)A)}_{=I}\delta_j \\ &= \sum_{i=1}^{j} \exp((\pi_j - \pi_i)A)\delta_i = x_j(\pi_j)\,. \end{aligned}$$

□

Based on the representation (3.12)-(3.13) one could easily code an algorithm for multiple dosing.

Now we consider equidistant dosing intervals with an equal amount for all doses. As mentioned before, this situation is the realistic scenario in drug development. Let $\tau \in \mathbb{R}_{>0}$ be the length of the dosing interval and $d \in \mathbb{R}^n_{\geq 0}$ the dose. Further, let $t_0 = 0$ be the first dosing time point.

We assume:

(A1) The eigenvalues of the matrix A are single, mutually distinct, real and negative.

This assumption is fulfilled for typical PK compartment models of mammillary or catenary type, compare [GP82] and [And83].

Applying the spectral theorem gives $\sigma(I - \exp(\tau A)) > 0$ for $\tau > 0$ and therefore, the invertibility of $I - \exp(\tau A)$ for $\tau > 0$.

Remark 3.3.2
Let (A1) hold. With equidistant dosing time points $j \cdot \tau$, $j \in \{1, ..., m\}$ and equal dose d one obtains for (3.13) the representation

$$\widetilde{x}_j(\xi) = \exp(\xi A)(I - \exp(j\tau A))\,(I - \exp(\tau A))^{-1}\,d \tag{3.14}$$

for $\xi \in [0, \tau]$.

Proof: With (3.13) for $s \in [\pi_j, \pi_{j+1}]$

$$\begin{aligned} x_j(s) &= \sum_{i=1}^{j} \exp((s - \pi_i)A)d \\ &= \exp((s - \pi_j)A)\big(\exp((\pi_j - \pi_1)A) + \cdots + \exp((\pi_j - \pi_{j-1})A) + \exp((\pi_j - \pi_j)A)\big)d \\ &= \exp((s - \pi_j)A)\big(\exp((j-1)\tau A) + \cdots + \exp(\tau A) + I\big)d \\ &= \exp((s - \pi_j)A)\left(\sum_{k=0}^{j-1} \exp(k\tau A)\right)d\,. \end{aligned}$$

With the geometric series for matrices

$$\begin{aligned} x_j(s) &= \exp((s - \pi_j)A)\left(\sum_{k=0}^{j-1} \exp(\tau A)^k\right)d \\ &= \exp((s - \pi_j)A)\left(I - \exp(\tau A)^j\right)(I - \exp(\tau A))^{-1}\,d \end{aligned}$$

for $s \in [\pi_j, \pi_j + \tau]$. Hence, with $\xi \in [0, \tau]$

$$\widetilde{x}_j(\xi) = \exp(\xi A)\,(I - \exp(j\tau A))\,(I - \exp(\tau A))^{-1}\,d\,.$$

□

The representation (3.14) for $\xi \in [0, \tau]$ is the generalized standard version used in pharmacokinetic modeling for multiple dosing, see [GP82]. Note that in (3.12)-(3.13)

$$x_j(t) = \widetilde{x}_j(t - (j-1)\tau) \quad \text{for } j = 1, ..., m.$$

An important situation in PK experiments is the so-called steady state concentration that means, if the number of equidistant administration tends to infinity. Because of assumption (A1), we have $Re\,\sigma(A) < \alpha < 0$ and therefore, $||\exp(tA)|| \leq \exp(t\alpha)$ for all $t \in \mathbb{R}_{>0}$, see [Ama95]. Hence,

$$\exp(tA) \to 0 \quad \text{for } t \to \infty.$$

The steady state function reads

$$\widetilde{x}^*(\xi) := \lim_{j\to\infty} \widetilde{x}_j(\xi) = \exp(\xi A)\left(I - \exp(\tau A)\right)^{-1} d \quad \text{for } \xi \in [0, \tau]\,.$$

In application, the steady state concentration is for example used to predict an appropriate *dose* for human based on inter-specific scaling. Therefore, we finally present an important property of equidistant dosing. With this feature the area under the curve of the steady state concentration for multiple dosing could be calculated based on just a single dose experiment.

Remark 3.3.3
Let (A1) hold. It yields for a single drug administration $x(t) = \exp(tA)d$ that

$$\int_0^\infty x(s)ds = \int_0^\tau \widetilde{x}^*(s)ds\,.$$

Proof: The area under the curve of single drug administration is

$$\int_0^\infty x(s)ds = \int_0^\infty \exp(sA)d\,ds = \left[A^{-1}\exp(sA)d\right]_0^\infty = -A^{-1}d\,.$$

For the steady state function we obtain

$$\begin{aligned}
&\int_0^\tau \exp(sA)\left(I - \exp(\tau A)\right)^{-1} ds \\
&= \left[A^{-1}\exp(sA)\left(I - \exp(\tau A)\right)^{-1}\right]_0^\tau \\
&= A^{-1}\exp(\tau A)\left(I - \exp(\tau A)\right)^{-1} d - A^{-1}\left(I - \exp(\tau A)\right)^{-1} d \\
&= \left(A^{-1}\exp(\tau A) - A^{-1}\right)\left(I - \exp(\tau A)\right)^{-1} d \\
&= -A^{-1}\left(-\exp(\tau A) + I\right)\left(I - \exp(\tau A)\right)^{-1} d \\
&= -A^{-1}d\,.
\end{aligned}$$

□

Note that the steady state concentration is a theoretical result for $j \to \infty$ and in practice one has for the drug concentration in blood

$$\lim_{t\to\infty} c(t) = 0\,.$$

Example 3.3.4
Consider the two-compartment model with single p.o. administration

$$c(t) = dose \cdot A_{po} \cdot \exp(-\alpha t) + dose \cdot B_{po} \cdot \exp(-\beta t) - dose \cdot (A_{po} + B_{po}) \cdot \exp(-k_{31} t)\,.$$

Following the results presented in this section, we obtain the multiple dosing representation for equidistant dosing time points

$$\begin{aligned} c_j(\xi) = {} & dose \cdot A_{po} \cdot \frac{1-\exp(-\tau j\alpha)}{1-\exp(-\tau\alpha)} \exp(-\alpha\xi) + dose \cdot B_{po} \cdot \frac{1-\exp(-\tau j\beta)}{1-\exp(-\tau\beta)} \exp(-\beta\xi) \\ & - dose \cdot (A_{po} + B_{po}) \cdot \frac{1-\exp(-\tau j k_{31})}{1-\exp(-\tau k_{31})} \exp(-k_{31}\xi) \end{aligned}$$

with $\xi \in [0, \tau]$ for the j-th dosing time point.

The property of Remark 3.3.3 is visualized in Figure 3.2.

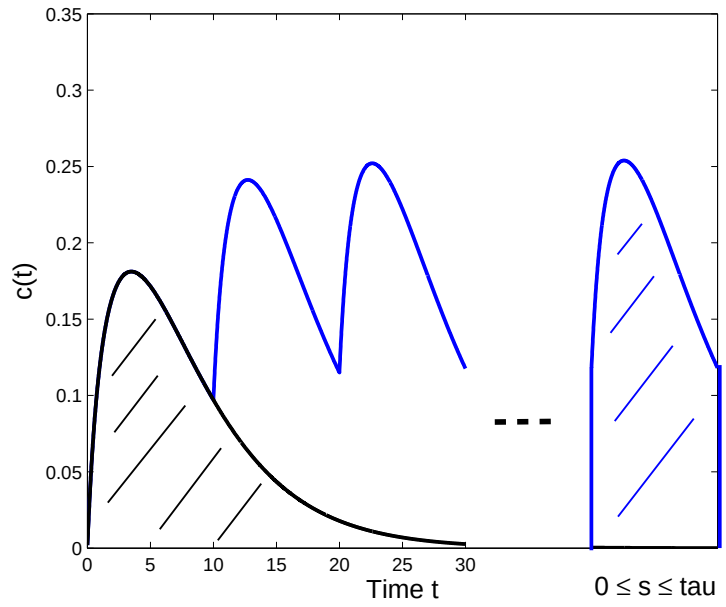


Figure 3.2: *Property of Corollar 3.3.3 for the two-compartment p.o. model from Example 3.3.4.*

3.4 Example for pharmacokinetic data of an antibody

We consider the GM-CSF monoclonal antibody 22E9 administered in mice, see Chapter 7 or [KWPZ+12] for more details. The antibody was applied several times with four different doses, see Table 3.2.

To simultaneously fit the measured data, the multiple dosing formula (3.12)-(3.13) with a two-compartment model i.v. in macro parameterization was used. In the fitting process the data was weighted, see Table 3.2, which equals their contribution to the model, see Chapter 5. In Table 3.3, the parameter estimates in the macro parameterization and in the equivalent physiological parameterization are presented. Additionally, we indicate the coefficient of variation, the 95%-confidence interval and the coefficient of determination, see Chapter 5 for more details. The simultaneous fit is presented in Figure 3.3.

Dose (mg/kg)	100	10	1	0.1
Time Points (hr)	0, 336	0, 168, 336	0, 168, 336	0, 168, 336
Weights	0.1	1	10	100

Table 3.2: *Dose and dosing time points of the antibody 22E9. Further the weights used in the fitting process are listed.*

Macro constants	Value (CV%) CI
A_{iv}	20.27 (5.2) [18.2, 22.4]
B_{iv}	17.54 (5.9) [15.48, 19.60]
α	0.2256 (12.4) [0.170, 0.281]
β	0.0065 (7.0) [0.005, 0.007]
Sum of squares	41009
R^2 (100 - 0.1)	0.99 / 0.97 / 0.96 / 0.99
Physiological constants	Value
Cl	0.0004
Cl_d	0.0029
V_1	0.0265
V_2	0.0270

Table 3.3: *Pharmacokinetic parameters of 22E9 for the two compartment model i.v.. The fitting parameters are in macro constant parameterization ($A_{iv}, B_{iv}, \alpha, \beta$). The physiological parameters Cl (clearance), Cl_d (intercompartmental distribution), V_1 (volume of distribution of the first compartment) and V_2 (volume of distribution of the second compartment) are calculated a posteriori.*

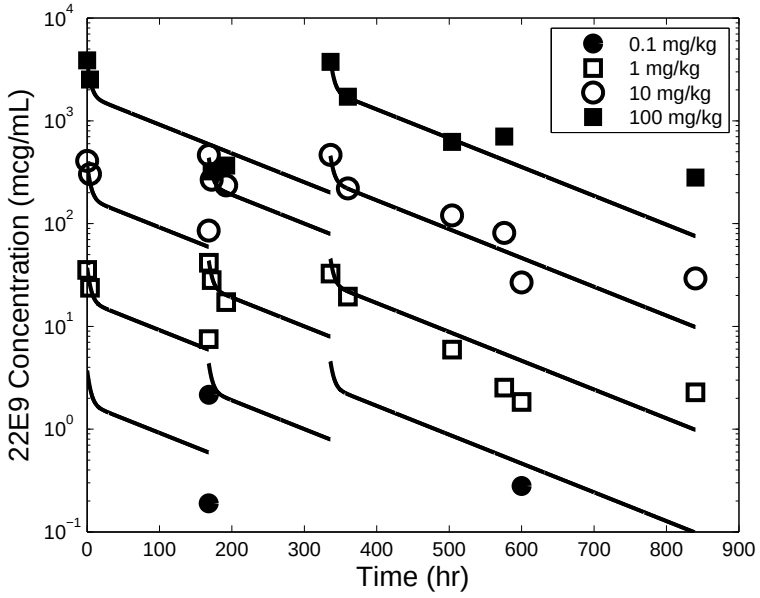


Figure 3.3: *Simultaneous fit of the antibody 22E9 concentration measured in blood for all available dosing schedules.*

3.5 Discussion and outlook

The compartment approach based on linear differential equations is the standard technique in pharmacokinetic modeling because it allows the identification of parts of the body with compartments in the model. More precisely, in PK studies mainly two-compartment models are applied to fit data because the data situation is usually sparse. We remark that the amount of data presented in Section 3.4 is uncommonly large.

The straightforwardness of analytically solving linear differential equations is of major importance in pharmacokinetic / pharmacodynamic modeling. Note that in the final PKPD model the drug concentration $c(t)$ has to be evaluated in a tremendous way.

However, from the modeling point of view there are several legitimate questions. For example, are the rate constants k_{ij} really constant (see e.g. [Jon06]) or do they maybe depend on outside influences (like temperature, age, weight,...)? In general, mice experiments are preformed under standardized laboratory conditions and the mice are from the same strain. In [MMN$^+$04] it is shown that different age and strain of mice significantly affect the levels of drug (in their work cocaine was observed) in brain and blood.

A new approach for pharmacokinetic modeling is based on the idea that the body behaves like a fractional system, which is from a physiological point of view a reasonable assumption, see [DM09] and [DMM10]. Such models are based on fractional calculus, that means, the derivative could be of real valued order instead of integers, as in classical calculus.

Anyway, such models were just recently introduced to the pharmacokinetic community and as far as we know not applied in a full PKPD model until now.

Chapter 4

Model Figures

In this chapter we present and deduce different model figures which will be used in our pharmacokinetic/pharmacodynamic models. We will also discuss the biological background and interpretation.

In Section 4.1 we present a general inflow/outflow model. Such a model has a zero order inflow into a state and a first order outflow from that state. An important property of this model is that, under realistic conditions, every solution runs into a stationary point. Such a stationary point is of fundamental biological importance and therefore, the models are frequently applied in PKPD modeling.

In Section 4.2 we consider transit compartment models (TCM). Such models consist of n states put in series and mimic signal transduction cascades. Roughly spoken, a TCM describes the pathway of a signal. However, in PKPD modeling TCMs are also just used to produce any kind of delay or to describe populations. In such applications, the different states of a TCM could lose their biological identification.

Therefore, we introduce in the next Section 4.3 the concept of lifespan models (LSM). In this model an individual enters a population and stays a certain time in this population, called the lifespan. After that lifespan the individual irrevocably has to leave the population. Lifespan models consist of exactly one state.

In Section 4.4 we present an important relationship between TCMs and LSMs. The obtained theoretical result enables the modeler to substitute TCMs (n states) by LSMs (one state). This result is accepted for publication in the Journal of Pharmacokinetics and Pharmacodynamcis, see [KS12].

In Section 4.5 a rough classification of models with an explicit delay is presented.

Section 4.6 is independent of the previous ones and deals with drug-effect terms. Such a term describes the effect of the drug on the target. This target could be the disease itself (e.g. proliferating cells) or the inhibition of messengers (e.g. cytokines, which have an indirect effect on the disease). Two classical and one new drug-effect term will be

presented. Our new drug-effect term is capable to describe non-monotonic drug effects, that means, a dosing group with a lower amount of administered drug shows a higher effect than the higher dosing group.

4.1 The inflow/outflow model

Consider a state $x(t)$ controlled by two processes, an inflow into that state and an outflow from that state. A reasonable realization is by a zero-order inflow and a first-order outflow. We call such a model an inflow/outflow model (IOM). See Figure 4.1 for a schematic representation.

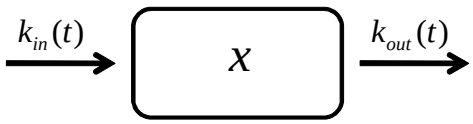


Figure 4.1: *Schematic representation of an inflow/outflow model.*

An important property of an IOM is that under realistic conditions every solution runs into a steady state or mathematically spoken, the system has a global asymptotically stable stationary point. Such a steady state behavior is of fundamental importance for pharmacological modeling.

The IOMs are part of the so-called class of turnover driven models, see [GW06]. Gabrielsson stated that *"turnover driven models are typically based on sound biological principles"* and *"the variables and parameters have ideally a physiological meaning and can often be related to ... physiological data"* in [GW06].

Let $k_{in} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and $k_{out} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ be the inflow and outflow, respectively. Let these functions be piecewise continuous and bounded. We assume that the limits

$$\lim_{t\to\infty} k_{in}(t) = k_{in}^* \geq 0 \quad \text{and} \quad \lim_{t\to\infty} k_{out}(t) = k_{out}^* > 0 \tag{4.1}$$

exist. An inflow/outflow model is of the form

$$x'(t) = k_{in}(t) - k_{out}(t) \cdot x(t)\,, \quad x(0) = x^0 \geq 0 \tag{4.2}$$

with the asymptotically stable stationary point

$$x^* = \lim_{t\to\infty} x(t) = \frac{k_{in}^*}{k_{out}^*}\,.$$

Because of the existence of the limits (4.1), (4.2) implies

$$0 = f(x^*) = k_{in}^* - k_{out}^* \cdot x^* \quad \Longrightarrow \quad x^* = \frac{k_{in}^*}{k_{out}^*}$$

and with $f'(x^*) = -k_{out}^* < 0$, x^* is asymptotically stable, see [HK96].

4.1.1 Application of inflow/outflow models - Indirect response models

In pharmacodynamics, one is often faced with a so-called indirect drug response, that means, the drug stimulates or inhibits factors which control the response, see [DGJ93]. Further, one assumes that the system is in a so-called baseline condition. For example, think of messengers in the body or heart rate. The aim is to describe a perturbation of this baseline by a drug $c(t)$. Moreover, if the perturbation vanishes, it is assumed that the response runs back into the baseline.

The basic equation of an indirect response model (IDR) is of the form (4.2) with constant positive inflow and outflow rates. This model reads

$$x'(t) = k_{in} - k_{out} \cdot x(t)\,, \qquad x(0) = x^0 \geq 0 \tag{4.3}$$

with the solution

$$x(t) = \frac{k_{in}}{k_{out}} + \left(x^0 - \frac{k_{in}}{k_{out}}\right) \exp(-k_{out} \cdot t)\,.$$

For a baseline condition the initial value is set equal to the steady state

$$x^0 = x^* = \frac{k_{in}}{k_{out}}\,.$$

In standard indirect response models, a Michaelis-Menten drug-effect term with Hill coefficient (see Section 4.6) is applied. Depending on which rate is stimulated or inhibited, one obtains four possible models, see originally Dayneka, Jusko and coworkers [DGJ93] or summarized [GW06] for the response $R(t)$, presented in compact form

$$\begin{aligned} R'(t) = k_{in} \cdot \left\{ \left(1 - \frac{I_{max} c(t)^h}{IC_{50}^h + c(t)^h}\right)\,,\, \left(1 + \frac{E_{max} c(t)^h}{EC_{50}^h + c(t)^h}\right) \right\} \\ - k_{out} \cdot \left\{ \left(1 - \frac{I_{max} c(t)^h}{IC_{50}^h + c(t)^h}\right)\,,\, \left(1 + \frac{E_{max} c(t)^h}{EC_{50}^h + c(t)^h}\right) \right\} \cdot R(t)\,, \qquad R(0) = \frac{k_{in}}{k_{out}} \end{aligned} \tag{4.4}$$

where $0 < I_{max} \leq 1$. Note that $\lim\limits_{t\to\infty} c(t) = 0$. IDRs (4.4) are one of the most popular models in PKPD and are extensively studied and applied by PD scientists in the last 20 years. Note that from the mathematical point of view, (4.4) is a special design of our general inflow/outflow model (4.2).

4.2 The transit compartment model

A widely used model in PKPD is the transit compartment model (TCM)

$$x_1'(t) = k_{in}(t) - k \cdot x_1(t)\,, \qquad x_1(0) = x_1^0 \geq 0 \tag{4.5}$$

$$x_2'(t) = k \cdot x_1(t) - k \cdot x_2(t)\,, \qquad x_2(0) = x_2^0 \geq 0 \tag{4.6}$$

$$\vdots$$

$$x_n'(t) = k \cdot x_{n-1}(t) - k \cdot x_n(t)\,, \quad x_n(0) = x_n^0 \geq 0 \tag{4.7}$$

where $k_{in} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a piecewise continuous and bounded function and $k \in \mathbb{R}_{>0}$ denotes the transit rate between the compartments. Roughly spoken, the states $x_2(t), ..., x_n(t)$ are delayed versions of $x_1(t)$. Note that a TCM actually consists of n inflow/outflow models (4.2) put in series.

A schematic representation is presented in Figure 4.2. In Figure 4.3 some solutions of the model (4.5)-(4.7) are plotted.

The application of (4.5)-(4.7) is versatile in PKPD modeling. TCMs are for example motivated based on signal transduction processes, see [SJ98], and therefore, mimic biological signal pathways. For example, in [FHM$^+$02] the maturation of cells for chemotherapy-induced myelosuppression was described by TCMs. But TCMs are also often used to just produce delays, see [LB02] (delayed drug course) or [EDM$^+$08b]-[EDM$^+$08a] (delayed cytokine growth). Hence, the states $x_i(t)$ often lose their pharmacological interpretation and the TCM concept is downgraded to a help technique. Historically, Sheiner was the first in 1979, see [SSV$^+$79], who suggested to apply a TCM with $n = 1$ to describe a delay between pharmacokinetics and effect.

TCMs are also applied to describe populations, see [SMC$^+$04] or Chapter 6. Because when looking at a TCM more precisely, one could assign a mean residence/transit time of $\frac{1}{k}$ for an individual to stay in the i-th compartment, $i \in 1, ..., n$, see e.g. [SJ98]. In this sense, a TCM could be reinterpreted as a model describing an age structured population and $x_i(t)$ describes the number of individuals with age a_i, where $a_i \in (\frac{i-1}{k}, \frac{i}{k}]$. Hence, spoken in population, the $x_1(t), ..., x_n(t)$ describe the age distribution of a total population

$$y_n(t) = x_1(t) + \cdots + x_n(t)\,.$$

Therefore, the secondary parameter

$$T = \frac{n}{k}$$

describes the mean transit/residence time needed for an object created by k_{in} to pass through all states $x_i(t)$ with $i = 1, ..., n$.

However, in most cases it is obvious that the choice of the number of compartments n is somehow arbitrary. In application, n is often chosen in such a way that the final PKPD model fits the data best. For example, Savic and Karlsson [SJKK07] presented a technique to determine the optimal number of compartments based on fitting results for delayed PK p.o. data.

We will show in Section 4.4 an important property of the TCM when the number of compartments tends to infinity. For that purpose we are interested in the analytical solution of system (4.5)-(4.7).

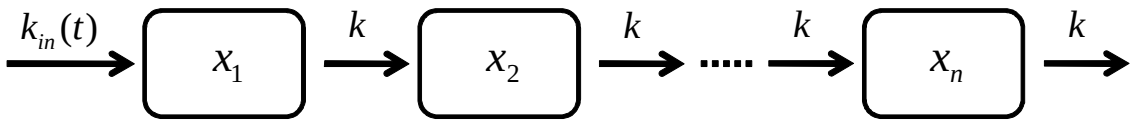


Figure 4.2: *Schematic representation of the transit compartment model (4.5)- (4.7).*

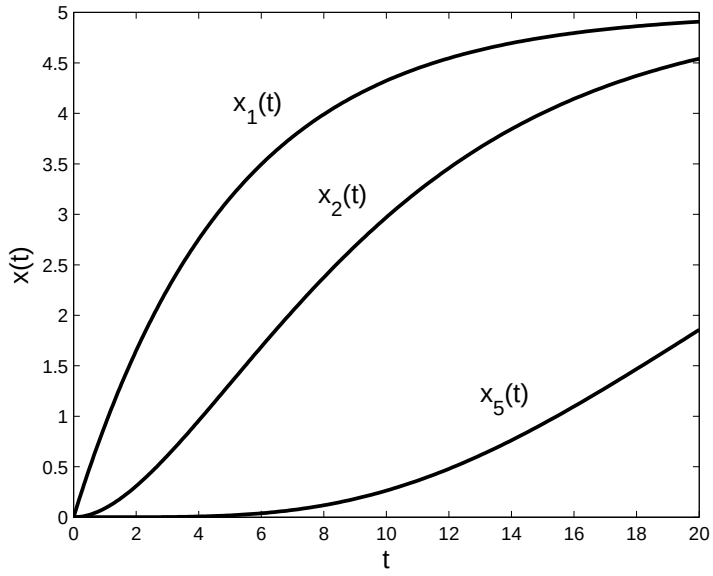


Figure 4.3: *Solutions of the transit compartment model (4.5)- (4.7) for* $n = 5$ *with the parameter* $k_{in} \equiv 1$, $k = 0.2$ *and* $x_1^0 = \cdots = x_5^0 = 0$.

Remark 4.2.1
The analytical solution of the transit compartment model (4.5)-(4.7) reads

$$x(t) = X(t) \cdot x^0 + \int_0^t X(t-s) \cdot k_{in}(s) \cdot e_1 \, ds \in \mathbb{R}^n \tag{4.8}$$

where $e_1 = (1, 0, ..., 0)^T \in \mathbb{R}^n$ *and*

$$X(t) = \frac{1}{k} \cdot \begin{pmatrix} g_k^1(t) & & & \\ g_k^2(t) & g_k^1(t) & & \\ \vdots & & \ddots & \\ g_k^n(t) & \dots & g_k^2(t) & g_k^1(t) \end{pmatrix} \in \mathbb{R}^{n,n}$$

with the gamma probability density functions

$$g_k^j(t) = \frac{k^j t^{j-1}}{(j-1)!} \exp(-kt) \,, \quad t \geq 0 \,, \quad j = 1, ..., n \,. \tag{4.9}$$

Proof: Consider system (4.5)-(4.7) in matrix notation

$$x'(t) = A \cdot x(t) + k_{in}(t) \cdot e_1 \,, \quad x(0) = x^0$$

with

$$A = \begin{pmatrix} -k & & & 0 \\ k & -k & & \\ & \ddots & \ddots & \\ 0 & & k & -k \end{pmatrix} \in \mathbb{R}^{n,n} \,.$$

By the variation of constants formula the solution reads

$$x(t) = X(t) \cdot x^0 + \int_0^t X(t-s) \cdot k_{in}(s) \cdot e_1 \, ds$$

with $X(t) = \exp(tA)$, $t \geq 0$. Now we calculate the explicit representation of the fundamental matrix $X(t)$. Let

$$N = \begin{pmatrix} 0 & & & 0 \\ 1 & 0 & & \\ & \ddots & \ddots & \\ 0 & & 1 & 0 \end{pmatrix} \in \mathbb{R}^{n,n} \,.$$

The matrix N is nilpotent. Further N has the property that for each multiplication with itself, the diagonal with the ones slides to the left lower corner. We split the compartment matrix A as follows

$$A = -k \cdot (I - N)$$

where $I \in \mathbb{R}^{n,n}$ is the identity matrix. Then we can compute

$$X(t) = \exp(tA) = \exp(-tk(I-N)) = \exp(-kt)\exp(ktN) = \frac{1}{k} g_k^1(t)\exp(ktN)\,.$$

Because N is nilpotent the matrix exponential $\exp(ktN)$ reads

$$\exp(ktN) = \sum_{j=0}^{\infty} \frac{(kt)^j}{j!} N^j = \sum_{j=0}^{n-1} \frac{(kt)^j}{j!} N^j = I + ktN + \frac{(kt)^2}{2!} N^2 + \cdots + \frac{(kt)^{n-1}}{(n-1)!} N^{n-1}\,.$$

With the additional property of the powers of N we obtain

$$\exp(ktN) = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ kt & 1 & 0 & & 0 & 0 \\ \frac{(kt)^2}{2!} & kt & 1 & & 0 & 0 \\ \vdots & & & & \ddots & \vdots \\ \frac{(kt)^{n-1}}{(n-1)!} & \frac{(kt)^{n-2}}{(n-2)!} & & \dots & kt & 1 \end{pmatrix}$$

and finally

$$X(t) = \frac{1}{k} g_k^1(t)\exp(ktN) = \frac{1}{k} \begin{pmatrix} g_k^1(t) & & & 0 \\ g_k^2(t) & g_k^1(t) & & \\ \vdots & & \ddots & \\ g_k^n(t) & \dots & g_k^2(t) & g_k^1(t) \end{pmatrix}.$$

□

Remark 4.2.2
For $\lim_{t\to\infty} k_{in}(t) = k_{in}^*$ *the asymptotically stable stationary points are*

$$x_i^* = \frac{k_{in}^*}{k} \quad \textit{for } i = 1, ..., n\,.$$

Finally, again by setting $x_i^0 = x_i^*$ for $i = 1, ..., n$ a baseline condition is obtained.

4.3 Lifespan models

Lifespan models (LSM) were introduced by Krzyzanski and Jusko in 1999 [KRJ99] to PKPD modeling. They applied this approach to cell populations in the context of indirect response models.

In this section we consider the lifespan approach from a more general point of view as e.g. in [KRJ99]. Let $y(t)$ be a state controlled by production (birth) and loss (death). The general form of such a model is

$$y'(t) = k_{in}(t) - k_{out}(t)\,, \quad y(0) = y^0 \tag{4.10}$$

where k_{in} and k_{out} are piecewise continuous and bounded functions.

Now we assume that every individual has a certain lifespan when it enters the state. After this lifespan the individual irrevocably vanishes from the state. This will lead to special forms of $k_{out}(t)$ in (4.10).

4.3.1 Lifespan models with constant lifespan

First, we consider the case that every object in the state y has the same lifespan $T > 0$. Then the outflow from state y at time t is equal to the inflow at time $t - T$ and we obtain the relation

$$k_{out}(t) = k_{in}(t - T) \quad \text{for } t \geq 0\,.$$

Hence, the LSM for constant lifespan reads

$$y'(t) = k_{in}(t) - k_{in}(t - T)\,, \quad y(0) = y^0\,. \tag{4.11}$$

In application one has seldom the freedom of choosing the initial value $y(0) = y^0$ arbitrarily. For example, when thinking in populations the initial value y_0 has to be set in such a way that it describes the amount of individuals already born and also died in the interval $[-T, 0]$.

In case of constant lifespan T no individual has died in the interval $[-T, 0]$ and therefore, we immediately obtain

$$y^0 = \int_{-T}^{0} k_{in}(s)\, ds\,. \tag{4.12}$$

The solution of (4.11)-(4.12) reads

$$y(t) = \int_{t-T}^{t} k_{in}(s)\, ds \quad \text{for } t \geq 0$$

which directly follows by differentiation.

An important case in application is a constant production in the past

$$k_{in}(s) = k_{in}^* \quad \text{for } s \leq 0\,.$$

Then the initial value (4.12) is

$$y^0 = T \cdot k_{in}^*\,.$$

However, it is immediately clear that the assumption of constant lifespan is idealized and in reality not true. Nevertheless, we will focus on this abstraction because in our experiments we do not have enough data to apply a non-constant lifespan approach. In Chapter 6 and 7 we apply a LSM with constant lifespan of the form (4.11).

Anyhow, we will also investigate LSMs with a distributed lifespan in the next section.

4.3.2 Lifespan models with distributed lifespan

The assumption that every individual in the population has the same lifespan is idealized and not really realistic. Hence, we also consider the case that individuals born at time t

will vanish around time $t + T$. For that purpose we introduce a distributed lifespan for individuals.

Let X be a random variable distributed based on a probability density function (PDF) $l : \mathbb{R} \to \mathbb{R}_{\geq 0}$ with $l(s) = 0$ for $s < 0$ and with expectation $T = \mathbb{E}[X]$. The outflow term then reads

$$k_{out}(t) = \int_0^\infty k_{in}(t-\tau)l(\tau)\,d\tau = \int_{-\infty}^\infty k_{in}(t-\tau)l(\tau)\,d\tau = (k_{in} * l)(t) \tag{4.13}$$

where " $*$ " denotes the convolution. In delay differential equation theory the second term in (4.13) is generally called a distributed delay, see [Smi10]. In our terminology we call the term a distributed lifespan, see e.g. [KWJ06] for derivation. The LSM with distributed lifespan reads

$$y'(t) = k_{in}(t) - (k_{in} * l)\,(t)\,, \quad y(0) = y^0\,. \tag{4.14}$$

Again when thinking in populations, the initial value y^0 has to be chosen in such a way that it describes the amount of individuals already born and died. We calculate the start amount y^0 based on a PDF $l(\tau)$.

Remark 4.3.1
The start amount for (4.14) reads

$$y^0 = \int_0^\infty \int_{-\tau}^0 k_{in}(s)\,ds\,l(\tau)\,d\tau\,. \tag{4.15}$$

Proof: To estimate the amount of the start population y^0, we define

$$\widetilde{k_{in}}(t) := \begin{cases} k_{in}(t) & \text{for } t \leq 0 \\ 0 & \text{for } t > 0 \end{cases}. \tag{4.16}$$

We claim

$$\lim_{t\to\infty} y(t) = 0 \tag{4.17}$$

in order to calculate the start population based on (4.16). Consider (4.14) substituted with (4.16)

$$y'(t) = \widetilde{k_{in}}(t) - \int_0^\infty \widetilde{k_{in}}(t-\tau)l(\tau)\,d\tau\,. \tag{4.18}$$

The solution of (4.18) reads

$$y(t) = y^0 + \int_0^t \widetilde{k_{in}}(s)\,ds - \int_0^t \int_0^\infty \widetilde{k_{in}}(s-\tau)l(\tau)\,d\tau\,ds\,.$$

Because of the definition in (4.16) we have $\int_0^t \widetilde{k_{in}}(s)\,ds = 0$ and therefore

$$y(t) = y^0 - \int_0^t \int_0^\infty \widetilde{k_{in}}(s-\tau) l(\tau)\,d\tau\,ds\,.$$

Changing the integration order gives

$$y(t) = y^0 - \int_0^\infty \int_0^t \widetilde{k_{in}}(s-\tau) l(\tau)\,ds\,d\tau = y^0 - \int_0^\infty \int_0^t \widetilde{k_{in}}(s-\tau)\,ds\,l(\tau)\,d\tau.$$

Because of the definition in (4.16) we obtain for the upper limit of the inner integral

$$y(t) = y^0 - \int_0^\infty \int_0^\tau \widetilde{k_{in}}(s-\tau)\,ds\,l(\tau)\,d\tau$$

and further

$$y(t) = y^0 - \int_0^\infty \int_0^\tau k_{in}(s-\tau)\,ds\,l(\tau)\,d\tau.$$

With the claim (4.17) we obtain for the limit

$$\lim_{t\to\infty} y(t) = \lim_{t\to\infty} \left(y^0 - \int_0^\infty \int_0^\tau k_{in}(s-\tau)\,ds\,l(\tau)\,d\tau \right) = 0$$

and therefore

$$y^0 = \int_0^\infty \int_{-\tau}^0 k_{in}(\xi)\,d\xi\,l(\tau)\,d\tau\,.$$

□

The solution of (4.14)-(4.15) reads

$$y(t) = \int_0^\infty \int_{t-\tau}^t k_{in}(s)\,ds\,l(\tau)\,d\tau$$

because

$$\begin{aligned}\frac{d}{dt}y(t) &= \int_0^\infty \left(\frac{d}{dt} \int_{t-\tau}^t k_{in}(s)\,ds \right) l(\tau)\,d\tau = \int_0^\infty \left(k_{in}(t) - k_{in}(t-\tau)\right) l(\tau)\,d\tau \\ &= \int_0^\infty k_{in}(t) l(\tau)\,d\tau - \int_0^\infty k_{in}(t-\tau) l(\tau)\,d\tau = k_{in}(t) - \left(k_{in} * l\right)(t)\,.\end{aligned}$$

For constant past, $k_{in}(s) = k_{in}^*$ for $s \leq 0$, we obtain with (4.15)

$$y^0 = \int_0^\infty \int_{-\tau}^0 k_{in}^* \, ds \, l(\tau) \, d\tau = k_{in}^* \int_0^\infty \tau l(\tau) \, d\tau = k_{in}^* T \, . \tag{4.19}$$

The result (4.19) for constant production in the past was also derived and published in [KWJ06].

Finally, we remark that in the overview article from Krzyzanski and Perez-Ruixo [KPR12] from 2012, LSMs with constant past in the context of indirect response are summarized.

4.4 General relationship between transit compartments and lifespan models

In this section we present an important relationship between transit compartments and lifespan models. Roughly spoken, we show that if the number of compartments tends to infinity and the parameter

$$T = \frac{n}{k}$$

is fixed, then in the limit the sum of all compartments is a lifespan model with constant lifespan $T > 0$.

An initial result was firstly presented by Krzyzanski in 2011, see [Krz11]. He investigated equal initial values for the TCM and constant past for the LSM.

Inspired by that work, we generalized the result to arbitrary initial values and non-constant past. The result of this section was established together with Prof. Dr. J. Schropp and is accepted for publication in the Journal of Pharmacokinetics and Pharmacodynamics, see [KS12].

For that purpose we introduce properties (see e.g. [LC98]) of the gamma probability function (PDF)

$$g_k^j(t) = \frac{k^j t^{j-1}}{(j-1)!} \exp(-kt) \, , \quad t \geq 0$$

for all $j \in \mathbb{N}$ and $k > 0$.

- Normalized integral:

$$\int_0^\infty g_k^j(s) \, ds = 1$$

- Expected value:

$$\int_0^\infty s \cdot g_k^j(s) \, ds = \frac{j}{k}$$

- Variance:
$$\int_0^\infty s^2 \cdot g_k^j(s)\, ds - \left(\int_0^\infty s \cdot g_k^j(s)\, ds \right)^2 = \frac{j}{k^2}$$
- Summation:
$$\sum_{l=1}^{j} \frac{1}{k} g_k^l(t) = 1 - \int_0^t g_k^j(s)\, ds$$

Now let $n = j$ and $k = \frac{n}{T}$ for a given $T > 0$. Then we obtain for the expected value

$$\lim_{n\to\infty} \frac{n}{k} = \lim_{n\to\infty} T = T$$

and for the variance

$$\lim_{n\to\infty} \frac{n}{k^2} = \lim_{n\to\infty} \frac{T^2}{n} = 0\,.$$

Hence, the gamma PDF $g_{\frac{n}{T}}^n(t)$ acts for $n \to \infty$ in the limit as the Dirac delta function (DDF) on integrable functions $h : \mathbb{R} \to \mathbb{R}$ via

$$\int_{-\infty}^{a} h(s) \cdot \delta(s - T)\, ds = \begin{cases} h(T) & \text{for } a > T \\ 0 & \text{for } a < T \end{cases}, \tag{4.20}$$

$a \in \mathbb{R}$ arbitrary.

Theorem 4.4.1
Consider the transit compartment model

$$x_1'(t) = k_{in}(t) - k \cdot x_1(t)\,, \qquad x_1(0) = x_1^0 \geq 0 \tag{4.21}$$
$$x_2'(t) = k \cdot x_1(t) - k \cdot x_2(t)\,, \qquad x_2(0) = x_2^0 \geq 0 \tag{4.22}$$
$$\vdots$$
$$x_n'(t) = k \cdot x_{n-1}(t) - k \cdot x_n(t)\,, \quad x_n(0) = x_n^0 \geq 0 \tag{4.23}$$

where $k_{in} : \mathbb{R} \to \mathbb{R}_{\geq 0}$ is a piecewise continuous and bounded function with finite many discontinuity points and $k \in \mathbb{R}_{>0}$ denotes the transit rate. Let $f : [0, 1] \to \mathbb{R}_{\geq 0}$ be an arbitrary piecewise continuous function (called initial density function) with $f(0) = k_{in}(0)$. Assume that the initial values of (4.21)-(4.23) satisfy

$$x_i(0) = \frac{1}{k} f\left(\frac{i}{n}\right) \quad \text{for } i = 1, ..., n\,. \tag{4.24}$$

Let

$$T = \frac{n}{k} > 0$$

be an arbitrary but fixed value where n and k are the TCM related parameter of (4.21)-(4.23). Further consider the total population based on (4.21)-(4.23)

$$y_n(t) = x_1(t) + \cdots + x_n(t) \tag{4.25}$$

where

$$y_n'(t) = x_1'(t) + \cdots + x_n'(t) = k_{in}(t) - k \cdot x_n(t)\,.$$

Then one obtains the limiting behavior

$$\lim_{n\to\infty} k \cdot x_n(t) = k_{in}(t-T) \quad \textit{for } t \geq 0 \tag{4.26}$$

and, as a consequence, the limit

$$y(t) = \lim_{n\to\infty} y_n(t) \quad \textit{for } t \geq 0 \tag{4.27}$$

fulfills the lifespan model

$$\begin{aligned} y'(t) &= k_{in}(t) - \lim_{n\to\infty} k \cdot x_n(t) \\ &= k_{in}(t) - k_{in}(t-T)\,, \qquad t \geq 0\,, \quad y(0) = y^0 \end{aligned} \tag{4.28}$$

with a constant lifespan T, provided the input function k_{in} satisfies

$$k_{in}(t) = f\left(-\frac{t}{T}\right) \quad \textit{for } -T \leq t \leq 0\,. \tag{4.29}$$

The initial value of (4.28) reads

$$y^0 = \lim_{n\to\infty} \sum_{i=1}^{n} x_i(0) = T \int_0^1 f(s)ds\,. \tag{4.30}$$

Proof: The TCM (4.21)-(4.23) with initial values (4.24) reads

$$x'(t) = \underbrace{\begin{pmatrix} -k & & & 0 \\ k & -k & & \\ & \ddots & \ddots & \\ 0 & & k & -k \end{pmatrix}}_{=:A} \cdot x(t) + k_{in}(t) \cdot e_1\,, \quad x(0) = \frac{1}{k} \cdot \hat{x}^0 \tag{4.31}$$

with

$$\hat{x}_i^0 = f\left(\frac{i}{n}\right) \quad \text{for } i = 1, ..., n\,. \tag{4.32}$$

With Remark 4.2.1 we obtain

$$x(t) = X(t) \cdot \frac{1}{k} \cdot \hat{x}^0 + \int_0^t X(t-s) \cdot k_{in}(s) \cdot e_1\, ds \;\in \mathbb{R}^n \tag{4.33}$$

where

$$X(t) = \frac{1}{k} \cdot \begin{pmatrix} g_k^1(t) & & & \\ g_k^2(t) & g_k^1(t) & & \\ \vdots & & \ddots & \\ g_k^n(t) & \ldots & g_k^2(t) & g_k^1(t) \end{pmatrix} \in \mathbb{R}^{n,n}\,. \tag{4.34}$$

The main idea of the proof is to rewrite $k \cdot x_n(t)$ as sum of two functions $v(t)$ and $w(t)$ which are itself solutions of linear differential equations.

The first differential equation with initial values depending on f reads

$$v'(t) = A \cdot v(t) \tag{4.35}$$

with

$$v(0) = v^0 = \hat{x}^0 \in \mathbb{R}^n \text{ where } \hat{x}_i^0 = f\left(\frac{i}{n}\right) \tag{4.36}$$

for $i = 1, ..., n$. The solution is

$$v(t) = X(t) \cdot \hat{x}^0 \,. \tag{4.37}$$

The second system includes the input function k_{in} and reads

$$w'(t) = A \cdot w(t) + k \cdot k_{in}(t) \cdot e_1 \,, \quad w(0) = 0 \in \mathbb{R}^n \tag{4.38}$$

with the solution

$$w(t) = \int_0^t k \cdot k_{in}(s) \cdot X(t-s) \cdot e_1 \, ds \,. \tag{4.39}$$

Obviously, by comparing the solution representation (4.33) of the TCM with (4.37) and (4.39) we obtain

$$k \cdot x(t) = v(t) + w(t) \quad \text{ for } t \geq 0 \,. \tag{4.40}$$

In order to investigate the limiting behavior of (4.40) for $n \to \infty$, we analyze the limiting behavior of $v(t)$ and $w(t)$ for $n \to \infty$.

We divide the proof in several parts:

Part 1: First of all we calculate the n-th component of $w(t)$. (4.39) can be written as

$$\begin{aligned} w(t) &= \int_0^t k \cdot k_{in}(s) \cdot X(t-s) \cdot e_1 \, ds = \int_0^t k \cdot k_{in}(s) \cdot \frac{1}{k} \cdot \begin{pmatrix} g_k^1(t-s) \\ g_k^2(t-s) \\ \vdots \\ g_k^n(t-s) \end{pmatrix} ds \\ &= \int_0^t k_{in}(s) \cdot \begin{pmatrix} g_k^1(t-s) \\ g_k^2(t-s) \\ \vdots \\ g_k^n(t-s) \end{pmatrix} ds = \int_0^t k_{in}(t-s) \cdot \begin{pmatrix} g_k^1(s) \\ g_k^2(s) \\ \vdots \\ g_k^n(s) \end{pmatrix} ds \,. \end{aligned}$$

Hence, the n-th component of $w(t)$ with $k = \frac{n}{T}$ is

$$w_n(t) = \int_0^t k_{in}(t-s) \cdot g^n_{\frac{n}{T}}(s)\, ds\,. \tag{4.41}$$

Part 2: A direct investigation of (4.37) is difficult because of the arbitrary initial values (4.36). Therefore, the idea is to prove the result indirect by considering a help differential equation with constant initial values. The system we have in mind is

$$u'(t) = A \cdot u(t) + k \cdot \widetilde{k_{in}}(t) \cdot e_1\,, \quad u(-T) = 0 \tag{4.42}$$

with

$$\widetilde{k_{in}}(t) = \begin{cases} f\left(-\frac{t}{T}\right) & \text{for } -T \le t \le 0 \\ 0 & \text{for } t > 0 \end{cases}\,. \tag{4.43}$$

Hence, in (4.42)-(4.43) the past $[-T, 0]$ is set in relationship with f. Note that according to (4.29) we have

$$\widetilde{k_{in}}(\xi) = k_{in}(\xi) \quad \text{for } \xi \in [-T, 0]\,. \tag{4.44}$$

The solution of (4.42) reads

$$u(t) = \int_{-T}^t X(t-s) \cdot k \cdot \widetilde{k_{in}}(s) \cdot e_1\, ds$$

and according to (4.34) the j-th component is

$$u_j(t) = \int_{-T}^t g^j_k(t-s) \cdot \widetilde{k_{in}}(s)\, ds = \int_0^{t+T} g^j_k(s) \cdot \widetilde{k_{in}}(t-s)\, ds \tag{4.45}$$

for $j = 1, ..., n$. Note that

$$||X(t)||_\infty = \sum_{j=1}^n \frac{1}{k} g^j_k(t) = 1 - \int_0^t g^n_k(s)\, ds \le 1$$

for $t \ge 0$ uniformly for $n \in \mathbb{N}$. We can compute for $t \ge 0$

$$\begin{aligned} |v_n(t) - u_n(t)| &= |e_n^T X(t) v(0) - e_n^T X(t) u(0)| \\ &\le \underbrace{||e_n^T||_\infty}_{=1} \cdot \underbrace{||X(t)||_\infty}_{\le 1} \cdot ||v(0) - u(0)||_\infty \\ &\le ||v(0) - u(0)||_\infty \end{aligned} \tag{4.46}$$

for any $n \in \mathbb{N}$ where $e_n = (0, 0, ..., 0, 1)^T \in \mathbb{R}^n$. The next step is to prove

$$\lim_{n\to\infty} u_n(t) = \lim_{n\to\infty} v_n(t) \quad \text{for } t \geq 0\,. \tag{4.47}$$

According to (4.46) it is sufficient to show the convergence of the initial values $v(0), u(0) \in \mathbb{R}^n$ towards the same limit as $n \to \infty$.

To that purpose we use a series of natural numbers $(j_n)_{n\in\mathbb{N}}$ describing the index of the compartments. Further let $z \in (0, 1]$ be fixed but arbitrary. This is necessary to show that the limit of the initial values is equal, compare (4.36) where $0 < i/n \leq 1$ for $i = 1, ..., n$. We choose $(j_n)_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} \frac{j_n}{n} = z$. Then with $k = \frac{n}{T}$

$$u_{j_n}(t) = \int_0^{t+T} g^{j_n}_{\frac{n}{T}}(s)\widetilde{k_{in}}(t-s)\,ds\,.$$

Because the expected value of a random variable $Y(n)$ with PDF g^n_k is $\mathbb{E}[Y(n)] = \frac{n}{k}$, we immediately obtain for the random variable $Z(n)$ with the PDF $g^{j_n}_{\frac{n}{T}}$ that

$$\lim_{n\to\infty} \mathbb{E}[Z(n)] = \lim_{n\to\infty} \frac{j_n T}{n} = zT$$

and

$$\lim_{n\to\infty} \mathrm{Var}[Z(n)] = \lim_{n\to\infty} \frac{j_n T^2}{n^2} = 0$$

hold. Thus we obtain

$$\begin{aligned}\lim_{n\to\infty} u_{j_n}(t) = \lim_{n\to\infty} \int_0^{t+T} g^{j_n}_{\frac{n}{T}}(s)\widetilde{k_{in}}(t-s)\,ds &= \int_0^{t+T} \widetilde{k_{in}}(t-s)\cdot\delta(s-zT)\,ds \\ &= \begin{cases} \widetilde{k_{in}}(t-zT) & \text{for } t+T > zT \\ 0 & \text{for } t+T < zT \end{cases}. \end{aligned} \tag{4.48}$$

In the case $t = 0$ with (4.29),(4.44) this leads to

$$\lim_{n\to\infty} u_{j_n}(0) = \widetilde{k_{in}}(-zT) = f(z) \quad \text{for } 0 < z \leq 1\,.$$

On the other hand we have

$$\lim_{n\to\infty} v_{j_n}(0) = \lim_{n\to\infty} f\left(\frac{j_n}{n}\right) = f(z) \quad \text{for } 0 < z \leq 1$$

and (4.47) is shown.

Part 3: We are now in the situation to show our main result. Using (4.41), (4.45) with $j = n$ and (4.47) we can compute

$$\begin{aligned}\lim_{n\to\infty} k \cdot x_n(t) &= \lim_{n\to\infty} (w_n(t) + v_n(t)) = \lim_{n\to\infty} (w_n(t) + u_n(t)) \\ &= \lim_{n\to\infty} \Big(\int_0^t k_{in}(t-s) g_{\frac{n}{T}}^n(s)\, ds \\ &\qquad + \int_0^t \underbrace{\widetilde{k_{in}}(t-s)}_{=0} g_{\frac{n}{T}}^n(s)\, ds + \int_t^{t+T} \underbrace{\widetilde{k_{in}}(t-s)}_{=k_{in}(t-s)} g_{\frac{n}{T}}^n(s)\, ds \Big)\,.\end{aligned}$$

Note that the relations for $\widetilde{k_{in}}$ are valid because of (4.43) and (4.44). Hence,

$$\begin{aligned}\lim_{n\to\infty} k \cdot x_n(t) &= \lim_{n\to\infty} \Big(\int_0^t k_{in}(t-s) g_{\frac{n}{T}}^n(s)\, ds + \int_t^{t+T} k_{in}(t-s) g_{\frac{n}{T}}^n(s)\, ds \Big) \\ &= \lim_{n\to\infty} \int_0^{t+T} k_{in}(t-s) g_{\frac{n}{T}}^n(s)\, ds \\ &= \int_0^{t+T} k_{in}(t-s)\delta(s-T)\, ds \\ &= k_{in}(t-T) \quad \text{for } t > 0 \end{aligned} \tag{4.49}$$

because of (4.20). For $t = 0$ we obtain with (4.24) and (4.29) directly

$$\lim_{n\to\infty} k \cdot x_n(0) = \lim_{n\to\infty} k \cdot \frac{1}{k} \cdot f\left(\frac{n}{n}\right) = f(1) = k_{in}(-T)\,. \tag{4.50}$$

Combining (4.49) and (4.50) yields

$$\lim_{t\to\infty} k \cdot x_n(t) = k_{in}(t-T) \quad \text{for } t \geq 0\,.$$

Part 4: The solution of (4.28) reads

$$y(t) = \int_{t-T}^{t} k_{in}(s) ds$$

and

$$y(0) = \lim_{n\to\infty} \sum_{i=1}^{n} x_i(0) = \lim_{n\to\infty} \sum_{i=1}^{n} \frac{1}{k} \cdot f\left(\frac{i}{n}\right) = T \cdot \lim_{n\to\infty} \sum_{i=1}^{n} \frac{1}{n} \cdot f\left(\frac{i}{n}\right) = T \cdot \int_0^1 f(s)\, ds\,.$$

□

It is remarked that the idea of splitting into two linear differential equations was contributed by Prof. Dr. J. Schropp.

In Chapter 6 the result will be applied to rewrite a PKPD model consisting of transit compartments into a model with a lifespan equation. In Chapter 7 we go vice versa and rewrite a model originally developed with a lifespan equation by transit compartments.

4.5 Classification of models with lifespan

In this section we categorize typical PKPD models based on a lifespan approach.

Let x_1 be a precursor or driver compartment which provokes a state x_2. All models in this section are coupled from x_1 towards x_2 in such a way that the outflow from x_1 is equal or delayed to the inflow of x_2. If also a back coupling from x_2 to x_1 exists, we call this a feedback. We describe by g a growth mechanism. In the following classification we set the rates time-independent.

Model I : Lifespan model with precursor and no feedback

Consider the model

$$x_1'(t) = g(t, x_1(t)) - k_1 \cdot x_1(t)\,, \qquad x_1(s) = x_1^0(s) \quad \text{for } -T \le s \le 0 \tag{4.51}$$
$$x_2'(t) = k_1 \cdot x_1(t) - k_1 \cdot x_1(t-T)\,, \qquad x_2(0) = x_2^0\,. \tag{4.52}$$

Note that the solution of (4.52) could be computed a posteriori by pure integration

$$x_2(t) = k_1 \int_{t-T}^{t} x_1(s)\,ds - k_1 \int_{-T}^{0} x_1^0(s)\,ds + x_2^0\,.$$

A structural equal model like (4.51)-(4.52) was developed and applied by Perez-Ruixo et al in 2005 [PRKC$^+$05] for cells in the context of indirect response (stimulation / inhibition of rate constants).

Model II : Delayed driver with no feedback

The development of x_2 is delayed due to the driver x_1. We obtain

$$x_1'(t) = g(t, x_1(t)) - k_1 \cdot x_1(t)\,, \qquad x_1(s) = x_1^0(s) \quad \text{for } -T \le s \le 0 \tag{4.53}$$
$$x_2'(t) = k_1 \cdot x_1(t-T) - k_2 \cdot x_2(t)\,, \qquad x_2(0) = x_2^0\,. \tag{4.54}$$

From a mechanistic point of view the delayed inflow into x_2 could be explained by a lifespan equation with inflow $k_{in}(t) = k_1 x_1(t)$, more precisely, if one adds

$$y'(t) = k_1 \cdot x_1(t) - k_1 \cdot x_1(t-T)\,, \qquad y(0) = y^0 \tag{4.55}$$

to (4.53)-(4.54). In Chapter 7 we present a structural similar model, build based on pharmacological assumptions for arthritis development.

Model III : Lifespan model with precursor and feedback

The model (4.51)-(4.52) reads with a feedback

$$x_1'(t) = g(t, x_1(t), x_2(t)) - k_1 \cdot x_1(t)\,, \qquad x_1(s) = x_1^0(s) \quad \text{for } -T \le s \le 0 \tag{4.56}$$
$$x_2'(t) = k_1 \cdot x_1(t) - k_1 \cdot x_1(t-T)\,, \qquad x_2(0) = x_2^0\,. \tag{4.57}$$

We will see in Chapter 6 the derivation of structure (4.56)-(4.57) based on a typical tumor growth PKPD model.

Model IV : Delayed driver with feedback

The model (4.53)-(4.54) with feedback reads

$$x_1'(t) = g(t, x_1(t), x_2(t)) - k_1 \cdot x_1(t)\,, \qquad x_1(s) = x_1^0(s) \quad \text{for } -T \le s \le 0 \tag{4.58}$$
$$x_2'(t) = k_1 \cdot x_1(t-T) - k_2 \cdot x_2(t)\,, \qquad x_2(0) = x_2^0\,. \tag{4.59}$$

A schematic representation of models I - IV is presented in Figure 4.4. The classification between feedback and no feedback is of importance when rewriting the equations by the method of steps, see [Dri77].

No feedback: Reformulation of model I and II as ordinary differential equation

Model I (4.51)-(4.52) and II (4.53)-(4.54) with no feedback or more general, models with no appearance of the delayed state in the corresponding equations are covered by the structure

$$x_1'(t) = f_1(t, x_1(t))\,, \qquad x_1(s) = x_1^0(s) \quad \text{for } -T \le s \le 0 \tag{4.60}$$
$$x_2'(t) = f_2(t, x_1(t), x_1(t-T), x_2(t))\,, \qquad x_2(0) = x_2^0\,. \tag{4.61}$$

For (4.60)-(4.61) the method of steps reduces to exactly two steps.

Step 1: $0 \le t \le T$
We include the explicitly known initial function x_1^0 for the delayed state and obtain

$$x_1'(t) = f_1(t, x_1(t))\,, \qquad x_1(0) = x_1^0(0) \tag{4.62}$$
$$x_2'(t) = f_2(t, x_1(t), x_1^0(t-T), x_2(t))\,, \qquad x_2(0) = x_2^0\,. \tag{4.63}$$

Step 2: $T \le t$
We add a further ordinary differential equation for x_3 which describes the state of x_1

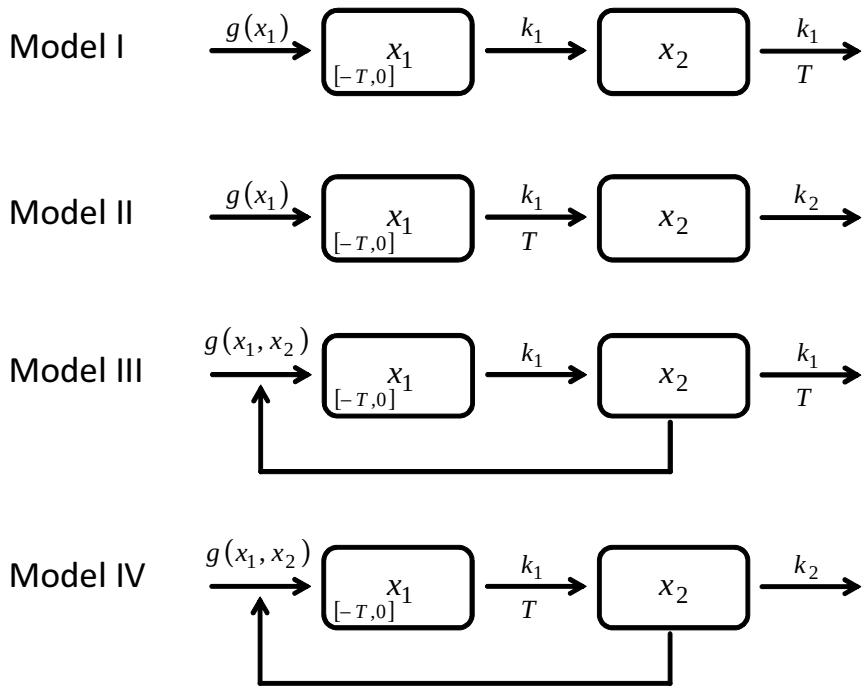


Figure 4.4: *Schematic representation of the models I - IV. The interval* $[-T, 0]$ *in compartment* x_1 *denotes the existence of a past. The parameter* T *denotes where the delay takes place.*

before $t - T$ time units, namely $x_3(t) = x_1(t - T)$ and denote by (x_1^T, x_2^T) the value at the time point T. Then the system reads

$$x_1'(t) = f_1(t, x_1(t)), \qquad x_1(T) = x_1^T \tag{4.64}$$
$$x_2'(t) = f_2(t, x_1(t), x_3(t), x_2(t)), \qquad x_2(T) = x_2^T \tag{4.65}$$
$$x_3'(t) = f_1(t - T, x_3(t)), \qquad x_3(T) = x_1^0(0). \tag{4.66}$$

Feedback: Reformulation of model III and IV as ordinary differential equation

Model III (4.56)-(4.57) and IV (4.58)-(4.59) with feedback are covered by the structure

$$x_1'(t) = f_1(t, x_1(t), x_2(t)), \qquad x_1(s) = x_1^0(s) \quad \text{for } -T \le s \le 0 \tag{4.67}$$
$$x_2'(t) = f_2(t, x_1(t), x_1(t - T), x_2(t)), \qquad x_2(0) = x_2^0. \tag{4.68}$$

Due to the feedback, the method of steps does not break off after two steps.

Consider the first step $0 \le t \le T$, then we directly obtain

$$x_1'(t) = f_1(t, x_1(t), x_2(t)), \qquad x_1(0) = x_1^0(0) \tag{4.69}$$
$$x_2'(t) = f_2(t, x_1(t), x_1^0(t - T), x_2(t)), \qquad x_2(0) = x_2^0. \tag{4.70}$$

In the second step we again use a delayed version of x_1 based on additional ordinary differential equations and obtain

$$x_1'(t) = f_1(t, x_1(t), x_2(t)), \qquad x_1(T) = x_1^T \tag{4.71}$$
$$x_2'(t) = f_2(t, x_1(t), x_3(t), x_2(t)), \qquad x_2(T) = x_2^T \tag{4.72}$$
$$x_3'(t) = f_1(t - T, x_3(t), x_4(t)), \qquad x_3(T) = x_1^0(0) \tag{4.73}$$
$$x_4'(t) = f_2(t - T, x_4(t), x_1^0(t - 2T), x_4(t)), \qquad x_4(T) = x_2^0. \tag{4.74}$$

Because in (4.73) the initial function is used, the interval for the second step is $T \leq t \leq 2T$. Hence, the method of steps does not break off and the number of steps depends on the ratio between the delay parameter and the length of the integration interval.

4.6 Modeling of the drug effect

We introduced models to describe the pharmacokinetics $c(t)$ of a drug in Chapter 3. The PK models have the property

$$\lim_{t\to\infty} c(t) = 0.$$

In this section we focus on drug-effect models describing the pharmacological effect of a drug at the target. Roughly spoken, such targets could be tumor cells or more indirectly, messengers in the body whose perturbation leads to an effect on the disease.

We denote a drug-effect term by

$$e(\sigma, c(t)) \tag{4.75}$$

where we call σ the drug-related parameter. The only assumption on (4.75) is

$$\lim_{c\to 0} e(\sigma, c) = 0$$

that means, if the drug concentration vanish, then the pharmacological effect will also vanish.

The simplest approach for a drug-effect term is a linear model

$$e_1(k_{pot}, c(t)) = k_{pot} \cdot c(t) \tag{4.76}$$

where k_{pot} describes the drug potency. Such a parameter could be used to rank different compounds (drug candidates) among each other in preclinical screening. The approach (4.76) is also useful if only few dosing groups are available for a simultaneous fit. In case of more dosing groups this model is only locally true around a certain amount of *dose* because the effect of a drug is in the majority of cases only linear in a small range of different doses. We will present the application of (4.76) in Chapter 6.

The classical drug-receptor binding theory based on the Michaelis-Menten theory, see

[Mur89] or [Boh06], states that the molecules of a drug bind to the appropriate target receptor and that the amount of binding possibilities at the receptor is limited. Therefore, the effect of a drug will saturate and more drug will not lead to more effect. Following the Michaelis-Menten theory, see e.g. [MI06], results in a drug-effect term for the pharmacological effect of the form

$$e_2(\sigma, c(t)) = \frac{E_{max} \cdot c(t)^h}{EC_{50}^h + c(t)^h} \tag{4.77}$$

with $\sigma = (E_{max}, EC_{50}, h)$. E_{max} is the maximal effect, EC_{50} is the concentration which is needed to produce the half-maximal effect and h is the Hill coefficient. In Figure 4.5 a plot of (4.77) is presented. The model (4.77) is one of the basic principles in PKPD modeling and is called the non-linear approach.

However, sometimes the drug effect relationship is highly non-linear and (4.77) will fail to describe the data. Therefore, we develop a new approach for drug-effect modeling covering the following situations:

- The drug-effect term could describe highly non-linear data.
- The drug-effect term is able to describe non-monotonic *dose* effect behavior, that means, if the effect for a lower dose is higher than for a higher dose. We are aware that this is a very special and seldom situation. However, we were faced with such a situation, see Chapter 7, experiment A.

The key to this new effect term is the observation that the potency parameter k_{pot} in the linear approach (4.76) obtained by fits from the control group together with one dosing group could depend on the drug level $c(t)$. In the experiments from Chapter 7 the potency k_{pot} decreases exponentially with an increasing level of $c(t)$. Therefore, we amend (4.76) by a term decreasing with c and obtain

$$k_{pot}(c) = \sigma_3 + \sigma_1 \exp(-\sigma_2 c)\,. \tag{4.78}$$

Note that (4.78) satisfies an inflow/outflow model (see Section 4.1) and therefore, is based on pharmacological first principles. By multiplying with $c(t)$ we obtain the drug effect term

$$e_3(\sigma, c(t)) = k_{pot}(c(t))c(t) = (\sigma_1 \exp(-\sigma_2 c(t)) + \sigma_3)\, c(t) \tag{4.79}$$

with $\sigma_1, \sigma_2, \sigma_3 > 0$. The parameter σ_3 describes the base potency of the drug and σ_1, σ_2 are potency decreasing parameters. See Figure 4.5 for a plot of (4.79).

4.7 Discussion and outlook

In this chapter we presented and summarized important model figures which we will apply in Chapters 6 and 7. In Section 4.1 we introduced the typical inflow / outflow model and in Section 4.2 the transit compartment approach. We showed in Section 4.4 an important

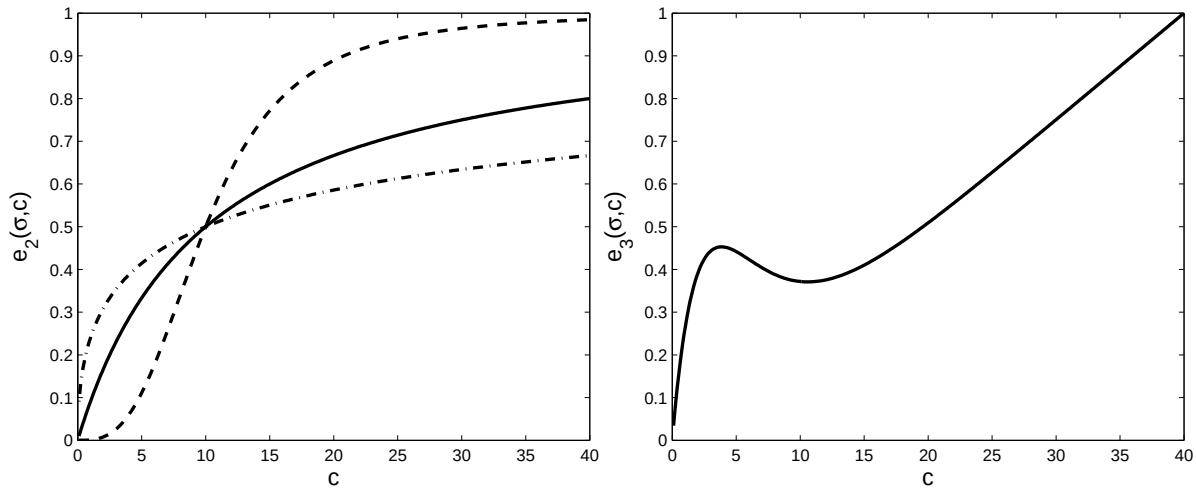


Figure 4.5: *In the left panel a plot of (4.77) with the parameters* $E_{max} = 1$, $EC_{50} = 10$ *and with* $h = 0.5$ *(dash dotted line),1 (solid line) and 3 (dashed line) is presented. In the right panel a plot of (4.79) with the parameter* $\sigma_1 = 0.328$, $\sigma_2 = 0.328$ *and* $\sigma_3 = 0.025$ *is presented.*

general relationship between transit compartments and lifespan models (Section 4.3) with constant delay. This result is also published in [KS12] coauthored by Prof. Dr. J. Schropp. Whether a result similar to Theorem 4.4.1 also holds in the case of distributed lifespan is an interesting open problem. In Section 4.5 we categorized typical models with a lifespan. Finally, in Section 4.6 we presented drug-effect terms.

Chapter 5

Point and Interval Estimation

This chapter deals with statistical properties of point estimators for model parameters. In application, one is not only interested in parameter estimates but also in the question, how reliable is the estimate obtained from a finite sample size?

Therefore, we calculate statistical characteristics for estimates, like confidence intervals and coefficient of variations. A confidence interval could be understood as follows. The confidence interval is a region around an estimate that indicates with a given probability the location of the true value for the parameter. The coefficient of variation is a normalized measure of dispersion.

For that purpose we introduce the weighted least squares sum and formulate statistical assumptions on the errors, see Sections 5.1, 5.2 and 5.3. In Section 5.4 we prove the consistency of the weighted least squares estimator. Based on this result, the asymptotically normal distribution of the weighted least squares estimator is obtained in Section 5.5. Equipped with this result we are able to construct the confidence interval and to calculate the coefficient of variation for parameter estimates in Section 5.6.

The presented proofs apply ideas from the results for ordinary least squares stated in [SW89], [DM93] and [Ame01]. Finally, we adopted the typical notation from statistics / econometrics in this chapter.

5.1 The weighted least squares sum

Our pharmacokinetic and pharmacodynamic experiments consist of measurements from individual subjects (animals) at different time points t_i. Usually, ten animals are used for each dosing group and the number of groups varies from two to four (placebo group and different drug schedules). In this work we average the measurements in every dosing group at each time point t_i denoted by y_i.

Let $t_i \in \mathbb{R}_{\geq 0}$ be fixed time points with $t_i \in [0, t_{end}]$ and y_i the corresponding averaged measurements for $i = 1, ..., n$. Let $x : \mathbb{R}_{\geq 0} \times \mathbb{R}^p_{>0} \to \mathbb{R}_{\geq 0}$ be a non-linear model with parameter $\theta \in \mathbb{R}^p_{>0}$. In our pharmacokinetics / pharmacodynamics applications the analytical representation of the model $x(t, \theta)$ is mostly not known because the model is based on differential equations.

The non-linear regression model reads (see [SW89])

$$y_i = x(t_i, \theta^*) + \varepsilon_i \tag{5.1}$$

where $\varepsilon_i \in \mathbb{R}$, $i = 1, ..., n$ are called errors and describe all kind of uncertainty. The unknown true parameter is denoted by θ^*.

The weighted least squares sum $S_n : \mathbb{R}^p_{>0} \to \mathbb{R}_{\geq 0}$ reads

$$S_n(\theta) = \sum_{i=1}^{n} w_i \left(y_i - x(t_i, \theta)\right)^2 \tag{5.2}$$

with bounded weights

(A1) $0 < w_i \leq \omega$ for $i = 1, ..., n$ for some $\omega \in \mathbb{R}_{>0}$.

The least squares estimator is defined as follows, compare with [Ame01].

Definition 5.1.1
Let $\Theta \subset \mathbb{R}^p$ be compact including θ^*. The map

$$\widehat{\theta}_n := \widehat{\theta}(y_1, y_2, ..., y_n) : \mathbb{R}^n \to \Theta$$

is called the weighted least squares (WLS) estimator of the model parameter θ, if it is the root of the normal equations

$$\frac{\partial S_n}{\partial \theta} = 0\,.$$

Note that this definition is in a local sense. But due to our construction of PKPD models based on pharmacological principles, only one set of parameters exists in a realistic range.

The following remark presents a standard technique in statistics, see e.g. [DM93].

Remark 5.1.2
In general the limit of $S_n(\theta)$ in (5.2) is infinity for $n \to \infty$. Hence, it is a common approach to investigate the averaged sum of squares

$$\frac{1}{n} S_n(\theta) = \frac{1}{n} \sum_{i=1}^{n} w_i \left(y_i - x(t_i, \theta)\right)^2 . \tag{5.3}$$

5.2 Statistical assumptions about the errors

We present three important assumptions about the error random variables ε_i for $i = 1, ..., n$:

(A2) The errors ε_i are independent for $i = 1, ..., n$.

(A3) The expected value of the error ε_i is $\mathbb{E}[\varepsilon_i] = 0$ for $i = 1, ..., n$.

(A4) The variance of the error ε_i is $0 < \mathrm{Var}[\varepsilon_i] = \sigma_i^2 < \infty$ for $i = 1, ..., n$.

(A2) is a classical assumption and is necessary to apply a central limit theorem. However, (A2) is questionable in our applications because our data primarily describes measured disease development. Hence, it is obvious that the erorrs are not independent. More generally spoken, Seber stated in [SW89] that serially correlated errors usually arise in the fitting of growth curves. But the amount of data measured in time is also often sparse and therefore, the application of more advanced statistical techniques to overcome the independently assumption is not possible.

(A3) states that the non-linear model for mean response is correctly specified, see [DM93] or [DG95].

(A4) is a generalization of the classical Gauss-Markov framework (see e.g. [SS90]) where it is assumed that the variances are all equal. However, in applications $\mathrm{Var}[\varepsilon_i] = \sigma^2$ for $i = 1, ..., n$ is often violated. Also in our simultaneous fits different variances could arise, see for example the pharmacokinetic data for different doses in Section 3.4.

Finally, note that we do not assume a distribution for ε_i.

5.3 Choice of weights

The weights could be arbitrarily chosen in our approach. We present three typical situations for setting weights in practice.

Situation 1: If there is evidence that some measurements are less reliable than others, then with the present approach the user is able to set weights a priori.

Situation 2: In a simultaneous fit, several levels of readouts are fitted. For example, for different pharmacokinetic dosing groups (see Section 3.4) every measurement should

contribute with the same importance to the sum of squares. Hence, one approach is that the weights correspond to the inverse of the administered dose for each dosing group.

Situation 3: A general approach in statistics is to set the weights in dependency of the variance σ_i^2 in order to construct a constant variance and therefore, to deal with heteroscedasticity, see e.g. [SS90]. For that purpose let

$$\mathrm{Var}[\varepsilon_i] = \sigma_i^2 = c_i^2\sigma^2 \quad \text{with } c_i \neq 0 \text{ for } i = 1, ..., n\,.$$

Then constant variance is achieved by dividing the non-linear regression model (5.1) with c_i

$$\frac{y_i}{c_i} = \frac{x(t_i, \theta^*)}{c_i} + \frac{\varepsilon_i}{c_i}\,.$$

One obtains the sum of squares

$$\sum_{i=1}^{n}\left(\frac{\varepsilon_i}{c_i}\right)^2 = \sum_{i=1}^{n}\frac{1}{c_i^2}(y_i - x(t_i,\theta))^2 = \sum_{i=1}^{n} w_i(y_i - x(t_i,\theta))^2$$

with the weights

$$w_i = \frac{1}{c_i^2}\,.$$

5.4 Consistent weighted least squares estimator

Roughly spoken, a consistent estimator $\widehat{\theta}_n$ converges towards the true value θ^*, if the sample size n enlarges to infinity.

To formulate this statement, we introduce the concept of convergence in probability of a random variable.

Definition 5.4.1
A sequence $X_1, X_2, ...$ of vector/matrix-valued random variables is said to converge in probability to a limit random variable X, denoted by

$$\operatorname*{plim}_{n\to\infty} X_n = X\,,$$

if

$$\forall \alpha > 0 \;:\; \lim_{n\to\infty} \mathbb{P}\left[||X_n - X|| \geq \alpha\right] = 0\,.$$

In this work we deal with so-called weak consistency which is defined in the following.

Definition 5.4.2
An estimator $\widehat{\theta}_n$ of a parameter θ^* is called a weak consistent estimator, if

$$\operatorname*{plim}_{n\to\infty} \widehat{\theta}_n = \theta^*\,.$$

In [GI83] or [SW89] the following property is stated:

Remark 5.4.3
Sufficient for weak consistency of an estimator $\widehat{\theta}_n$ is that the limit of the averaged sum of squares,

$$\operatorname*{plim}_{n\to\infty} \frac{1}{n} S_n(\theta)\,, \tag{5.4}$$

is minimized at the true value θ^*.

A useful tool to prove convergence in probability is Chebyshev's inequality.

Theorem 5.4.4 (Chebyshev's inequality)
Let X be a real-valued random variable with $\mathbb{E}[X] < \infty$ and $Var[X] < \infty$. Then for any $\alpha \in \mathbb{R}_{>0}$

$$\mathbb{P}\left[|X - \mathbb{E}[X]| \geq \alpha\right] \leq \frac{Var[X]}{\alpha^2}\,.$$

We further assume:

(A5) Let $\Theta \subset \mathbb{R}^p$ compact and $\theta^* \in \Theta$. We assume $x(t,\theta)$ is continuous in θ for all $\theta \in \Theta$.

Hence, $x(t,\theta)$ is bounded in θ for $\theta \in \Theta$.

Except for the pharmacokinetics model for multiple i.v. administration, $x(t,\theta)$ is also continuous in t for $t \in [0, t_{end}]$. However, in general

(A6) $\sup\{|x(t,\theta_1) - x(t,\theta_2)| \,:\, \theta_1, \theta_2 \in \Theta\} < M_t < \infty$ for all $t \in [0, t_{end}]$.

The main result of this section is formulated in the following proposition:

Theorem 5.4.5 (Weak consistent estimator)
Let (A1)-(A6) hold. Further we assume:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} w_i \sigma_i^2 = C_1 < \infty\,, \tag{5.5}$$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} w_i^2 Var[\varepsilon_i^2] = C_2 < \infty \tag{5.6}$$

and for $\theta \in \Theta$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} w_i (x(t_i,\theta^*) - x(t_i,\theta))^2 = 0 \quad \textit{if and only if } \theta = \theta^*\,. \tag{5.7}$$

Then

$$\operatorname*{plim}_{n\to\infty} \frac{1}{n} S_n(\theta)$$

is minimized at the true value θ^*.

Proof: We rewrite the averaged sum of squares $\frac{1}{n}S_n(\theta)$ in terms of errors ε_i from (5.1)

$$\begin{aligned}\frac{1}{n}S_n(\theta) &= \frac{1}{n}\sum_{i=1}^{n} w_i \left(y_i - x(t_i,\theta)\right)^2 \\ &= \frac{1}{n}\sum_{i=1}^{n} w_i \left(y_i - x(t_i,\theta^*) + x(t_i,\theta^*) - x(t_i,\theta)\right)^2 \\ &= \frac{1}{n}\sum_{i=1}^{n} w_i \left(\varepsilon_i + x(t_i,\theta^*) - x(t_i,\theta)\right)^2 \qquad (5.8) \\ &= \frac{1}{n}\sum_{i=1}^{n} w_i\varepsilon_i^2 + \frac{2}{n}\sum_{i=1}^{n} w_i\varepsilon_i(x(t_i,\theta^*) - x(t_i,\theta)) + \frac{1}{n}\sum_{i=1}^{n} w_i \left(x(t_i,\theta^*) - x(t_i,\theta)\right)^2 \qquad (5.9)\end{aligned}$$

and investigate the convergence in probability of (5.9) term by term.

Term 1: We define

$$Y_n := \frac{1}{n}\sum_{i=1}^{n} w_i\varepsilon_i^2 \,.$$

Because of (A4) and (5.5)

$$\mathbb{E}[Y_n] = \frac{1}{n}\sum_{i=1}^{n} w_i\mathbb{E}[\varepsilon_i^2] = \frac{1}{n}\sum_{i=1}^{n} w_i\mathrm{Var}[\varepsilon_i] = \frac{1}{n}\sum_{i=1}^{n} w_i\sigma_i^2 < \infty$$

and with (5.6)

$$\mathrm{Var}[Y_n] = \frac{1}{n^2}\sum_{i=1}^{n} w_i^2\mathrm{Var}[\varepsilon_i^2] < \infty \,.$$

We show

$$\operatorname*{plim}_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} w_i\varepsilon_i^2 = \lim_{n\to\infty} \mathbb{E}[Y_n] = C_1$$

by applying Chebyshev's inequality. For all $\alpha > 0$

$$\mathbb{P}\left[|Y_n - \mathbb{E}[Y_n]| \geq \alpha\right] \leq \frac{1}{\alpha^2}\mathrm{Var}\left[Y_n\right] = \frac{1}{\alpha^2}\frac{1}{n^2}\sum_{i=1}^{n} w_i^2\mathrm{Var}\left[\varepsilon_i^2\right] \,. \qquad (5.10)$$

Now consider $n \to \infty$. Because of (5.5) and (5.6) we have for all $\alpha > 0$

$$\lim_{n\to\infty} \mathbb{P}[|Y_n - \mathbb{E}[Y_n]| \geq \alpha] = 0$$

which is the result.

Term 2: We show for a $\theta \in \Theta$

$$\operatorname*{plim}_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} w_i\varepsilon_i \left(x(t_i,\theta^*) - x(t_i,\theta)\right) = 0 \,. \qquad (5.11)$$

Note that the only stochastic source in (5.11) is ε_i for $i = 1, ..., n$. We define

$$Y_n := \frac{1}{n} \sum_{i=1}^{n} w_i \varepsilon_i \left(x(t_i, \theta^*) - x(t_i, \theta) \right) .$$

Because of (A3)

$$\mathbb{E}\left[Y_n \right] = \frac{1}{n} \sum_{i=1}^{n} w_i \mathbb{E}[\varepsilon_i] \left(x(t_i, \theta^*) - x(t_i, \theta) \right) = 0 .$$

With the assumptions we have

$$\mathrm{Var}\left[Y_n \right] = \frac{1}{n^2} \sum_{i=1}^{n} w_i^2 \sigma_i^2 \underbrace{\left(x(t_i, \theta^*) - x(t_i, \theta) \right)^2}_{< M_{t_i}^2} < \infty .$$

Chebyshev's inequality gives for all $\alpha > 0$

$$\mathbb{P}\left[|Y_n| \geq \alpha \right] \leq \frac{1}{\alpha^2} \frac{1}{n^2} \sum_{i=1}^{n} w_i^2 \sigma_i^2 \left(x(t_i, \theta^*) - x(x_i, \theta) \right)^2 . \tag{5.12}$$

With (5.5)

$$\lim_{n \to \infty} \frac{1}{n^2} \sum_{i=1}^{n} w_i^2 \sigma_i^2 \underbrace{\left(x(t_i, \theta^*) - x(t_i, \theta) \right)^2}_{< M_{t_i}^2 \leq M} \leq \lim_{n \to \infty} \frac{1}{n} \omega M \frac{1}{n} \sum_{i=1}^{n} w_i \sigma_i^2 = 0$$

and therefore, we showed (5.11).

Term 3: Apply assumption (5.7) to term 3 in (5.9).

Summarized we obtain

$$\operatorname*{plim}_{n \to \infty} \frac{1}{n} S_n(\theta) = C_1 + \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} w_i \left(x(t_i, \theta^*) - x(t_i, \theta) \right)^2$$

is minimized at $\theta = \theta^*$.

□

Finally, we shortly comment on the stated assumptions in the above proposition.

Assumption (5.7) is reasonable because it reflects that the model is not over-parameterized. That means, no arbitrary sets of parameters in a local range produce the same fitting results. This is the basis for physiological interpretation of model parameters and finally, to apply them for predictions for higher species. Because our models are build on pharmacological principles, we take special care on over-parameterization and therefore, (5.7) is fulfilled in our PKPD models.

Assumptions (5.5) and (5.6) are technical conditions. Note that these assumptions are formally different from the conditions stated in [SW89]. However, for example, choosing the weights as in situation 3 in Section 5.3, (5.5) results in

$$C_1 = \sigma^2$$

which is one of the results from the consistency observation in ordinary least squares, compare [SW89]. Hence, (5.5) is reasonable because the weights are mostly chosen in such a way that they somehow correspond to the inverse of the variance, see Section 5.3. Finally, (5.6) is actually a condition on the fourth moment of ε_i. Note that this moment exists for most distributions used in application.

Summarizing, all stated assumptions are connected to practice.

5.5 Asymptotic normality of the weighted least squares estimator

In this section we investigate the asymptotic normality of the WLS estimator $\widehat{\theta}_n$. For that purpose we first of all introduce the concept of convergence in distribution, see [DM93].

Definition 5.5.1
A sequence $X_1, X_2, ...$ of vector-valued random variables is said to converge in distribution to a limit random variable X, if

$$\lim_{n\to\infty} \mathbb{P}\left[X_n \leq Z\right] = \mathbb{P}[X \leq Z]$$

for all vectors Z such that the limiting distribution function $\mathbb{P}\left[X \leq Z'\right]$ is continuous in Z' at $Z' = Z$. One writes:

$$X_n \xrightarrow{d} X \quad \text{for } n \to \infty\,.$$

An important relationship between convergence in probability and distribution is formulated in Slutsky's theorem.

Theorem 5.5.2 (Slutsky)
Let $X_1, X_2, ...$ be a sequence of matrix-valued random variables and C a constant matrix. Further let $Y_1, Y_2, ...$ be a sequence of vector-valued random variables and Y a vector-valued random variable. If $\operatorname{plim}_{n\to\infty} X_n = C$ and $Y_n \xrightarrow{d} Y$, then

$$X_n Y_n \xrightarrow{d} CY \quad \text{for } n \to \infty\,.$$

Now we present the standard definition of an asymptotically normal distributed estimator for a sample size tending to infinity.

Definition 5.5.3
An estimator $\widehat{T}_n$ is asymptotically normal distributed, if

$$\sqrt{n}(\widehat{T}_n - T^*) \xrightarrow{d} \mathcal{N}(0, Cov) \quad \text{for } n \to \infty \tag{5.13}$$

where Cov is the finite covariance matrix.

In practice, one constructs from the limiting distribution in (5.13) the approximate distribution [Gre12] for large n denoted by

$$\widehat{T}_n \overset{a}{\sim} \mathcal{N}\left(T^*, \frac{1}{n}\mathrm{Cov}\right) . \tag{5.14}$$

The key theorem to prove asymptotic normality is the following version of the central limit theorem, compare e.g. with [Gre12].

Theorem 5.5.4 (Multivariate central limit theorem)
Let $\Gamma_1, \Gamma_2, ...$ be a sequence of independent vector-valued random variables such that $\mathbb{E}[\Gamma_i] = \mu_i$, $Cov[\Gamma_i] = Q_i$ and all mixed third moments of the multivariate distribution are finite. Let

$$\overline{\Gamma}_n = \frac{1}{n}\sum_{i=1}^{n}\Gamma_i\,, \quad \overline{\mu}_n = \frac{1}{n}\sum_{i=1}^{n}\mu_i \quad \text{and} \quad \overline{Q}_n = \frac{1}{n}\sum_{i=1}^{n}Q_i\,.$$

Assume that

$$\lim_{n\to\infty}\overline{Q}_n = Q \tag{5.15}$$

where Q is a finite, positive definite matrix and that for every i

$$\lim_{n\to\infty}(n\overline{Q}_n)^{-1}Q_i = 0\,. \tag{5.16}$$

Then

$$\sqrt{n}\left(\overline{\Gamma}_n - \overline{\mu}_n\right) \xrightarrow{d} \mathcal{N}(0, Q) \qquad \text{for } n \to \infty\,.$$

Before we formulate the main result of this section, we present another important tool. Due to [Ame01] it yields:

Theorem 5.5.5
Let $X_1, X_2, ...$ be a sequence of random variables with $X_n : \mathbb{R}^p \to \mathbb{R}$ depending on a parameter θ. Further let $X_1(\theta), X_2(\theta), ...$ converge uniformly in probability in $\theta \in D$ that means

$$\lim_{n\to\infty}\mathbb{P}[\sup_{\theta\in D}|X_n(\theta)| < \alpha] = 1 \quad \text{for any } \alpha > 0\,,$$

to a non-stochastic function $X(\theta)$. Let $\widehat{\theta}_n$ be a consistent estimator of θ^* and $X(\theta)$ continuous at θ^*. Then

$$\operatorname*{plim}_{n\to\infty} X_n(\widehat{\theta}_n) = X\left(\operatorname*{plim}_{n\to\infty}\widehat{\theta}_n\right) = X(\theta^*)\,.$$

Now we consider our situation.

Let U_{θ^*} be an open neighborhood of θ^* in Θ. In order to prove asymptotic normality we need the following condition:

- Let $x(t,\theta)$ be two times partial differentiable with respect to θ for $t \in [0, t_{end}]$, $\theta \in U_{\theta^*}$ and the partial derivatives are continuous.

For a better notation we introduce the following matrices. Let

$$P(\theta) := \begin{pmatrix} \frac{\partial x(t_1,\theta)}{\partial \theta_1} & \cdots & \frac{\partial x(t_1,\theta)}{\partial \theta_p} \\ \vdots & & \vdots \\ \frac{\partial x(t_n,\theta)}{\partial \theta_1} & \cdots & \frac{\partial x(t_n,\theta)}{\partial \theta_p} \end{pmatrix} \in \mathbb{R}^{n,p}$$

be the Jacobian matrix, $W :=\text{diag}(w_1, w_2, ..., w_n) \in \mathbb{R}^{n,n}$ a diagonal matrix consisting of the weights and $G :=\text{diag}(\sigma_1^2, \sigma_2^2, ..., \sigma_n^2) \in \mathbb{R}^{n,n}$ a diagonal matrix describing the variances. Note the dependency on the sample size n in the dimension of the matrices.

The following proposition shows that the WLS estimator $\widehat{\theta}_n$ is asymptotically normal distributed under certain conditions.

Theorem 5.5.6 (Asymptotic normality)
Let $\widehat{\theta}_n$ be a consistent WLS estimator. We assume:

- *The matrix function series*

$$\overline{X}_n(\theta) := \frac{1}{n} \sum_{i=1}^{n} w_i \frac{\partial x(t_i,\theta)}{\partial \theta} \frac{\partial x(t_i,\theta)}{\partial \theta^T} \in \mathbb{R}^{p,p} \tag{5.17}$$

  converges uniformly for all $\theta \in U_{\theta^*}$ towards a regular matrix $\Omega(\theta) \in \mathbb{R}^{p,p}$ for $n \to \infty$.

- *The matrix function series*

$$\overline{Y}_n(\theta_1, \theta_2) := \frac{1}{n} \sum_{i=1}^{n} w_i x(t_i, \theta_1) \frac{\partial^2 x(t_i,\theta_2)}{\partial \theta \partial \theta^T} \in \mathbb{R}^{p,p} \tag{5.18}$$

  converges uniformly for all $\theta_1, \theta_2 \in U_{\theta^*}$ towards a finite matrix Y for $n \to \infty$.

- *The independent stochastic sequence*

$$\Gamma_n = -w_n \varepsilon_n \frac{\partial x(t_n,\theta^*)}{\partial \theta} \in \mathbb{R}^p \tag{5.19}$$

  fulfills the assumptions of Theorem 5.5.4.

Then $\widehat{\theta}_n$ is asymptotically normal distributed, that means

$$\sqrt{n}\left(\widehat{\theta}_n - \theta^*\right) \xrightarrow{d} \mathcal{N}(0, Cov) \quad \text{for } n \to \infty$$

with $Cov = \Omega(\theta^*)^{-1}\Sigma(\theta^*)\Omega(\theta^*)^{-1}$ where

$$\Omega(\theta^*) = \operatorname*{plim}_{n\to\infty} \frac{1}{n} P^T(\widehat{\theta}_n) W P(\widehat{\theta}_n) \tag{5.20}$$

and

$$\Sigma(\theta^*) = \lim_{n\to\infty} \frac{1}{n} P^T(\theta^*) W G W P(\theta^*)\,. \tag{5.21}$$

Proof: The sum of squares reads

$$S_n(\theta) = \sum_{i=1}^{n} w_i \left(y_i - x(t_i, \theta)\right)^2 .$$

We apply the mean value theorem to $\frac{\partial S_n}{\partial \theta}$ in order to obtain a representation of the difference between $\widehat{\theta}_n$ and θ^*. Hence, we receive the existence of the random variable $\widetilde{\theta}_n$ between $\widehat{\theta}_n$ and θ^* such that

$$\frac{\partial S_n(\widehat{\theta}_n)}{\partial \theta} - \frac{\partial S_n(\theta^*)}{\partial \theta} = \frac{\partial^2 S_n(\widetilde{\theta}_n)}{\partial \theta \partial \theta^T} \left(\widehat{\theta}_n - \theta^*\right) . \tag{5.22}$$

Because $\widehat{\theta}_n$ is a WLS estimator

$$\frac{\partial S_n(\widehat{\theta}_n)}{\partial \theta} = 0 \in \mathbb{R}^p . \tag{5.23}$$

With $\frac{\partial^2 S_n(\tilde{\theta}_n)}{\partial \theta \partial \theta^T}$ invertible (as we will see in the following) we obtain

$$\left(\widehat{\theta}_n - \theta^*\right) = - \left(\frac{\partial^2 S_n(\widetilde{\theta}_n)}{\partial \theta \partial \theta^T} \right)^{-1} \cdot \frac{\partial S_n(\theta^*)}{\partial \theta} . \tag{5.24}$$

Multiplying (5.24) by $\sqrt{n}$ gives the representation for asymptotically normal distribution

$$\sqrt{n} \left(\widehat{\theta}_n - \theta^*\right) = - \left(\frac{1}{n} \frac{\partial^2 S_n(\widetilde{\theta}_n)}{\partial \theta \partial \theta^T} \right)^{-1} \cdot \frac{1}{\sqrt{n}} \frac{\partial S_n(\theta^*)}{\partial \theta} . \tag{5.25}$$

We split the investigation of the behavior of the right hand side in (5.25) into two parts.

Part 1: We show for the first term in (5.25)

$$\operatorname*{plim}_{n\to\infty} \left(\frac{1}{n} \frac{\partial^2 S_n(\widetilde{\theta}_n)}{\partial \theta \partial \theta^T} \right)^{-1} = \frac{1}{2} \Omega(\theta^*)^{-1} \in \mathbb{R}^{p,p} . \tag{5.26}$$

Because $\widehat{\theta}_n$ is a consistent estimator we have $\widehat{\theta}_n \in U_{\theta^*}$ for a large n and due to the convexity of U_{θ^*}, $\widetilde{\theta}_n \in U_{\theta^*}$ and therefore, $\widetilde{\theta}_n$ is a consistent estimator,

$$\operatorname*{plim}_{n\to\infty} \widetilde{\theta}_n = \theta^* .$$

We differentiate (5.8)

$$\frac{1}{n} S_n(\theta) = \frac{1}{n} \sum_{i=1}^{n} w_i (\varepsilon_i + x(t_i, \theta^*) - x(t_i, \theta))^2$$

twice with respect to θ and obtain

$$\begin{aligned}\frac{1}{n}\frac{\partial^2 S_n(\theta)}{\partial\theta\partial\theta^T} &= \frac{2}{n}\sum_{i=1}^{n} w_i \frac{\partial x(t_i,\theta)}{\partial\theta}\frac{\partial x(t_i,\theta)}{\partial\theta^T} - \frac{2}{n}\sum_{i=1}^{n} w_i x(t_i,\theta^*)\frac{\partial^2 x(t_i,\theta)}{\partial\theta\partial\theta^T} \\ &\quad + \frac{2}{n}\sum_{i=1}^{n} w_i x(t_i,\theta)\frac{\partial^2 x(t_i,\theta)}{\partial\theta\partial\theta^T} - \frac{2}{n}\sum_{i=1}^{n} w_i \varepsilon_i \frac{\partial^2 x(t_i,\theta)}{\partial\theta\partial\theta^T} \\ &=: 2\overline{X}_n(\theta) - 2\overline{Y}_n(\theta^*,\theta) + 2\overline{Y}_n(\theta,\theta) - 2\overline{Z}_n(\theta)\,. \end{aligned} \tag{5.27}$$

We investigate (5.27) term by term. The $\mathbb{R}^{p,p}$ is equipped with the matrix norm

$$||A||_\infty = \max\{\sum_{\beta=1}^{p} |A_{\alpha,\beta}| \,|\, \alpha = 1,...,p\}\,.$$

Note that for a sequence of matrices $A_1, A_2, ...$

$$||A_n - A||_\infty \to 0 \quad\Longleftrightarrow\quad |(A_n)_{\alpha,\beta} - A_{\alpha,\beta}| \to 0\,, \quad \alpha,\beta \in \{1,...,p\}\,. \tag{5.28}$$

This opens the route to apply Theorem 5.5.5 component-by-component.

We have that $\overline{X}_n(\theta), \overline{Y}_n(\theta)$ and $\overline{Z}_n(\theta)$ are continuous in $\theta \in U_{\theta^*}$ for $i = 1,...,n$. Because of the assumed uniformly convergence also the limits are continuous.

Consider the <u>first term</u> $\overline{X}_n : \mathbb{R}^p \to \mathbb{R}^{p,p}$ in (5.27). Then we obtain for each component with (5.17) for $\widetilde{\theta}_n$

$$\operatorname*{plim}_{n\to\infty} \overline{X}_n^{\alpha,\beta}(\widetilde{\theta}_n) = \Omega^{\alpha,\beta}\left(\operatorname*{plim}_{n\to\infty}\widetilde{\theta}_n\right) = \Omega^{\alpha,\beta}(\theta^*)\,, \quad \alpha,\beta \in \{1,...,p\}\,.$$

Therefore, with (5.28)

$$\operatorname*{plim}_{n\to\infty} 2\overline{X}_n(\widetilde{\theta}_n) = \operatorname*{plim}_{n\to\infty}\frac{2}{n}\sum_{i=1}^{n} w_i \frac{\partial x(t_i,\widetilde{\theta}_n)}{\partial\theta}\frac{\partial x(t_i,\widetilde{\theta}_n)}{\partial\theta^T} = \operatorname*{plim}_{n\to\infty}\frac{2}{n}P^T(\widetilde{\theta}_n)WP(\widetilde{\theta}_n) = 2\Omega(\theta^*)\,.$$

For the <u>second and third term</u> in (5.27) we obtain by analogous argumentation with (5.18)

$$\begin{aligned}&\operatorname*{plim}_{n\to\infty}\overline{Y}_n(\theta^*,\widetilde{\theta}_n) - \operatorname*{plim}_{n\to\infty}\overline{Y}_n(\widetilde{\theta}_n,\widetilde{\theta}_n) \\ &= Y\left(\theta^*, \operatorname*{plim}_{n\to\infty}\widetilde{\theta}_n\right) - Y\left(\operatorname*{plim}_{n\to\infty}\widetilde{\theta}_n, \operatorname*{plim}_{n\to\infty}\widetilde{\theta}_n\right) = Y(\theta^*,\theta^*) - Y(\theta^*,\theta^*) = 0\,.\end{aligned}$$

For the <u>fourth term</u> we consider a compact convex neighborhood $B \subset U_{\theta^*}$ with $\theta^* \in B$. The components for $\alpha,\beta \in \{1,...,p\}$ read with $\theta \in B$

$$\left|\overline{Z}_n^{\alpha,\beta}(\theta)\right| = \left|\tfrac{1}{n}\sum_{i=1}^{n} w_i\varepsilon_i\left(\tfrac{\partial^2 x(t_i,\theta)}{\partial\theta\partial\theta^T}\right)_{\alpha,\beta}\right| \le \frac{1}{n}\left|\sum_{i=1}^{n} w_i\varepsilon_i M_{\alpha,\beta}\right|\,.$$

We have

$$\mathbb{E}\left[\left|\overline{Z}_n^{\alpha,\beta}(\theta)\right|\right] = 0$$
$$\mathrm{Var}\left[\left|\overline{Z}_n^{\alpha,\beta}(\theta)\right|\right] \leq \frac{1}{n^2} M_{\alpha,\beta}^2 \sum_{i=1}^{n} w_i^2 \sigma_i^2 \leq \frac{1}{n^2} M_{\alpha,\beta}^2 \omega \sum_{i=1}^{n} w_i \sigma_i^2 < \infty$$

because of (5.5). Applying Chebyshev's inequality gives

$$\mathbb{P}\left[\left|\overline{Z}_n^{\alpha,\beta}(\theta)\right| \geq \alpha\right] \leq \frac{1}{n}\left(M_{\alpha,\beta}^2 \omega \frac{1}{n}\sum_{i=1}^{n} w_i \sigma_i^2\right)$$

and with (5.5) for $n \to \infty$

$$\operatorname*{plim}_{n\to\infty} \overline{Z}_n^{\alpha,\beta}(\theta) = 0 \quad \text{for } \theta \in B\,.$$

Hence, for large n we have $\operatorname*{plim}_{n\to\infty} \overline{Z}_n^{\alpha,\beta}(\widetilde{\theta}_n) = 0$ and finally, with (5.28)

$$\operatorname*{plim}_{n\to\infty} \overline{Z}_n(\widetilde{\theta}_n) = 0\,.$$

Summarizing, we showed

$$\operatorname*{plim}_{n\to\infty} \frac{1}{n} \frac{\partial^2 S_n(\widetilde{\theta}_n)}{\partial\theta\partial\theta^T} = 2\Omega(\theta^*)$$

where $\Omega(\theta^*)$ is invertible because of assumption (5.17).

Using the matrix inverse rule for probability limits (see e.g. [Gre12]) gives

$$\operatorname*{plim}_{n\to\infty} \left(\frac{1}{n} \frac{\partial^2 S_n(\widetilde{\theta}_n)}{\partial\theta\partial\theta^T}\right)^{-1} = \frac{1}{2}\Omega(\theta^*)^{-1} \in \mathbb{R}^{p,p}\,.$$

Moreover, the inverse in (5.24) exists.

Part 2: The second term in (5.25) reads

$$\frac{1}{\sqrt{n}} \frac{\partial S_n(\theta^*)}{\partial\theta} = -\frac{2}{\sqrt{n}} \sum_{i=1}^{n} w_i \left(y_i - x(t_i,\theta^*)\right) \frac{\partial x(t_i,\theta^*)}{\partial\theta} = -\frac{1}{\sqrt{n}} \sum_{i=1}^{n} 2 w_i \varepsilon_i \frac{\partial x(t_i,\theta^*)}{\partial\theta}\,.$$

We set

$$\Gamma_i := -2 w_i \varepsilon_i \frac{\partial x(t_i,\theta^*)}{\partial\theta} \in \mathbb{R}^p$$

and apply the central limit theorem. Note that (5.19) ensures the applicability of Theorem 5.5.4. The expected value of Γ_i is $\mathbb{E}[\Gamma_i] = 0 \in \mathbb{R}^p$ for $i = 1, ..., n$. The covariance matrix reads

$$\mathrm{Cov}[\Gamma_i] = Q_i = 4\sigma_i^2 w_i^2 \frac{\partial x(t_i,\theta^*)}{\partial\theta} \frac{\partial x(t_i,\theta^*)}{\partial\theta^T}\,.$$

We obtain

$$\overline{Q}_n = \frac{1}{n}\sum_{i=1}^{n} Q_i = \frac{1}{n}\sum_{i=1}^{n} 4\sigma_i^2 w_i^2 \frac{\partial x(t_i,\theta^*)}{\partial \theta}\frac{\partial x(t_i,\theta^*)}{\partial \theta^T} = \frac{4}{n}P^T(\theta^*)WGWP(\theta^*)\,.$$

Because of (5.19) we have by Theorem 5.5.4

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\Gamma_i - \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[\Gamma_i]\right) = \frac{1}{\sqrt{n}}\frac{\partial S_n(\theta^*)}{\partial \theta} \xrightarrow{d} \mathcal{N}\left(0, 4\Sigma(\theta^*)\right) \quad \text{for } n\to\infty\,. \tag{5.29}$$

Summarizing we are able to consider the overall behavior of (5.25). We apply Slutsky's theorem with the results (5.26) and (5.29). Note that the limit in (5.26) is non-stochastic and symmetric. Using $A\mathrm{Cov}(X)A^T = \mathrm{Cov}(AX)$ gives

$$\sqrt{n}\left(\widehat{\theta}_n - \theta^*\right) \xrightarrow{d} \frac{1}{2}\Omega(\theta^*)^{-1}\mathcal{N}\left(0, 4\Sigma(\theta^*)\right) = \mathcal{N}\left(0, \Omega(\theta^*)^{-1}\Sigma(\theta^*)\Omega(\theta^*)^{-1}\right)\,.$$

for $n\to\infty$.
□

The assumptions in Theorem 5.5.4 are typical statistical conditions, compare with the different formulations of the central limit theorem in literature, see [Rao87]. Also assumptions (5.17) and (5.18) are typical statistical conditions, compare with the ordinary least squares approach in literature, see [SW89].

5.6 Confidence interval and coefficient of variation

A confidence interval could be understood as follows, compare e.g. [Bar74]. Consider an experiment repeated one hundred times. Each experiment would yield an estimate $\widehat{\theta}_n$. Calculate for example the 95% confidence interval for each estimate. Then the true value θ^* should be contained in about 95 of these 100 intervals. A large confidence interval indicates that the sample size is maybe to small, the data are not sufficent for estimating a special parameter or it exists a strong variability in reality.

The true covariance matrix is never known in practice. To account for this uncertainty, the Student's t-distribution is used to calculate the confidence interval, see [Fox97].

To introduce the Student's t-distribution, we firstly define the Chi-Quadrat distribution.

Definition 5.6.1
Let $X_1, ..., X_n$ be independent random variables with $X_i \sim \mathcal{N}(0,1)$ for $i = 1, ..., n$. Then

$$\sum_{i=1}^{n} X_i^2 \sim \chi_n^2$$

is Chi-Quadrat distributed with n degrees of freadom.

Definition 5.6.2
Let $X \sim \mathcal{N}(0,1)$, $Y \sim \chi^2_n$ and X, Y independent. The Student's t-distribution is defined as the distribution of the random variable

$$t_n = \frac{X}{\sqrt{\frac{Y}{n}}}\,. \tag{5.30}$$

The proved asymptotic normality of the WLS estimator $\widehat{\theta}_n$ gives approximately

$$\widehat{\theta}_n \overset{a}{\sim} \mathcal{N}\left(\theta^*, \frac{1}{n}\Omega(\theta^*)^{-1}\Sigma(\theta^*)\Omega(\theta^*)^{-1}\right) =: \mathcal{N}\left(\theta^*, \mathrm{Cov}(\theta^*)\right)\,.$$

Now we denote by θ_j the jth entry of the vector $\theta \in \mathbb{R}^p$ for $j = 1, ..., p$ and suppress the sample size n. Let $e_j \in \mathbb{R}^p$ be the vector with a one at the jth place and zero elsewhere. Hence

$$\widehat{\theta}_j = e_j^T\widehat{\theta} \overset{a}{\sim} \mathcal{N}\left(e_j^T\theta^*, e_j^T\mathrm{Cov}_{jj}(\theta^*)e_j\right) = \mathcal{N}(\theta_j^*, \mathrm{Cov}_{jj}(\theta^*)) \quad \Longleftrightarrow \quad \frac{(\widehat{\theta}_j - \theta_j^*)}{\sqrt{\mathrm{Cov}_{jj}(\theta^*)}} \overset{a}{\sim} \mathcal{N}(0,1)\,. \tag{5.31}$$

An appropriate estimator for $\mathrm{Cov}(\theta^*)$ reads with (5.20), (5.21) and $\widehat{G} = \mathrm{diag}(s_1^2, ..., s_n^2)$

$$\begin{aligned}\widehat{\mathrm{Cov}} = \mathrm{Cov}(\widehat{\theta}) &= \frac{1}{n}\left(\frac{1}{n}P^T(\widehat{\theta})WP(\widehat{\theta})\right)^{-1}\frac{1}{n}P^T(\widehat{\theta})W\widehat{G}WP(\widehat{\theta})\left(\frac{1}{n}P^T(\widehat{\theta})WP(\widehat{\theta})\right)^{-1} \\ &= \left(P^T(\widehat{\theta})WP(\widehat{\theta})\right)^{-1}P^T(\widehat{\theta})W\widehat{G}WP(\widehat{\theta})\left(P^T(\widehat{\theta})WP(\widehat{\theta})\right)^{-1}\end{aligned} \tag{5.32}$$

where s_i^2 denotes the sample variance. Compare representation (5.32) e.g. with [DSW09] and see Section 5.7 for the design of the variance estimator. Now we further assume that

$$\varepsilon_i \sim \mathcal{N}(0, \sigma_i^2) \quad \text{ for } i = 1, ..., n\,.$$

With appropriate weights (see e.g. situation 3 in Section 5.3) we obtain approximately $\varepsilon_i \overset{a}{\sim} \mathcal{N}(0, \sigma^2)$ and following [SL77] and [SW89] gives approximately

$$\frac{(n-p)\widehat{\mathrm{Cov}}_{jj}}{\mathrm{Cov}_{jj}(\theta^*)} \overset{a}{\sim} \chi^2_{n-p} \quad \text{for } j = 1, ..., p \tag{5.33}$$

and $\widehat{Cov}_{jj}$ independent from $\widehat{\theta}_j$ for $j = 1, ..., p$. Hence, definition (5.30) with (5.31) and (5.33) gives approximately

$$\frac{(\widehat{\theta}_j - \theta_j^*)}{\sqrt{\widehat{\mathrm{Cov}}_{jj}}} \overset{a}{\sim} t_{n-p}$$

compare [SW89]. In terms of probability we obtain with $\alpha \in [0,1]$ the confidence interval for the confidence level $1-\alpha$

$$\mathbb{P}\left[\widehat{\theta}_j - t_{n-p}^{1-\frac{\alpha}{2}}\sqrt{\widehat{\mathrm{Cov}}_{jj}} \leq \theta_j^* \leq \widehat{\theta}_j + t_{n-p}^{1-\frac{\alpha}{2}}\sqrt{\widehat{\mathrm{Cov}}_{jj}}\right] = 1-\alpha\,.$$

Hence, the approximated $(1-\alpha)$-confidence interval reads

$$\left[\widehat{\theta}_j - t_{n-p}^{1-\frac{\alpha}{2}}\sqrt{\widehat{\mathrm{Cov}}_{jj}}\,,\ \widehat{\theta}_j + t_{n-p}^{1-\frac{\alpha}{2}}\sqrt{\widehat{\mathrm{Cov}}_{jj}}\right]. \tag{5.34}$$

Note that unrealistic negative values could appear for the left interval boundary. In the following we mark such values by zero.

The estimator for the coefficent of variation of the WLS estimator $\widehat{\theta}_j$ in percent reads

$$\mathrm{CV}\%(\widehat{\theta}_j) = 100 \cdot \frac{\widehat{\mathrm{Cov}}_{jj}}{\widehat{\theta}_j}. \tag{5.35}$$

see [DSW09]. If CV% is larger than 100, we mark it by "$>$100" in the following.

5.7 Application to simultaneous PKPD fits

Data from PKPD experiments consists of a control group (placebo) and different dosing groups (administered drug). Let m be the number of all groups, then the sum of squares reads

$$S(\theta) = \sum_{k=1}^{m}\sum_{i=1}^{n_k} w_{ik}(y_{ik} - x(t_{ik}, dose_k, \theta))^2$$

where n_k, $k = 1, ..., m$, is the number of the measurements of each group with $n_1 + \cdots + n_m = n$. Note that we emphasize the dependency of *dose* in the PKPD model. A common assumption (see e.g. [DSW09]) is that every dosing group has a separate variance σ_k^2. An appropriate estimator reads

$$s_k^2 = \frac{1}{n_k - \frac{p}{m}}\sum_{i=1}^{n_k} w_{ik}(y_{ik} - x(t_{ik}, dose_k, \widehat{\theta}))^2$$

where p is the number of model parameters, see [DSW09]. Hence, the variance estimator matrix reads

$$\widehat{G} = \mathrm{diag}\left(\underbrace{s_1^2, \ldots, s_1^2}_{n_1}, \ldots, \underbrace{s_m^2, \ldots, s_m^2}_{n_m}\right). \tag{5.36}$$

Statistical a posteriori analysis for a point estimator:

Let $\widehat{\theta}$ be a weighted least squares estimator obtained by a fitting process for a sample of size n. The a posteriori statistical analysis is obtained as follows:

The covariance matrix is calculated by (5.32) with the variance estimates (5.36). The coefficient of variation for a single parameter estimator $\widehat{\theta}_j$, $j = 1, ..., p$, is calculated by (5.35). The (1-α)-confidence interval for $\widehat{\theta}_j$ is obtained with (5.34).

Finally, we use the coefficient of determination $R^2 \in [0,1]$ as a goodness of fit criteria, see [DSW09]. A value close to 1 indicates that the model describes the data well. For each dosing group $k = 1, ..., m$ we calculate

$$R_k^2 = \frac{\sum_{i=1}^{n_k} \left(x(t_i, dose_k, \widehat{\theta}) - \overline{x} \right) (y_{ik} - \overline{y}_k)}{\sum_{i=1}^{n_k} \left(x(t_i, dose_k, \widehat{\theta}) - \overline{x}_k \right)^2 \sum_{i=1}^{n} (y_{ik} - \overline{x}_k)^2}$$

where $\overline{x}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} x\left(t_i, dose_k, \widehat{\theta}\right)$ and $\overline{y}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} y_{ik}$.

5.8 Discussion and outlook

In this chapter we investigated the asymptotical behavior of the weighted least squares estimator based on reasonable statistical assumptions and derived statistical characteristics, like confidence intervals and coefficient of variation for parameter estimates.

We remark that in statistics, the weights sometimes are also considered as a function of time, see [Rao97]. However, in our data situation such extensions are of no further help. Typically, in PKPD weights are used to deal with unequal variances observed in data.

We want to emphasize that on the way towards calculation of a confidence interval many assumptions and approximations are performed. Further, in our situation the amount of data is limited. Hence, we suggest not to overate the presented statistical characteristics obtained from a fitting process.

Note that in this work, we are not interested in a statistical population analysis (see [DG95] or [Bon06]) as applied in clinical phase II/III, due to several reasons. First, our amount of individuals (ranging from 20-80) is not comparable to the situation in clinics Second, the existence of typical individual covariates (e.g. age, weight, smoker,...) is not given in preclinics.

Chapter 6

Modeling of Tumor Growth and Anticancer Effects of Mono- and Combination Therapy

It is reported in a study from the Council of the European Union (see [dEU08]) in 2008, that every third European develops cancer once in his lifetime. Therefore, a huge field in drug development deals with the understanding of the mechanisms of cancer development and the design of appropriate anticancer drugs.

It is generally stated in literature that the work from Anna Kane Laird [Lai64] "*Dynamics of tumor growth*" published in 1964 initiated the mathematical modeling of tumor growth. Laird applied the Gompertz equation (here presented in the original formulation)

$$\frac{W}{W_0} = e^{\frac{A}{\alpha}(1-e^{-\alpha t})}$$

to describe unperturbed tumor growth in a test tube. W denotes the tumor size in time, W_0 is the initial tumor size and A, α are growth related parameters. This model realizes a sigmoid growth behavior and therefore, describes the three significant phases of tumor growth. First, a tumor grows exponentially, after a while the tumor growth gets linear due to limits of nutrient supply and finally, the tumor growth completely saturates. Laird applied the Gompertz equation to data from mice, rats and rabbits.

In the book of T. E. Wheldon titled "*Mathematical models and cancer research*" [Whe88] from 1988, an overview and analysis of the Gompertz model could be found. An important statement from Wheldon in [Whe88] is that the saturation property of tumors could often never measured in patients in practice because the host dies in the majority of cases before this saturation phase begins. Also in preclinics, the experiments have to be canceled if a specific tumor size is reached due to ethical reasons. Hence, in this work we present a tumor growth model without a saturation property and focus on the first two tumor growth phases, namely exponential growth followed by linear growth.

In our experiments, drugs inhibiting the histone deacetylase (class of enzymes) activity, see [KOM02], were tested. This inhibition leads to the blocking of the cell cycle of

proliferating cells and therefore, drives the cancer cells in the apoptose (process of a programmed cell death). At the publication date of this work, the compounds are in phase I/II and have a promising outlook. But due to confidentiality reasons, no details about the compounds are presented in this work. In the following, we encrypt all anticancer compounds by capitals.

The presented experiments were performed in xenograft mice (see e.g. [BD11]). Such mice develop human solid tumors based on implantation of human cancer cells. The tumor grows on the back of the mice and is measured as volume and recalculated to weight based on tissue consistency assumptions.

The aim of our project was to develop a PKPD model for tumor growth, describing mono-therapy (single drug administration) as well as combination therapy (several drugs are administered). The combination approach is widely applied in cancer treatment. The main motivation to combine anticancer drugs in clinics is to obtain a synergistic response. Based on this synergistic effect the amount of dose could be reduced in order to lower toxicity in patients. Hence, also in early drug development one main objective is to identify drug combinations which have an enhanced pharmacological effect and to rank them according to their interaction intensity.

This chapter is structured in the following way. In Section 6.1 a brief introduction to the experimental setup is presented. In Section 6.2 we develop a tumor growth model for unperturbed growth (no drug administration). Then we extend this tumor growth model by a drug-effect term for mono-therapy in Section 6.3. We introduce assumptions about drug effects and include the pharmacokinetics in the model. The resulting model consists of $n+1$ ordinary differential equations based on a n-dimensional transit compartment system (compare Section 4.2) describing different damaging stages of the attacked tumor cells. Section 6.4 deals with the extension of the PKPD model to combination therapy. In Section 6.5 an important secondary parameter describing the necessary mean concentration for tumor eradication is calculated. This parameter opens the route to compare experiments in animals and humans and hence, could be used to perform animal to human predictions. For example, Rocchetti presented in 2007, see [RSP$^+$07], based on a structural similar tumor growth model the scaling between animals and humans for known anticancer drugs.

In the next section we refine the results from Section 6.1-6.5 and reformulate our PKPD model for mono-therapy by delay differential equations of lifespan type based on the theoretical results from Section 4.4. This reformulation reduces the number of differential equations to exactly two, where one describes the proliferating cells and the other the attacked tumor cells by a drug.

6.1 Experimental setup

For pharmacokinetic measurements, single oral doses of the drugs were administered to five mice. The PK was either modeled by a one or two compartment model with p.o. administration.

The different dosing groups (including the placebo group) consists at any time of 8-10 mice. The tumor grows at the back of the mice and is visually accessible. The dimension of the tumors, more precisely, the length and width, were measured by an electronic caliber. Using this information the volume was calculated and recalculated as weight based on tissue consistency assumptions.

6.2 Unperturbed tumor growth

It was observed in more than 70 data sets in our experiments as well as in several data from literature ([SMC+04], [MSP+06], [SPRVC07]) that unperturbed tumor growth in xenograft mice consists of an initial exponential growth phase followed by a linear growth phase. The aim of this section is to model this behavior with a realistic right hand side of the differential equation

$$w'(t) = f(w(t)), \quad w(0) = w_0$$

where w_0 is the inoculated tumor weight (the amount of implanted human tumor cells into the xenograft mouse). The tumor weight is denoted by $w(t)$.

In 2004, Simeoni et al [SMC+04] presented a model consisting of an exponential and a linear growth phase in order to describe the tumor growth in xenograft mice in time. Their tumor growth function $g_s : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ reads

$$g_s(w) = \begin{cases} \lambda_0 w, & w \leq w_{th} \\ \lambda_1, & w > w_{th} \end{cases}, \qquad w_{th} = \frac{\lambda_1}{\lambda_0}. \tag{6.1}$$

The parameter $\lambda_0 > 0$ describes the exponential growth rate and $\lambda_1 > 0$ the linear growth rate. This model has the property that if the weight w reaches a threshold w_{th}, the exponential growth switches immediately to linear growth. This produces a fast transition between the exponential and linear phase in $w(t)$. But this fast transition was found unrealistic in our working group and therefore, we present a new approach with a more pronounced transition. Further, one directly observes that $g_s(w)$ suffers at a lack of differentiability for $w = w_{th}$. To overcome this issue, Simeoni et al suggested to apply the approximation

$$g_a(w) = \frac{\lambda_0 w}{\left[1 + \left(\frac{\lambda_0}{\lambda_1} w\right)^{20}\right]^{1/20}} \tag{6.2}$$

for $g_s(w)$ in practice. Unfortunately, this approximation consists of several costly exponentiations and is therefore, questionable also numerically.

Our aim is to model the unperturbed tumor growth by a Michaelis-Menten type of function, see e.g. [Boh06]. Such an approach is biologically more evident because it allows a more pronounced transition between the exponential and linear phase. Note that the Michaelis-Menten approach is a basic equation in biology, see e.g. [Mur89]. The Michaelis-Menten function reads

$$g(w) = \frac{aw}{b+w} \tag{6.3}$$

with $a, b > 0$. First, (6.3) is everywhere differentiable and second, consist of no questionable raise to higher power.

To achieve physiological meaningful parameters with the same meaning as in Simeoni's model we adjust actual values of a and b. We set

$$\lim_{w\to\infty} g(w) = \lambda_1$$

which ensures that the maximal growth rate of g_s coincides with g, hence, $a = \lambda_1$. The parameter b in (6.3) describes the point w when $g(w)$ gets equal to $\frac{a}{2}$. Therefore, we identify the half value of the threshold w_{th} by the parameter b and set

$$g\left(\frac{\lambda_1}{2\lambda_0}\right) = \frac{a}{2} \tag{6.4}$$

which ensures that the half maximal value of g_s and g coincide. We end up with the expression

$$g(w) = \frac{\lambda_1 w}{\frac{\lambda_1}{2\lambda_0} + w} = \frac{2\lambda_0\lambda_1 w}{\lambda_1 + 2\lambda_0 w}\,. \tag{6.5}$$

Differentiation of (6.5) at $w = 0$ gives $g'(0) = 2\lambda_0$.

With (6.5) we obtain the model

$$w'(t) = \frac{2\lambda_0\lambda_1 w(t)}{\lambda_1 + 2\lambda_0 w(t)}\,, \qquad w(0) = w_0 \tag{6.6}$$

for unperturbed tumor growth $w(t)$ with the three parameter

$$\theta = (\lambda_0, \lambda_1, w_0)\,.$$

Hence, λ_0 describes the rate of exponential growth and λ_1 the rate of linear growth.

In Figure 6.1 the tumor growth function (6.1) and the tumor growth function (6.5) are plotted.

In the following we present measurements from four different human tumor cell lines, namely RKO (cancer of the colon), PC3 (prostate cancer), MDA (breast cancer) and A459 (lung cancer). These data were fitted by the model (6.6), see Figure 6.2. The parameter estimates are listed in Table 6.1.

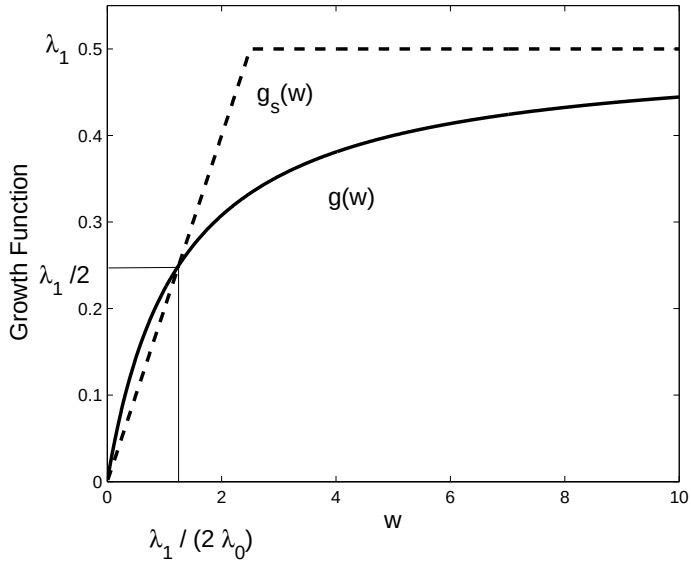


Figure 6.1: *The dashed line is the tumor growth function (6.1) and the solid line is the tumor growth function (6.5). The parameters are* $\lambda_0 = 0.2$ *and* $\lambda_1 = 0.5$.

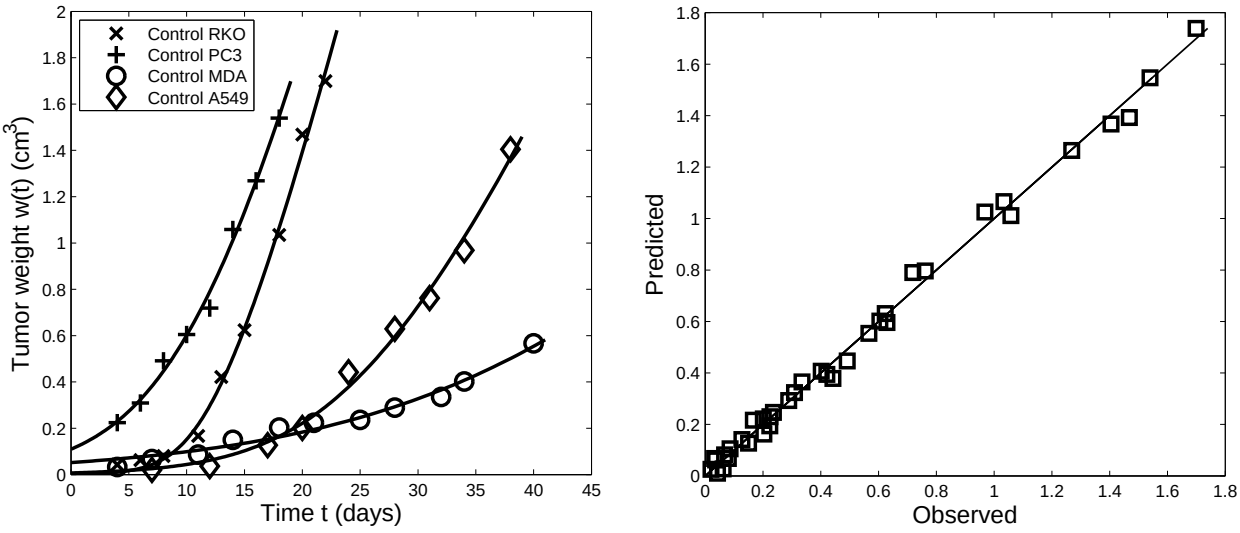


Figure 6.2: *Different tumor cell lines in xenograft mice fitted with (6.6). In the left panel the tumor growth is plotted and the right panel shows an observed vs. predicted plot.*

Parameter	Explanation	RKO	PC3	MDA	A549
λ_0	Exponential rate	0.250	0.104	0.033	0.093
(CV%)		(32.4)	(27.8)	(29.8)	(27.1)
CI		[0.052, 0.448]	[0.030, 0.178]	[0.011, 0.057]	[0.031, 0.155]
λ_1	Linear rate	0.223	0.275	0.104	0.139
(CV%)		(18.3)	(43.6)	(>100)	(30.6)
CI		[0.123, 0.333]	[0, 0.582]	[0, 0.472]	[0.035, 0.244]
w_0	Inoculated tumor weight	0.001	0.110	0.052	0.007
(CV%)		(>100)	(28.7)	(26.2)	(93.5)
CI		[0, 0.007]	[0.029, 0.191]	[0.021, 0.084]	[0, 0.024]
	Sum of Squares	0.014	0.010	0.006	0.012
	R^2	0.99	0.99	0.98	0.99

Table 6.1: *Model parameter estimates, coefficient of variation and 95%-confidence interval of the fit of unperturbed tumor growth for different human tumor cell lines in xenograft mice.*

6.3 Perturbed tumor growth for mono-therapy

The next step towards a PKPD tumor growth model is to include the pharmacokinetics of a drug, or more precisely, the perturbation of the tumor growth by an anticancer agent. It is generally observed that the pharmacological effect is delayed due to the drug concentration. For example, Lobo and Baltasar [LB02] stated that *"chemotherapeutic effects often appear days or weeks following drug exposure"* in humans. Therefore, our first assumption reads:

(A1) The effect (death of proliferating cells) due to drug concentration is delayed.

Lobo and Baltasar presented in 2002 a PKPD model for delayed chemotherapeutic effects with respect to drug concentration. They applied a transit compartment model (see Chapter 4) to delay the drug concentration. In 2004, Simeoni and co-workers presented a tumor growth model for data measured in xenograft mice. They also applied a transit compartment model, but instead of delaying the effect of the drug, it is assumed that proliferating cells attacked by the drug will pass through different damaging stages until the cells finally and irrevocably die. Hence, the attacked tumor cells have a lifespan. We formulate this as second assumption:

(A2) Tumor cells affected by drug action stop to proliferate and will irrevocably die after a certain lifespan.

The effect of the drug on proliferating cells is described by a linear drug-effect term (see Section 4.6) of the form

$$e(k_{pot}, c(t)) = k_{pot} \cdot c(t)$$

where $c(t)$ is the pharmacokinetics of the drug and $k_{pot} > 0$ the potency parameter of the drug. Because we have only two dosing groups (placebo and one drug administration) in our performed experiments, the linear effect term is an appropriate choice. We remark that the main focus in our experiments laid on testing compounds and to rank them among each other by their potency.

For the first, we also apply a transit compartment model to describe different stages of dying tumor cells. We denote by $p(t)$ the amount of proliferating tumor cells and by $d_1(t), ..., d_n(t)$ the different stages of dying tumor cells attacked by an anticancer agent. Since, the non-proliferating cells $d_1, ..., d_n$ still add to total tumor mass, the total tumor volume w is the sum of proliferating tumor cells p and non-proliferating tumor cells $d_1, ..., d_n$. In Figure 6.3 we present the schematic overview of the PKPD model for unperturbed and perturbed tumor growth based on transit compartments.

Only proliferating cells that are not affected by drug action contribute to the tumor growth. Therefore, the tumor growth function $g(w)$ of the total tumor consisting of proliferating and non-proliferating cells is slowed down by the factor $\frac{p}{w}$ because p represents that portion of total tumor volume w that is actually proliferating.

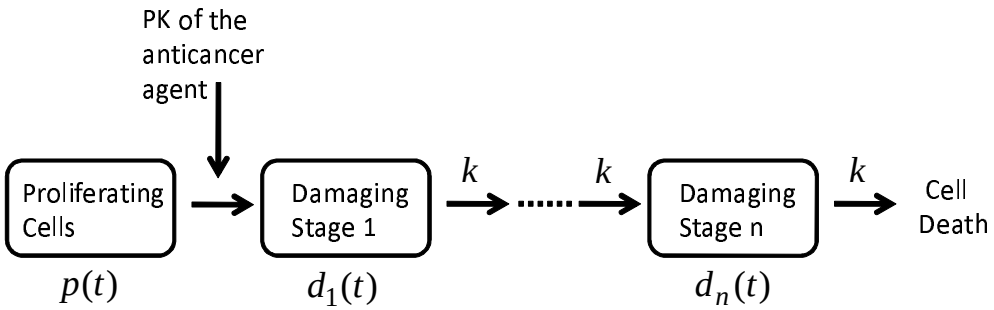


Figure 6.3: *Schematic overview of the PKPD model for unperturbed and perturbed tumor growth based on mono-therapy.*

In addition to the three parameters λ_0, λ_1 and w_0 for unperturbed tumor growth we have the drug potency parameter k_{pot} and the transit rate k between the compartments of the dying tumor cells.

The PKPD model for mono-therapy with transit compartments reads

$$p'(t) = \frac{2\lambda_0\lambda_1 p(t)}{(\lambda_1 + 2\lambda_0 p(t))} \cdot \frac{p(t)}{w(t)} - k_{pot} \cdot c(t) \cdot p(t)\,, \qquad p(0) = w_0 \tag{6.7}$$

$$d_1'(t) = k_{pot} \cdot c(t) \cdot p(t) - k \cdot d_1(t)\,, \qquad d_1(0) = 0 \tag{6.8}$$

$$d_i'(t) = k \cdot d_{i-1}(t) - k \cdot d_i(t)\,, \qquad d_i(0) = 0\,, \quad i = 2, ..., n \tag{6.9}$$

$$w(t) = p(t) + d_1(t) + \cdots + d_n(t)\,, \tag{6.10}$$

with the model parameters

$$\theta = (\lambda_0, \lambda_1, w_0, k_{pot}, k)\,.$$

The total tumor weight is denoted by $w(t)$.

The model provides information regarding the lifespan (called time-to-death in [KWLS09]) T of the attacked tumor cells, which is the mean transit time that it takes for the tumor cells, affected by the action of a drug, to go through the cascade of damaging events to cell death. In our case the average lifespan is computed by

$$T = \frac{n}{k} \tag{6.11}$$

after a successful fitting process as a secondary parameter.

We present different simultaneous fits of unperturbed and perturbed data with (6.7)-(6.10) and $n = 3$ in Figure 6.2 and the parameter estimates in Table 6.2. See Appendix B for pharmacokinetic parameter.

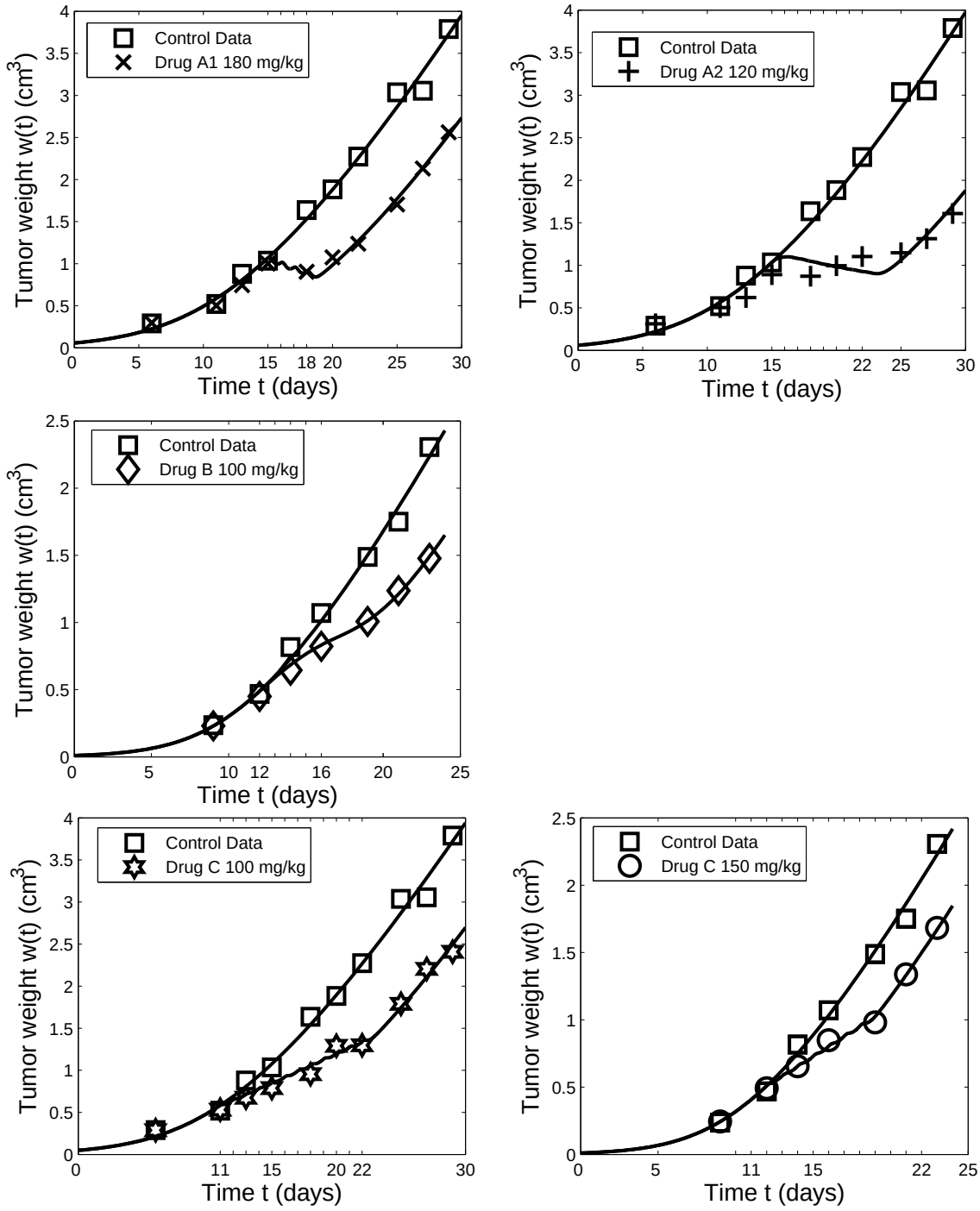


Figure 6.4: *In every plot the unperturbed and perturbed tumor growth data was simultaneously fitted. In the left upper panel the drug A1 was administered at day 15,16,17 and 18 with 180 mg/kg. In the right upper panel the drug A2 was administered 8 times every day starting from day 15 with 120 mg/kg. In the left middle panel the drug B was administered at day 12,13,14,15 and 16 with 100 mg/kg. In the left lower panel the drug C was administered 12 times every day starting from day 11 with 100 mg/kg. In the right lower panel the drug C was administered 12 times every day starting from day 11 with 150 mg/kg.*

Parameter	Explanation	A1 180 mg/kg	A2 120 mg/kg	B 100 mg/kg	C 100 mg/kg	C 150 mg/kg
λ_0		0.127	0.121	0.194	0.136	0.196
(CV%)	Exponential rate	(26.4)	(30.5)	(35.7)	(28.5)	(36.6)
CI		[0.056, 0.198]	[0.042, 0.199]	[0.038, 0.351]	[0.054, 0.218]	[0.034, 0.359]
λ_1		0.293	0.313	0.246	0.280	0.238
(CV%)	Linear rate	(18.3)	(22.8)	(22.1)	(17.9)	(21.8)
CI		[0.179, 0.408]	[0.161, 0.465]	[0.123, 0.370]	[0.174, 0.387]	[0.120, 0.355]
w_0		0.057	0.059	0.010	0.049	0.010
(CV%)	Inoculated tumor weight	(57.2)	(67.5)	(>100)	(65.6)	(>100)
CI		[0, 0.126]	[0, 0.143]	[0, 0.038]	[0, 0.118]	[0, 0.040]
k_{pot}		0.180	0.007	0.007	0.015	0.012
(CV%)	Drug potency	(15.1)	(12.1)	(12.7)	(45.5)	(24.5)
CI		[0.122, 0.234]	[0.005, 0.009]	[0.006, 0.010]	[0.001, 0.030]	[0.005, 0.018]
k		7.22	2.97	0.666	13.7	10.6
(CV%)	Transit rate	(>100)	(60.7)	(31.4)	(>100)	(>100)
CI		[0, 46.7]	[0, 6.805]	[0.193, 1.134]	[0, 488]	[0, 130]
	Sum of Squares	0.150	0.269	0.032	0.166	0.031
	R^2	0.99 / 0.99	0.99 / 0.91	0.99 / 0.99	0.99 / 0.99	0.99 / 0.99

Table 6.2: *Model parameter estimates, coefficient of variation and 95%-confidence interval of the simultaneous fit for unperturbed and perturbed tumor growth with $n = 3$. Further the secondary parameter T is presented. The sum of squares as well as the goodness of fit criteria R^2 (first unperturbed, second perturbed) are presented.*

6.4 Perturbed tumor growth for combination therapy

In cancer drug development, a major aim is to combine different drugs in order to maximize the effect. For example, if the combination of two drugs has a synergistic effect, then the administered doses could be decreased in order to lower the toxicity in patients, see [BFK+09]. In this section we present an approach to model data from combination therapy and to explicitly quantify the effect of combination therapy.

To adjust the PKPD tumor growth model for combination therapy we follow a suggestion of Chakraborty and Jusko [CJ02] and include a combination effect parameter ψ into the model. This parameter ψ quantifies the interaction of the drugs. Again we assume a linear drug-effect term for each drug and multiply one of the linear drug-effect terms by the effect parameter ψ. The drug-effect term then reads

$$e(\psi) = k_{pot}^A c_A(t) + k_{pot}^B c_B(t)\psi \tag{6.12}$$

where k_{pot}^A, k_{pot}^B are the drug potency parameter of drug A and B. The concentrations of drug A and B are denoted by $c_A(t)$ and $c_B(t)$, respectively.

The combination effect parameter ψ in (6.12) has the following meaning:

$$\psi \begin{cases} < 1 & \text{antagonistic effect} \\ = 1 & \text{additive effect} \\ > 1 & \text{synergistic effect} \end{cases} .$$

The investigation of combination therapy data is divided into two steps:

Step 1: Estimate the potency parameter k_{pot}^A and k_{pot}^B of drug A and B from mono-therapy.

Step 2: The PKPD model for combination therapy based on transit compartments then reads with (6.12)

$$p'(t) = \frac{2\lambda_0\lambda_1 p(t)^2}{(\lambda_1 + 2\lambda_0 p(t))w(t)} - (k_{pot}^A c_A(t) + k_{pot}^B c_B(t)\psi)p(t)\,, \qquad p(0) = w_0 \tag{6.13}$$

$$d_1'(t) = (k_{pot}^A c_A(t) + k_{pot}^B c_B(t)\psi)p(t) - k \cdot d_1(t)\,, \qquad d_1(0) = 0 \tag{6.14}$$

$$d_i'(t) = k \cdot d_{i-1}(t) - k \cdot d_i(t)\,, \qquad d_i(0) = 0\,, \quad i = 2, ..., n \tag{6.15}$$

$$w(t) = p(t) + d_1(t) + \cdots + d_n(t)\,, \tag{6.16}$$

with the five model parameter

$$\theta = (\lambda_0, \lambda_1, w_0, \psi, k)\,.$$

Steps 1 and 2 could also be merged to one simultaneous fit, consisting of mono-therapy data of drug A and B as well as the combination therapy data.

Parameter	Explanation	Comb 1	Comb 2	Comb 3
λ_0 (CV%) CI	Exponential rate	0.151 (31.8) [0.049, 0.253]	0.141 (33.1) [0.042, 0.241]	0.172 (33.7) [0.041, 0.033]
λ_1 (CV%) CI	Linear rate	0.271 (17.8) [0.168, 0.374]	0.282 (19.8) [0.163, 0.400]	0.260 (24.4) [0.117, 0.404]
w_0 (CV%) CI	Inoculated tumor weight	0.036 (82.2) [0, 0.098]	0.041 (82.6) [0, 0.114]	0.015 (>100) [0, 0.051]
ψ (CV%) CI	Combination effect parameter	0.837 (28.1) [0.336, 1.334]	0.461 (53.1) [0, 0.983]	1.783 (19.7) [0.991, 2.576]
k (CV%) CI	Transit rate	1.61 (31.4) [0.533, 2.687]	2.58 (56.5) [0, 5.695]	0.491 (16.2) [0.311, 0.672]
	Sum of Squares R^2	0.197 0.99 / 0.95	0.271 0.99 / 0.85	0.031 0.99 / 0.98

Table 6.3: *Model parameter estimates, coefficient of variation and 95%-confidence interval of the simultaneous fit with unperturbed and perturbed tumor growth for combination therapy with $n = 3$. In Comb 1, the drug A1 180 mg/kg and drug C 100 mg/kg is given in combination. In Comb 2, the drug A2 120 mg/kg and drug C 100 mg/kg is given in combination. In Comb 3, the drug B 100 mg/kg an drug C 150 mg/kg is given in combination. The dosing time points are equal to those from mono-therapy. The sum of squares as well as the goodness of fit criteria R^2 (first: unperturbed growth, second: perturbed growth) are presented.*

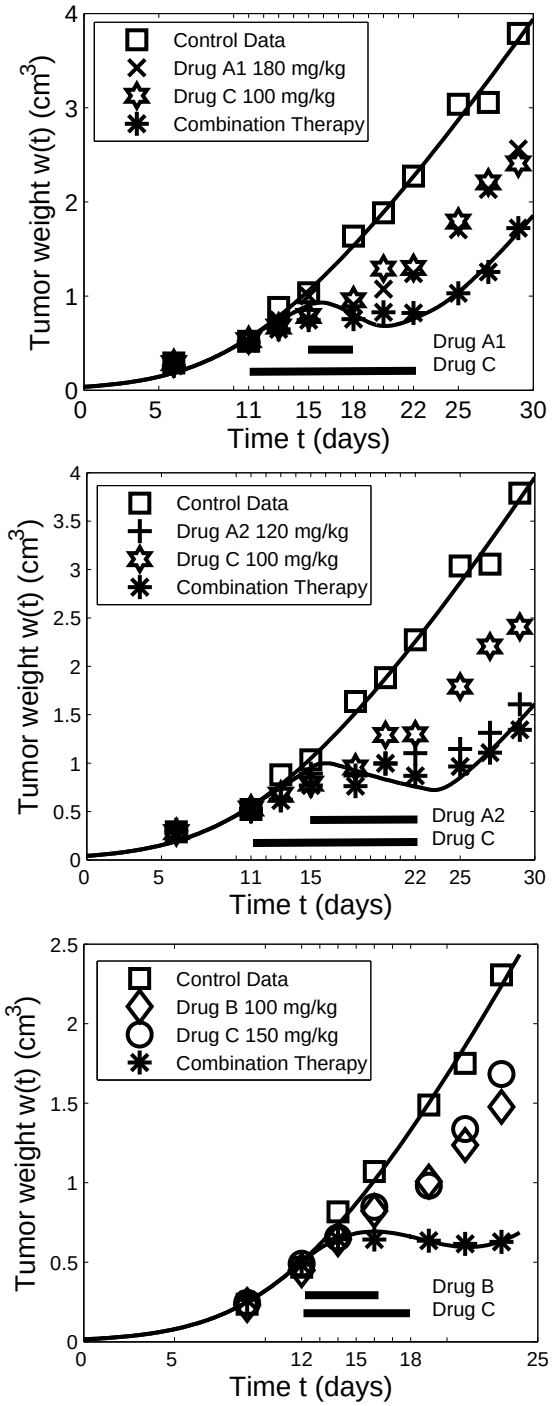


Figure 6.5: *In every plot the unperturbed and perturbed tumor growth data was simultaneously fitted. The bars in each panel denote the dosing interval of the drugs. In the upper panel drug A1 and drug C was administered in combination (Comb 1). In the middle panel drug A2 and drug C was administered in combination (Comb 2). In the lower panel drug B and drug C was administered in combination (Comb 3).*

6.5 The threshold concentration

A structural property of the models (6.7)-(6.10) and (6.13)-(6.16) is that the tumor always starts to re-grow, if the drug concentration vanishes. An interesting theoretical secondary parameter of the models is the so-called threshold concentration $\overline{c}$ required for tumor eradication. Rocchetti and colleagues [RSP+07] have shown a correlation of $\overline{c}$ from experiments in xenograft animals with the human active dose of several drugs for mono-therapy available on the market based on a structural similar PKPD model.

Remark 6.5.1
The constant threshold concentration $\overline{c}$ for tumor eradication reads for mono-therapy

$$k_{pot}\overline{c} = \frac{1}{2T}\left(-1+\sqrt{1+8T\lambda_0}\right)$$

and for combination therapy

$$k_{pot}^A\overline{c}^A + k_{pot}^B\overline{c}^B\psi = \frac{1}{2T}\left(-1+\sqrt{1+8T\lambda_0}\right).$$

Proof: The stationary equations of (6.7)-(6.10) and (6.13)-(6.16) read

$$0 = \frac{2\lambda_0\lambda_1 p^2}{(\lambda_1+2\lambda_0 p)w} - \gamma p \tag{6.17}$$

$$0 = \gamma p - k d_1 \tag{6.18}$$

$$0 = k(d_{i-1}-d_i)\,, \qquad i = 2,...,n \tag{6.19}$$

with

$$\gamma = \begin{cases} k_{pot}\overline{c} & \text{mono-therapy} \\ k_{pot}^A\overline{c}_A + k_{pot}^B\overline{c}_B\psi & \text{combination therapy} \end{cases}.$$

(6.18)-(6.19) directly implies

$$d_1 = \cdots = d_n = \frac{\gamma p}{k}\,. \tag{6.20}$$

Inserting (6.20) into (6.17) leads to

$$\gamma^2\left(n\frac{\lambda_1}{k}+2n\frac{\lambda_0 p}{k}\right)+\gamma(\lambda_1+2\lambda_0 p)-2\lambda_0\lambda_1 = 0\,. \tag{6.21}$$

We set $p=0$ in (6.21) due to the eradication of the tumor and find the unique positive solution

$$\gamma = \frac{-k+k\sqrt{1+8n\frac{\lambda_0}{k}}}{2n}$$

of (6.21) and with $T=\frac{n}{k}$, see (6.11), the stated result.
□

6.6 Tumor growth model for mono-therapy in the lifespan type formulation

In Chapter 4 we presented the relationship between transit compartments (TCM) and lifespan models (LSM), if the number of compartments tends to infinity.

The aim of this section is to reformulate the presented model (6.7)-(6.10) as delay differential equation of lifespan type. For that purpose we consider the tumor cells attacked by an anticancer agent as a population, where every cell has a lifespan T. After that lifespan T the attacked tumor cells have to leave the population.

Consider the general representation of the PKPD tumor growth model

$$p'(t) = g(\eta, p(t), d_1(t) + \cdots + d_n(t)) - e(\sigma, c(t)) \cdot p(t)\,, \qquad p(0) = w_0 \tag{6.22}$$
$$d_1'(t) = e(\sigma, c(t)) \cdot p(t) - k \cdot d_1(t)\,, \qquad d_1(0) = 0 \tag{6.23}$$
$$d_2'(t) = k \cdot d_1(t) - k \cdot d_2(t)\,, \qquad d_2(0) = 0 \tag{6.24}$$
$$\vdots \tag{6.25}$$
$$d_n'(t) = k \cdot d_{n-1}(t) - k \cdot d_n(t)\,, \qquad d_n(0) = 0 \tag{6.26}$$
$$w(t) = p(t) + d_1(t) + \cdots + d_n(t) \tag{6.27}$$

where w is the tumor weight, g denotes a tumor growth function and e is a drug-effect term. The model parameters are

$$\theta_1 = (\eta, \sigma, w_0, k)$$

where η are unperturbed tumor growth related parameters, σ the drug-effect related parameters, w_0 is the inoculated tumor weight and k is the transit rate.

Now we apply Theorem 4.4.1 from Chapter 4.4 to (6.22)-(6.27). The PKPD model (6.22)-(6.27) is a system with a TCM represented by (6.23)-(6.26), input

$$k_{in}(t) = e(\sigma, c(t)) \cdot p(t) \tag{6.28}$$

and initial density function $f \equiv 0$. Moreover, the proliferating cells $p(t)$ are governed by (6.22). On the way to a description of the pharmacological process with an LSM we set

$$d(t) = d_1(t) + \cdots + d_n(t)$$

representing the totality of cells attacked by the anticancer agent and replace the TCM (6.23)-(6.26) by a LSM for the population $d(t)$. Using (6.28) leads to

$$d'(t) = k_{in}(t) - k_{in}(t - T) = e(\sigma, c(t)) \cdot p(t) - e(\sigma, c(t - T)) \cdot p(t - T)$$

completed by the initial condition

$$d(0) = T \cdot \int_0^1 0\, ds = 0$$

and the history function

$$e(\sigma, c(s)) \cdot p(s) = 0\,, \quad -T \leq s < 0\,. \tag{6.29}$$

In applications (6.29) is fulfilled because no drug is administered before inoculation of the tumor cells.

Then the reformulation of (6.22)-(6.27) in the lifespan model context reads

$$p'(t) = g(\eta, p(t), d(t)) - e(\sigma, c(t)) \cdot p(t)\,, \qquad p(0) = w_0 \tag{6.30}$$
$$d'(t) = e(\sigma, c(t)) \cdot p(t) - e(\sigma, c(t-T)) \cdot p(t-T)\,, \qquad d(0) = 0 \tag{6.31}$$
$$w(t) = p(t) + d(t)\,. \tag{6.32}$$

In the LSM formulation (6.29)-(6.32) we have exactly two differential equations, one for the proliferating cells $p(t)$ and one governing the population of the attacked tumor cells $d(t)$. Note that it is not necessary to provide information about $p(s)$ for $-T \leq s < 0$ due to (6.29). See Figure 6.6 for schematic representation. The parameters are

$$\theta_2 = (\eta, \sigma, w_0, T)$$

where T is the lifespan of the dying tumor cells which is now fitted directly from the data.

This new formulation is from the modeling point of view a serious alternative to the classical formulation. Here the number of dying tumor stages is reduced to exactly one stage for the total population of cells attacked by the anticancer agent. This coincides with the situation in practice, where the choice of the number of compartments n is more or less arbitrary because the different stages could not be measured.

Finally note, that (6.29)-(6.31) is structurally equal to model III (4.56)-(4.57) from Section 4.5 but with a time-dependent first order rate $k(t) = e(\sigma, c(t))$.

We applied the model (6.29)-(6.32) with our tumor growth function and the linear drug-effect term to our data and summarized the fitting results in Table 6.4.

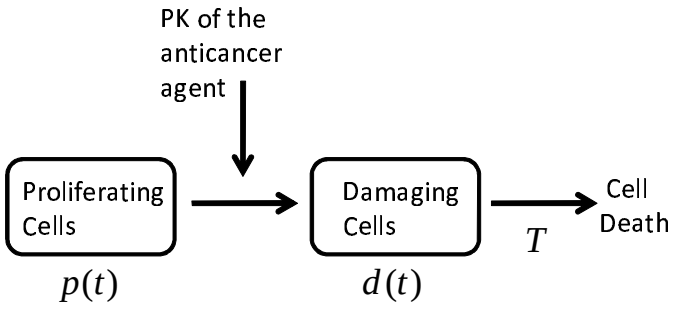


Figure 6.6: *Schematic overview of the PKPD model for unperturbed and perturbed tumor growth based on mono-therapy in lifespan representation.*

Parameter	Explanation	A1 180 mg/kg	A2 120 mg/kg	B 100 mg/kg	C 100 mg/kg	C 150 mg/kg
λ_0		0.127	0.121	0.194	0.134	0.192
(CV%)	Exponential rate	(26.2)	(30.6)	(36.6)	(28.4)	(34.6)
CI		[0.056, 0.198]	[0.042, 0.199]	[0.033, 0.356]	[0.053, 0.216]	[0.042, 0.342]
λ_1		0.293	0.313	0.245	0.282	0.240
(CV%)	Linear rate	(18.3)	(27.8)	(22.4)	(18.2)	(20.5)
CI		[0.179, 0.408]	[0.160, 0.465]	[0.121, 0.369]	[0.172, 0.391]	[0.129, 0.352]
w_0		0.057	0.059	0.010	0.051	0.011
(CV%)	Initial tumor weight	(56.9)	(67.5)	(>100)	(64.6)	(>100)
CI		[0, 0.126]	[0, 0.144]	[0, 0.038]	[0, 0.120]	[0,0.41]
k_{pot}		0.172	0.007	0.007	0.016	0.011
(CV%)	Drug potency	(12.8)	(12.3)	(14.6)	(81.9)	(30.5)
CI		[0.125, 0.219]	[0.005, 0.009]	[0.005, 0.010]	[0, 0.043]	[0.004, 0.020]
T		0.811	0.917	3.61	0.044	0.116
(CV%)	Lifespan	(>100)	(27.8)	(19.1)	(>100)	(>100)
CI		[0, 3.026]	[0.373, 1.461]	[2.054, 5.173]	[0, 11.8]	[0, 4.754]
	Sum of Squares	0.150	0.269	0.037	0.166	0.031
	R^2	0.99 / 0.99	0.99 / 0.91	0.99 / 0.99	0.99 / 0.99	0.99 / 0.99

Table 6.4: *Fitting values of the model (6.32)-(6.32) for the simultaneous fit with unperturbed and perturbed tumor growth for different drugs in xenograft mice. The sum of squares as well as the goodness of fit criteria R^2 (first: unperturbed growth, second: perturbed growth) are presented.*

6.7 Numerics

We solved the mono-therapy model (6.7)-(6.10) and the combination therapy model (6.13)-(6.16) in ordinary differential equation formulation with the `Matlab` internal solver `ode45`. The mono-therapy model (6.29)-(6.32) in delay differential equation formulation was solved with the `Matlab` internal solver `dde23`.

The fitting process was performed with the procedure `lsqcurvefit` from the `Matlab` Optimization Toolbox, where large- and medium-scale algorithms are applied. The gradient was calculated by variation equations (see e.g. [Ama95]) in case of ordinary differential equations and solved numerically in case of delay differential equations. The weights are $W = I$.

Also the PKPD software `ADAPT` (a `Fortran` based package) from the Biomedical Simulations Resource (BMSR) in the Department of Biomedical Engineering at the University of Southern California was used to fit data. `ADAPT` applies the Simplex Nelder-Mead algorithm (gradient free method) for optimization and uses `LSODA` (solver with automatic method switching for stiff and non-stiff problems) for solving ordinary differential equations, see [Pet84]. `ADAPT` was applied to the ordinary differential equation formulation (6.7)-(6.10) and (6.13)-(6.16).

The results from `Matlab` and `ADAPT` coincide within numerical errors.

6.8 Project structure

The results from Section (6.1)-(6.5) were carried out together with coworkers from Nycomed (A Takeda Company) namely, Dr. A. Walz (Biologist) and Dr. G. Lahu (Head of the Department Pharmacometrics), and Prof. Dr. J. Schropp from the University of Konstanz. Dr. Walz contributed to this project with fundamental biological knowledge in tumor growth. The results were published in the Journal of Pharmacokinetics and Pharmacodynamics (JPKPD) in April 2009, see [KWLS09]. On the website of JPKPD data and code is available for download. Section (6.1)-(6.5) were part of the collaboration *Numerical simulation of drug designing experiments (Project no. 735/06)* between Nycomed and the University of Konstanz.

Section 6.6 is part of the forthcoming publication [KS12] appearing in Spring 2012 in JPKPD.

6.9 Discussion and outlook

To our knowledge, our combination therapy PKPD tumor growth model for data measured in xenograft mice was the first published in the PKPD community.

Our published manuscript [KWLS09] was already frequently cited, where the most notable

citations are from P. Bonate and D. Howard [BD11] in the book *"Pharmacokinetics in Drug Development Volume 3: Advances and Applications"* and the overview article from Zhou et al [ZG11] named *"The Pharmacokinetic/Pharmacodynamic Pipeline: Translating Anticancer Drug Pharmacology to the Clinic"*.

However, due to the importance of combination therapy (especially in cancer development) this field is subject of strong ongoing research, also in mathematical modeling. Hence, several modeling approaches have been published in the last three years.

In 2009, Rocchetti and colleagues [RBG$^+$09] published a model to test additivity of anticancer agents. They proposed also a model with transit compartments but here every state of dying cells is connected among each other. Their approach does not fit combination therapy data and also does not quantify the effect of combination by an explicit parameter. Instead, the model simulates based on the parameter obtained from mono-therapy, if the model prediction lies above or under the combination therapy data. They recommend visual checks as well as statistical hypothesis testing to conclude for additive effects.

In 2010, Goteti and colleagues [GEGU$^+$10] presented a PKPD tumor growth model with a build-in synergistic term, especially designed to describe synergistic combination therapy data.

In 2011, Frances et al [FCBI11] published a model for clinical trials to describe the interaction between capecitabine and docetaxel used in combination in metastatic breast cancer. This model structurally differs from our approach and the cited PKPD models above.

Finally, we mention that the lifespan formulation (6.29)-(6.32) of our tumor growth model for mono-therapy opens the route to a new combination therapy approach. In this new approach not only the drug potencies k_{pot}^A and k_{pot}^B could be used to characterize drugs but also the lifespan T^A and T^B of the attacked tumor cells by either drug A or B from mono-therapy. Therefore, one would include more information of the drugs, which is actually available but still not used, in a new combination therapy approach. The development and application of such a model is currently ongoing.

Chapter 7

Modeling of Arthritis and Anti-GM-CSF Effects

Rheumatoid arthritis (RA) is an immune-mediated inflammatory disease and is characterized by a chronic inflammation and synovial hyperplasia leading to destruction of cartilage and bone, see [FM98]. Approximately one percent from the world-wide population suffers on RA. Interestingly, women are three times more affected than men. In most cases the disease starts at an age between 40 and 50 years.

The cytokine granulocyte-macrophage colony-stimulating factor (GM-CSF) is a key activator of the innate part of the immune system and as such involved in chronic stages of inflammatory and autoimmune diseases like RA, asthma or multiple sclerosis and GM-CSF was found aberrantly overproduced in such diseases, compare [AGZB$^+$91], [CPR$^+$98] and [GWJ03]. In arthritis, administration of recombinant GM-CSF for therapeutic purposes was found to aggravate arthritis and is therefore, one of the potential main drivers, see [VWB$^+$91] and [BZHC00]. In experimental models of arthritis, inhibition of GM-CSF reduces the intensity of the inflammation and thereby also lowers articular cartilage and bone destruction, see [CBC$^+$01] and [PZJH$^+$07].

In this project the arthritis development in mice is observed. In our experiments collagen induced arthritic (CIA) mice, which is a widely accepted model, are used. This model shares several clinical, histopathological and immunological features with human RA, see [HBBY02], [Wil04] and [BVT10]. The monoclonal antibody 22E9 was used in order to neutralize the biological activity of murine GM-CSF in CIA mice.

In 2008, Earp and colleagues [EDM$^+$08b], [EDM$^+$08a] presented a mathematical model for arthritis development in CIA rats. They measured and modeled different cytokines and the paw size. To describe existing delays they used transit compartments.

In this project we develop a multi-response PKPD model to describe either a total arthritic score as well as a ankylosis score. To capture existing delays we use delay differential equations.

In Section 7.1 we present the experimental setup. In Section 7.2 we develop the PKPD model based on pharmacological assumptions with delay differential equations. In the next Section 7.3 the fitting results are presented. In Section 7.4 we use the model as example and apply the results from Section 4.4.

7.1 Experimental setup

The CIA mouse model (DBA/1 mice, Taconic Farms, 8-9 weeks old) consists of an arthritic induction phase and an arthritic development phase. The arthritic induction phase starts with an initial immunization with collagen. 21 days later a booster injection with collagen is administered. The time point $t = 0$ is the day before the day of onset ($t = 1$) when first signs of arthritis are detected, which is the start of the arthritic development phase. Around 10 days later, also first signs of ankylosis, describing the bone and cartilage destruction in joints, is detected. See Figure 7.1 for a schematic overview of the CIA mouse model. In the experiments two different scorings were performed in the CIA mouse model,

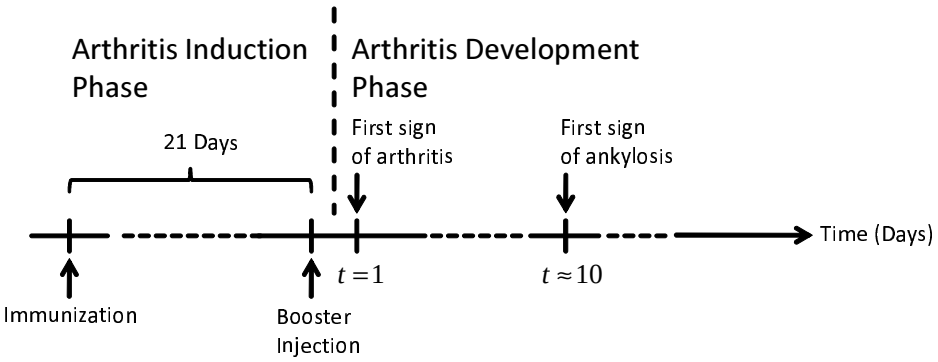


Figure 7.1: *Schematic overview of the CIA mouse model.*

namely a discrete total arthritic score (TAS), which is a descriptive overall measurement of the disease development and a discrete ankylosis score (AKS), representing bone and cartilage destruction in the joints. The appearance of ankylosis was delayed to the first sign of inflammation. The readout TAS consists of integers ranging from 0 to 4 for each individual paw and hence a range of 0-16 per animal is possible. The readout AKS was originally scored from 0-2 for each individual paw and results in a range of 0-8 per animal, see Table 7.1. In order to achieve a consistent scoring scale between TAS and AKS we multiplied the AKS scores by a factor 2. Both TAS and AKS are visual scores. In all experiments the drug treatment was administered i.v. and the PK and the PD readouts where measured in the same mice. In experiment A the PK was measured at 8 time points per mouse and at the terminal bleed. In experiment B the PK was measured only once per mouse and at the terminal bleed for all animals. The PD readouts were taken every other day in both experiments. Table 7.1 shows the dosing schedule for the different experiments. In experiment A and B some mice were found dead and were therefore excluded from our analysis. Overall 82 mice were treated in the experiments with either

Total Arthritic Score (TAS)	Observation	Ankylosis Score (AKS)	Observation
0	Normal Paw	0	No Ankylosis
1	One toe inflamed and swollen.	1	Mild Ankylosis
2	More than one toe, but not entire paw, inflamed and swollen, or mild swelling of entire paw.	2	Severe Ankylosis
3	Entire paw inflamed and swollen.		
4	Very inflamed and swollen paw or ankylosed paw. If the paw is ankylosed, the mouse cannot grip the wire top of the cage.		

Table 7.1: *The total arthritic score (TAS) and the ankylosis score (AKS).*

Experiment A	Experiment B
Placebo on day 1,8,15	Placebo on day 1,8,15
1 mg/kg on day 1,8,15	0.1 mg/kg on day 1,8,15
10 mg/kg on day 1,8,15	1 mg/kg on day 1,8,15
100 mg/kg on day 1,15	10 mg/kg on day 1,8,15

Table 7.2: *Dosing schedules of monoclonal antibody 22E9 for experiment A and B.*

vehicle or 22E9 and 72 were included in the analysis.

In experiment A a non-monotonic data behavior regarding *dose* and effect was observed. In more detail the TAS and the AKS over time of the 10 mg/kg group laid above the 1 mg/kg group. On the other hand in experiment B the data showed a monotonic behavior. This discrepancy might be explained by the intense blood sampling for PK analysis in experiment A, which induced stress and possible infections to the animals thereby altering the PD readout. Therefore, sparse PK sampling was applied in experiment B.

7.2 Model development

Starting point of the arthritis development is the inflammation $I(t)$ driven by the cytokine GM-CSF denoted by $G(t)$. After a while the destruction of the joints starts as a result of the existing inflammation and becomes more and more important. The bone destruction $D(t)$ is visualized by joint ankylosis. The underlying inflammation does not subside and remains in a steady state as an important hallmark of the disease, see [Wil04]. The ad-

ministered antibody 22E9 neutralizes the cytokine GM-CSF.

Thus the following assumptions were made:

(A1) The cytokine GM-CSF drives the inflammation and the cartilage and bone destruction.

(A2) The arthritis starts with the inflammatory part which dominates the disease for 1-2 weeks, afterwards inflammation decreases but does not vanish completely and remains at a certain level.

(A3) The destructive part of the disease is delayed.

(A4) The antibody 22E9 directly acts on the cytokine GM-CSF.

The task to develop the full PKPD model is divided into three steps. Firstly, we present an equation for the cytokine GM-CSF. Secondly, we model the responses TAS and AKS based on the assumptions (A1)-(A3) and finally we include an effect term, which describes the inhibition of GM-CSF in our model based on (A4).

Modeling Step 1: The cytokine behavior in time

Based on assumption (A1) the cytokine GM-CSF drives the disease. Hence, the first step is to set up a model to describe the time course of GM-CSF. This is done with the use of a classical inflow/outflow model. First, we consider the following coupled inflow/outflow models

$$x_1'(t) = a_1 - a_2 x_1(t)\,, \qquad x_1(0) = 0 \tag{7.1}$$
$$x_2'(t) = a_3 - a_4 x_1(t) x_2(t)\,, \quad x_2(0) = x_2^0 \geq 0\,. \tag{7.2}$$

We explicitly solve (7.1) and obtain the system

$$x_2'(t) = a_3 - a_4 \frac{a_1}{a_2}(1 - \exp(-a_2 t)) x_2(t)\,, \quad x_2(0) \geq 0\,. \tag{7.3}$$

One immediately notices that the system (7.3) is over-parameterized because of the product $a_4 \cdot a_1$ and due to the fact that a_4 and a_1 do not appear elsewhere on the right hand side of (7.3). Hence, we set $k_1 = a_4 a_1$, $k_2 = k_2$ and $k_3 = a_3$ and obtain the model

$$G'(t) = k_3 - \frac{k_1}{k_2}\left(1 - \exp(-k_2 t)\right) G(t)\,, \quad G(0) \geq 0 \tag{7.4}$$

for the cytokine GM-CSF. The solution $G(t)$ of (7.4) can realize either monotonic or non-monotonic behavior in time, regarding to the actual values of the parameters k_1, k_2 and k_3. Earp and colleagues, see [EDM+08b] and [EDM+08b], measured different cytokines in CIA rats which show monotone or non-monotone behavior. We could not measure GM-CSF in plasma due to volume constraints in the CIA mice but expect the same qualitative behavior as the cytokines measured by Earp et al.

Modeling Step 2: Multi-response approach to model the TAS and AKS

In the next step towards a mathematical model the arthritic disease is split into an inflammatory part $I(t)$ and an ankylosis (bone and cartialge destruction) part $D(t)$ and the sum

$$R_1(t) = I(t) + D(t) \tag{7.5}$$

is defined as the response $R_1(t)$, which is fitted against the measured TAS. In addition the second response function is defined as

$$R_2(t) = D(t) \tag{7.6}$$

which is fitted against the AKS data.

To build a model for the time course of the inflammation $I(t)$ and the ankylosis $D(t)$ we adapt the concept of lifespan modeling introduced in Section 4.3. Similar to that concept the overall inflammation $I(t)$ is controlled by two processes, the inflow $k_{in}(t)$ and the outflow $k_{out}(t)$. Assuming that the inflammation caused by these processes remains a certain time period T and is driven by the amount of GM-CSF, one obtains

$$k_{out}(t) = k_{in}(t - T)$$

and

$$k_{in}(t) = k_4 G(t)$$

where $k_4 > 0$ is a first order rate constant. Then the total balance equation for the inflammation reads

$$I'(t) = k_{in}(t) - k_{out}(t) = k_4 G(t) - k_4 G(t - T)\,. \tag{7.7}$$

Finally, for the ankylosis $D(t)$ one obtains based on the assumption (A1)-(A2)

$$k_{in}(\text{ankylosis}) = k_{out}(\text{inflammation})\,.$$

Applying a first order loss term

$$k_{out}(\text{ankylosis}) = k_5 D(t)$$

leads to the equation

$$D'(t) = k_4 G(t - T) - k_5 D(t)\,. \tag{7.8}$$

The presence of $G(t)$ and $G(t - T)$ in (7.7) and (7.8), respectively, reflect that the inflammation and the ankylosis is driven by GM-CSF. Moreover, the action of GM-CSF in respect to the ankylosis is delayed by T.

It is realistic to assume that an increase of GM-CSF already starts after the immunization in the CIA model. Therefore, it is reasonable to take the initial function $G_0(s)$, $-T \leq s \leq 0$ monotonic increasing. Furthermore, because the GM-CSF-producing cells

divide and proliferate in response to the collagen immunization, it is realistic to assume an exponential growth of the cytokine. The initial function is then of the form

$$G_0(s) = a\exp(bs) \quad \text{for } -T \leq s \leq 0\,,\ a, b > 0. \tag{7.9}$$

Because the cytokine GM-CSF is modeled in a qualitative and not a quantitative manner the parameter a and b in (7.9) will be fixed during the fitting process.

A consequence of the early start of the GM-CSF production described by (7.9) is that at $t = 0$ already some weak inflammation exists, but is still scored with the value 0 in the discrete TAS scheme by the experimenter. Therefore, the initial value of the inflammation in (7.7) is set to $I(0) = I_0 > 0$. Finally, the model equations (7.4),(7.7)-(7.8) completed by the initial conditions (7.9), $I(0) = I_0 > 0$ and $D(0) = 0$ describe the unperturbed arthritis development.

Modeling Step 3: The final PKPD model with the influence of the antibody 22E9 at the cytokine GM-CSF

The antibody 22E9 acts directly on the cytokine GM-CSF $G(t)$, see (A4). Thus it is obvious that equation (7.4) has to be amended by an effect term to obtain a PKPD model. In experiment A and B different *dose* levels were administered. It turned out in our experiments that the effect of the drug is highly non-linear in respect to the amount of *dose*. Sometimes it was observed that the effect is not even monotonic regarding to *dose*-effect relationship. Therefore, we apply the drug-effect term (4.79) derived in Chapter 4.6.

The final PKPD model

The final PKPD model for unperturbed and perturbed arthritis development in CIA mice reads

$$\begin{aligned}
G'(t) &= k_3 - \left(\sigma_1 \exp(-\sigma_2 c(t)) + \sigma_3\right) c(t) G(t) \qquad G(s) = a\exp(bs) \qquad (7.10)\\
&\quad -\frac{k_1}{k_2}\left(1 - \exp(-k_2 t)\right) G(t)\,, \qquad \text{for } 0 \geq s \geq -T\\
I'(t) &= k_4 G(t) - k_4 G(t-T)\,, \qquad I(0) = I_0 > 0 \qquad (7.11)\\
D'(t) &= k_4 G(t-T) - k_5 D(t)\,, \qquad D(0) = 0 \qquad (7.12)\\
R_1(t) &= I(t) + D(t) \qquad (7.13)\\
R_2(t) &= D(t) \qquad (7.14)
\end{aligned}$$

with the model parameters

$$\theta = (k_1, k_2, k_3, k_4, k_5, T, I_0, a, b, \sigma_1, \sigma_2, \sigma_3)\,. \tag{7.15}$$

Hence, the final model consists of three meaningful compartments, namely GM-CSF, inflammation and ankylosis. In Figure 7.2 a basic diagram of the model (7.10)-(7.14) is presented. Note that (7.10)-(7.12) is structurally similar to model II (4.53)-(4.55) from Section 4.5.

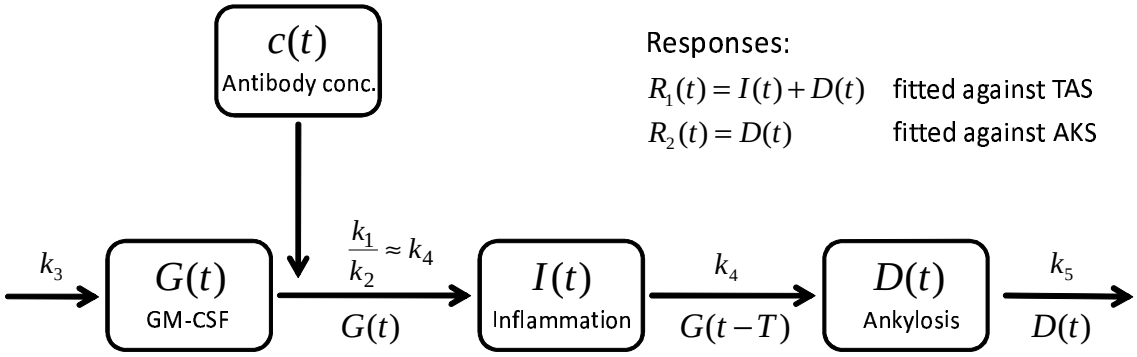


Figure 7.2: *Schematic overview of the PKPD model (7.10)-(7.14).*

Proposition 7.2.1
The steady state of (7.10)-(7.12) reads

$$G^* = \frac{k_3 k_2}{k_1}, \quad I^* = I_0 + k_4 T G^* - \frac{k_4 a}{b}\left(1 - \exp(-bT)\right) \quad \textit{and} \quad D^* = \frac{k_4}{k_5} G^*.$$

Proof: Because $\lim_{t\to\infty} c(t) = 0$ we obtain for Eq. (7.10)

$$0 = k_3 - \frac{k_1}{k_2} G^* \quad \Rightarrow \quad G^* = \frac{k_3 k_2}{k_1}. \tag{7.16}$$

The solution of (7.11) is of the form

$$\begin{aligned} I(t) = k_4 \int_0^t G(\tau) - G(\tau - T) d\tau + I_0 &= k_4 \left(\int_0^t G(\tau) d\tau - \int_0^t G(\tau - T) d\tau \right) + I_0 \\ &= k_4 \left(\int_0^t G(\tau) d\tau - \int_{-T}^{t-T} G(z) dz \right) + I_0 . \end{aligned} \tag{7.17}$$

Because $G(t)$ is defined for $t \leq 0$ we could split the integrals in (7.17) and obtain

(7.18)

$$\begin{aligned} I(t) &= k_4 \left(\int_0^{t-T} G(\tau) d\tau + \int_{t-T}^{t} G(\tau) d\tau - \int_{-T}^{0} G(\tau) d\tau - \int_0^{t-T} G(\tau) d\tau \right) + I_0 \\ &= k_4 \left(\int_{t-T}^{t} G(\tau) d\tau - \int_{-T}^{0} G(\tau) d\tau \right) + I_0 . \end{aligned} \tag{7.19}$$

Then using $\lim_{t\to\infty} G(t) = G^*$ we have

$$\lim_{t\to\infty} \int_{t-T}^{t} G(\tau) d\tau = T G^* .$$

Therefore,

$$I^* = \lim_{t\to\infty} I(t) = k_4 T G^* - k_4 \int_{-T}^{0} G_0(\tau)d\tau + I_0 \tag{7.20}$$

and with $G_0(s) = a\exp(bs)$ for $s \leq 0$ we obtain

$$\int_{-T}^{0} G_0(\tau)d\tau = \int_{-T}^{0} a\exp(b\tau)d\tau = \frac{a}{b}\left(1-\exp(-bT)\right) \tag{7.21}$$

leading to

$$I^* = k_4 T G^* - k_4 \frac{a}{b}\left(1-\exp(-bT)\right) + I_0\,. \tag{7.22}$$

Finally, using (7.12) we immediately obtain

$$0 = k_4 G^* - k_5 D^* \quad \Longrightarrow \quad D^* = \frac{k_4}{k_5} G^*\,. \tag{7.23}$$

Hence, the steady states of (7.10)-(7.12) are (7.16), (7.22) and (7.23).
□

Finally, we reformulate the DDE (7.10)-(7.12) as ordinary differential equation. Due to the special structure of the model the method of steps (see for example [Dri77]) reduces to exactly two steps. We formulate the DDE (7.10)-(7.12) slightly more general and denote by $g(c(t))$ an arbitrary effect term and by $h(t)$ an initial function. The model then reads

$$\begin{aligned}
y_1'(t) &= k_3 - g(c(t))y_1(t) - \frac{k_1}{k_2}\left(1-\exp(-k_2 t)\right) y_1(t)\,, && y_1(s) = h(s) && (7.24)\\
&&& \text{for } s \in [-T, 0] \\
y_2'(t) &= k_4 y_1(t) - k_4 y_1(t-T)\,, && y_2(0) = y_2^0 && (7.25)\\
y_3'(t) &= k_4 y_1(t-T) - k_5 y_3(t)\,, && y_3(0) = 0\,. && (7.26)
\end{aligned}$$

The first step is to substitute the initial function $h(s)$ for $y_1(t-T)$ into the right hand side of (7.24)-(7.26). Hence, we obtain the following ODE system for $0 \leq t \leq T$

$$\begin{aligned}
x_1'(t) &= k_3 - g(c(t))x_1(t) - \frac{k_1}{k_2}\left(1-\exp(-k_2 t)\right) x_1(t)\,, && x_1(0) = h(0) && (7.27)\\
x_2'(t) &= k_4 x_1(t) - k_4 h(t-T)\,, && x_2(0) = x_2^0 && (7.28)\\
x_3'(t) &= k_4 h(t-T) - k_5 x_3(t)\,, && x_3(0) = 0\,. && (7.29)
\end{aligned}$$

Let $(x_1^T, x_2^T, x_3^T) = (x_1(T), x_2(T), x_3(T))$ be the solution of (7.27)-(7.29) for the time point $t = T$. By adding an additional ordinary differential equation with the property $x_4(t) =$

$x_1(t-T)$ for all $t \geq T$ we obtain the system

$$\begin{aligned}
x_1'(t) &= k_3 - g(c(t))x_1(t) - \frac{k_1}{k_2}\left(1-\exp(-k_2 t)\right)x_1(t)\,, & x_1(T) = x_1^T \quad (7.30)\\
x_2'(t) &= k_4 x_1(t) - k_4 x_4(t)\,, & x_2(T) = x_2^T \quad (7.31)\\
x_3'(t) &= k_4 x_4(t) - k_5 x_3(t)\,, & x_3(T) = x_3^T \quad (7.32)\\
x_4'(t) &= k_3 - g(c(t-T))x_4(t) \\
&\quad -\frac{k_1}{k_2}\left(1-\exp(-k_2(t-T))\right)x_4(t)\,, & x_4(T) = x_1^0 = h(0) \quad (7.33)
\end{aligned}$$

for all $t \geq T$.

Hence, the ODE formulation of the DDE (7.24)-(7.26) is (7.27)-(7.29) for $0 \leq t \leq T$ and (7.30)-(7.33) for $t \geq T$.

7.3 Fitting results

We simultaneously fitted in experiment A and in experiment B all available dosing groups. To describe the results of the experiments A and B four outputs are created, the fit of the TAS data (left upper panel), the fit of the AKS data (right upper panel) as well as the prediction of the qualitative behavior of GM-CSF (left lower panel) and of the inflammation (right lower panel). The bars denote the standard deviation, calculated from the individual data.

In Figure 7.3 the results for experiment A are presented. Here we have the non-monotonic *dose*-effect relationship, more precisely, the 1 mg/kg group shows a higher effect than the 10 mg/kg group. In Figure 7.4 the results for experiment B are presented. The bars denote the standard deviation calculated on the basis of the individual data. For both experiments the developed PKPD model describes the data adequately. Also our simulations of GM-CSF (see Figure 7.3 and 7.4) show the same qualitative behavior as the cytokines measured by Earp et al [EDM$^+$08b],[EDM$^+$08a]. In Table 7.3 the estimates of the parameters are presented. In both experiments the delay parameter T has the lowest coefficient of variation compared to all other parameters. This shows the stability of the estimate of the delay T. Also the actual estimates of T completely coincide with the situation seen in the data.

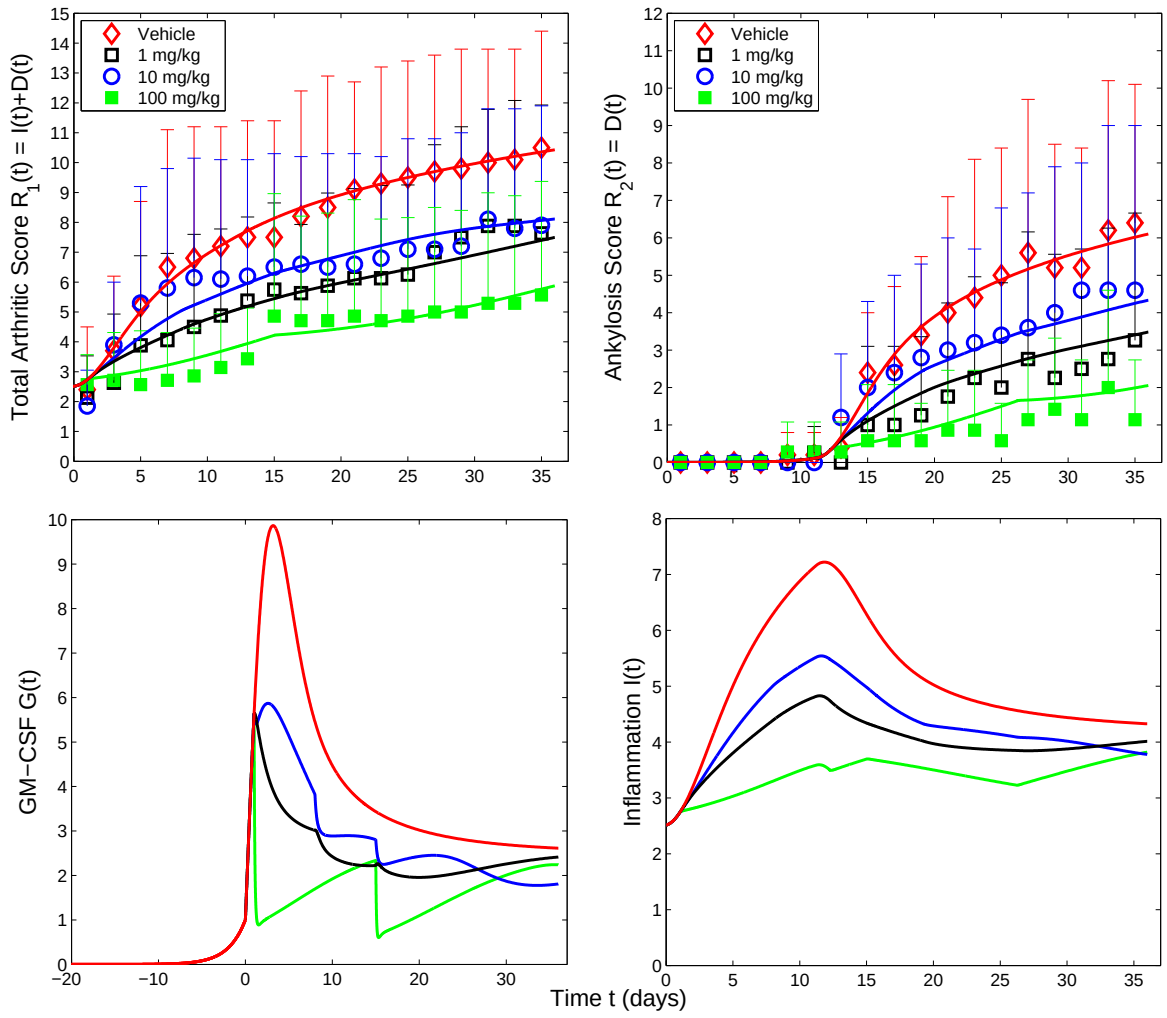


Figure 7.3: *Experiment A: Note that in this experiment the effect of the 1 mg/kg is higher than the effect of 10 mg/kg. The red diamonds denote the vehicle group, the blue open circles denote the 10 mg/kg group, the black open squares denote the 1 mg/kg group and the green filled squares denote the 100 mg/kg group. The left and right upper panel show the fits of the total arthritic score and the ankylosis score, respectively. In the left lower panel the qualitative behavior of the GM-CSF is presented and in the right lower panel the inflammation is plotted.*

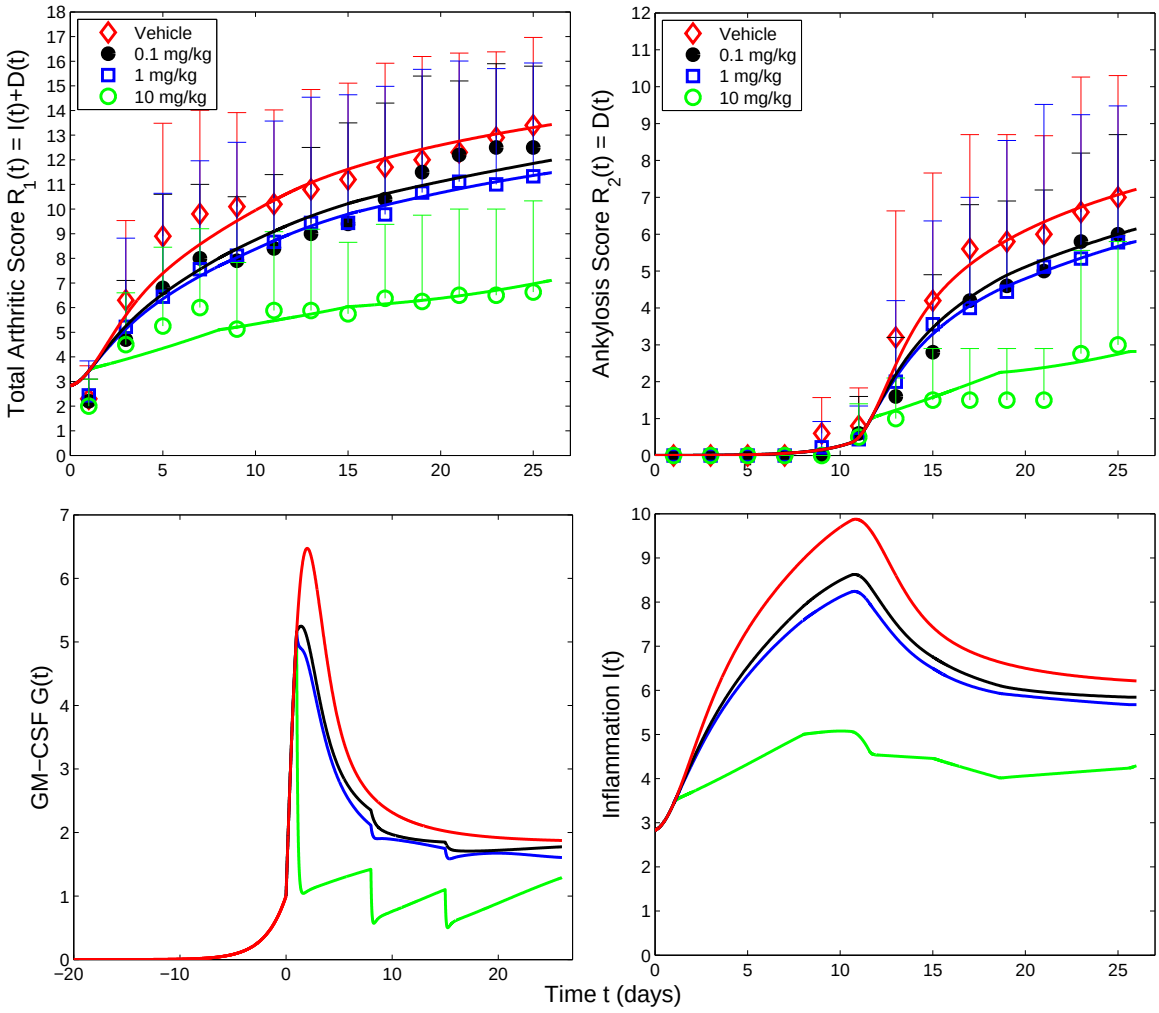


Figure 7.4: *Experiment B: The red diamonds denote the vehicle group, the black filled circles denote the 0.1 mg/kg group, the blue open squares denote the 1 mg/kg group and the green open circles denote the 10 mg/kg group. The panels are defined as in Figure 7.3.*

Parameter	Explanation	Experiment A Value (CV%) CI	Experiment B Value (CV%) CI
k_1	Outflow GM-CSF	0.183 (32) [0.065, 0.302]	0.456 (42) [0.079, 0.832]
k_2	Outflow GM-CSF	0.092 (26) [0.045, 0.140]	0.169 (25) [0.085, 0.252]
k_3^+	Inflow GM-CSF	5	5
k_4	Inflow/Outflow Inflammation; Inflow Ankylosis	0.064 (23) [0.034, 0.093]	0.185 (30) [0.076, 0.292]
k_5	Outflow Ankylosis	0.016 (26) [0.008, 0.021]	0.031 (21) [0.018, 0.044]
σ_1	Effect term parameter	0.154 (44) [0.019, 0.289]	0.328 (41) [0.063, 0.593]
σ_2	Effect term parameter	0.065 (39) [0.014, 0.116]	0.328 (25) [0.165, 0.491]
σ_3	Effect term parameter	0.003 (23) [0.002, 0.005]	0.025 (35) [0.008, 0.042]
T	Delay Ankylosis	11.2 (4.4) [10.24, 12.21]	10.6 (4.3) [9.688, 11.48]
I_0	Initial value Inflammation	2.52 (5.4) [2.24, 2.79]	2.83 (8.5) [2.35, 3.31]
a^+	Initial function parameter GM-CSF	1	1
b^+	Initial function parameter GM-CSF	0.5	0.5
	Sum of squares	21.15	23.72
	R^2 for $R_1(t)$	0.99 / 0.98 / 0.87 / 0.91	0.94 / 0.96 / 0.99 / 0.73
	R^2 for $R_2(t)$	0.99 / 0.96 / 0.98 / 0.82	0.99 / 0.99 / 0.99 / 0.90

Table 7.3: *PKPD model parameters of (7.10)-(7.14). Parameters denoted by superscript + were fixed during the fitting process. Model parameter estimates, coefficient of variation, 95%-confidence interval and R^2 of the simultaneous fit are presented.*

7.4 Reformulation as transit compartment based model

In this section we apply Theorem 4.4.1 from Section 4.4 and rewrite the model (7.10) - (7.14) as a transit compartment model.

The inflammation (see (7.11)) in the arthritis model is governed by an LSM with input

$$k_{in}(t) = k_4 \cdot G(t)$$

On the way towards a TCM realization of the model we replace the lifespan equation (7.11) by a TCM with n states and obtain

$$x_1'(t) = k_4 \cdot G(t) - k \cdot x_1(t) \tag{7.34}$$
$$x_2'(t) = k \cdot x_1(t) - k \cdot x_2(t) \tag{7.35}$$
$$\vdots$$
$$x_n'(t) = k \cdot x_{n-1}(t) - k \cdot x_n(t) \tag{7.36}$$

completed by the initial values

$$x_i(0) = \frac{1}{k} \cdot f\left(\frac{i}{n}\right), \quad i = 1, ..., n, \quad k = \frac{n}{T}. \tag{7.37}$$

According to equation (4.29) the initial density function f is based on the past $G(s)$, $-T \leq s \leq 0$ that is

$$f(t) = k_{in}(-Tt) = k_4 \cdot G(-Tt) = k_4 \cdot a \exp(-bTt) \quad \text{for } 0 \leq t \leq 1.$$

Finally, as a consequence we eliminate the term $k_4 \cdot G(t-T)$ from equation (7.12) as well as $I(t)$ from (7.13) and obtain by the use of (4.25) and (4.28)

$$G'(t) = k_3 - (\sigma_1 \exp(-\sigma_2 c(t)) + \sigma_3)\, c(t) G(t), \qquad G(0) = a \tag{7.38}$$
$$- \frac{k_1}{k_2}(1 - \exp(-k_2 t))\, G(t)$$
$$x_1'(t) = k_4 G(t) - k x_1(t), \qquad x_1(0) = \frac{1}{k} f\left(\frac{1}{n}\right) \tag{7.39}$$
$$x_2'(t) = k x_1(t) - k x_2(t), \qquad x_2(0) = \frac{1}{k} f\left(\frac{2}{n}\right) \tag{7.40}$$
$$\vdots$$
$$x_n'(t) = k x_{n-1}(t) - k x_n(t), \qquad x_n(0) = \frac{1}{k} f(1) \tag{7.41}$$
$$D'(t) = k x_n(t) - k_5 D(t), \qquad D(0) = 0 \tag{7.42}$$
$$R_1(t) = x_1(t) + \cdots + x_n(t) + D(t) \tag{7.43}$$
$$R_2(t) = D(t) \tag{7.44}$$

with the model parameter

$$\theta_2 = (k_1, k_2, k_3^+, k_4, k_5, a^+, b^+, \sigma_1, \sigma_2, \sigma_3, k). \tag{7.45}$$

A short look on the list of parameters (see (7.15) and (7.45)) shows that there is a difference between the arthritis model in LSM and TCM formulation. Equations (7.10)-(7.14) make sense for the initial condition $I(0) = I_0$ with $I_0 > 0$ arbitrary. But the TCM arthritis equations (7.38)-(7.44) converge in the limit $n \to \infty$ towards the LSM model (7.10)-(7.14) with I_0 fixed according to

$$I_0 = T \int_0^1 f(s)\, ds = T \int_0^1 k_4 a \exp(-bTs)\, ds = \frac{k_4 a}{b} \left(1 - \exp(-bT)\right) \tag{7.46}$$

(see 4.30) to which the LSM realization is really equivalent to.

Nevertheless, to demonstrate the potential of the TCM formulation even with one parameter less we fix the number of compartments to $n = 5$ and fit the parameters (7.45) with the data from experiment B, see Figure 7.4. The estimates of the parameter are listed in Table 7.4. Note that the significant difference in the parameter estimates is due to two reasons. First, the original arthritis model in LSM formulation consists of an additional fitting parameter I_0. Second, the arthritis model in TCM formulation is also from the pharmacological point of view a different model. In the original model the drug acts not until time T on the ankylosis, whereas this is not the case in the TCM formulation.

Parameter	Explanation	Experiment B Value (CV%) CI
k_1	Outflow GM-CSF	1.608 (19) [1.011, 2.204]
k_2	Outflow GM-CSF	0.131 (23) [0.069, 0.192]
k_3^+	Inflow GM-CSF	5
k_4	Inflow/Outflow Inflammation; Inflow Ankylosis	0.746 (12) [0.567, 0.926]
k_5	Outflow Ankylosis	0.035 (22) [0.020, 0.05]
σ_1	Effect term parameter	0.884 (33) [0.303, 1.465]
σ_2	Effect term parameter	0.307 (26) [0.148, 0.488]
σ_3	Effect term parameter	0.066 (26) [0.032, 0.100]
k	Transit rate	0.329 (3.7) [0.305, 0.353]
a^+	Initial function parameter GM-CSF	1
b^+	Initial function parameter GM-CSF	0.5
	Sum of Squares	30.07
	R^2 for $R_1(t)$	0.97 / 0.96 / 0.99 / 0.76
	R^2 for $R_2(t)$	0.96 / 0.98 / 0.97 / 0.92

Table 7.4: *PKPD model parameters of the TCM formulation (7.38)-(7.44). Parameters denoted by superscript + were fixed during the fitting process. Model parameter estimates, coefficient of variation, 95%-confidence interval and R^2 of the simultaneous fit are presented.*

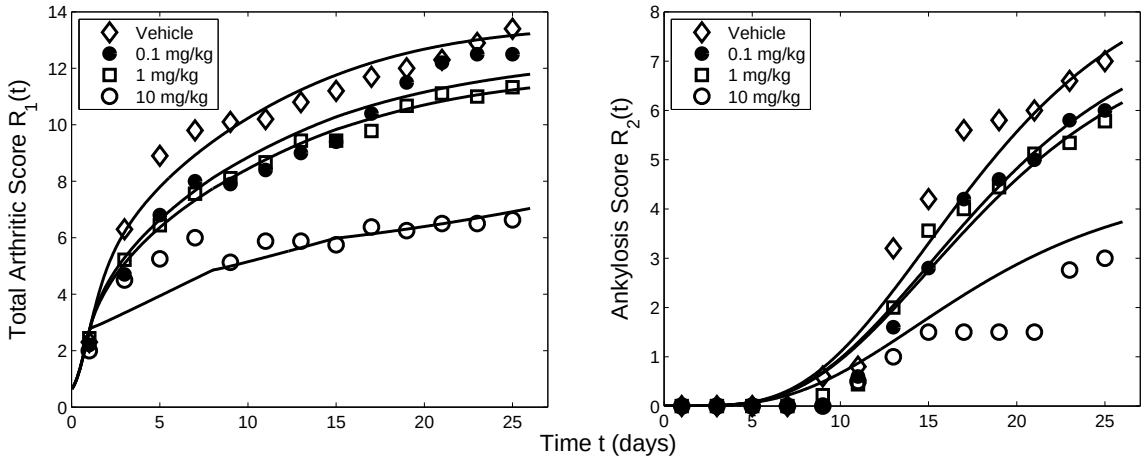


Figure 7.5: *Experiment B: TCM formulation (7.38)-(7.44).*

7.5 Numerics

We solved the arthritis model in DDE formulation (7.10)-(7.12) by the `Matlab` solver `dde23` and the ODE formulation (7.27)-(7.33) and (7.38)-(7.44) with the `Matlab` solver `ode45`.

The fitting process was performed with the procedure `lsqcurvefit` from the `Matlab` Optimization Toolbox where large- and medium-scale algorithms were applied. The gradient was calculated numerically. The weights are $W = I$.

We also applied the PKPD software `ADAPT` which is a `Fortran` based package from the Biomedical Simulations Resource (BMSR) in the Department of Biomedical Engineering at the University of Southern California to the ODE formulation (7.27)-(7.33) and (7.38)-(7.44), see [DSW09]. `ADAPT` applies the Simplex Nelder-Mead algorithm for optimization and uses LSODA (solver with automatic method switching for stiff and non-stiff problems) for solving ordinary differential equations, see [Pet84].

The results from `Matlab` and `ADAPT` coincide within numerical errors.

7.6 Project structure

The results from Section (7.1)-(7.3) were carried out together with coworkers from Nycomed (A Takeda Company) namely, Dr. T. Wagner (Chemist) and Dr. G. Lahu (Head of the Department Pharmacometrics), from Micromet namely, Dr. C. Plater-Zyberk and Prof. Dr. J. Schropp from the University of Konstanz. Dr. Wagner supported the project with his pharmacological knowledge. Dr. Plater-Zyberk has more than 25 years experience in arthritis development in mice and contributed with her biological knowledge to formulate the model assumptions (A1)-(A4).

Sections (7.1)-(7.3) were published in the Journal of Pharmacokinetics and Pharmacodynamics (JPKPD) in Janury 2012, see [KWPZ+12]. On the website of JPKPD all data is available for download. Sections (7.1)-(7.3) were part of the collaboration *Numerical simulation of drug designing experiments (Project no. 735/06)* between Nycomed and the University of Konstanz.

Section 7.4 is part of the forthcoming publication [KS12] appearing in Spring 2012 in JPKPD.

7.7 Discussion and outlook

To our knowledge, the model from Earp et al, see [EDM+08b]-[EDM+08a], for CIA data to handle cytokines as well as paw swelling from 2009 was the first mathematical PKPD model for arthritis development. In this model, transit compartments with $n \in \{19, 24, 29\}$ where applied to account for delays.

In our experiments, the readouts TAS and AKS are visual scores. But our model also simulates the qualitative behavior of the cytokine (could not be measured in mice) as well as the inflammation (part of the TAS score) of the paws. We identified the inflammation as a population and applied the lifespan approach. Therefore, our model consists of just three differential equations.

We remark that the CIA mouse model with the two performed readouts is the perfect example for the use of delay differential equations. First, it exists an important past of the cytokines (arthritis induction phase) before the first measurements. This is modeled by the initial function which is described by a realistic exponential approach. Second, the strongly delayed appearance of the ankylosis driven by the cytokines is perfectly implemented by a lifespan approach.

Moreover, to our knowledge our presented model is the first which describes simultaneously the total arthritis disease as well as the bone and cartilage destruction (which is an important property of the disease).

Appendix A

Laplace transform

The Laplace transform (see [Wid66] or [Doe76]) is an important tool to analytically solve linear ordinary differential equations.

A function $f(t)$ living in the so-called time domain is transformed by the Laplace transform (integral transformation) into the so-called image domain. In the image domain, the transformed function could be easier treated. Finally, the modified object from the image domain is transformed back to the time domain.

Definition A.0.1 (Laplace transform)
Let $f : [0, \infty) \to \mathbb{C}$. The Laplace transform of $f(t)$ is defined by

$$F(s) = \mathcal{L}\{f(t)\} = \mathcal{L}\{f\}(s) = \int_0^\infty \exp(-st) f(t)\, dt$$

where $s \in \mathbb{C}$.

The Laplace transform in the image domain is linear

$$\mathcal{L}\{af(t) + bg(t)\} = a\mathcal{L}\{f(t)\} + b\mathcal{L}\{g(t)\}.$$

Proposition A.0.2 (Existence of the Laplace transform)
Let $f : [0, \infty) \to \mathbb{C}$ and $C > 0, s_0 > 0$. Further it exists a $T > 0$ with

$$|f(t)| \leq C \exp(s_0 t) \quad \text{for } t > T\,.$$

If also

$$\int_0^T |f(t)| dt \leq \infty$$

then the Laplace transform $\mathcal{L}\{f\}(s)$ exists in the half-plane $Re(s) > s_0$.

Definition A.0.3 (Inverse Laplace transform)

$$f(t) = \mathcal{L}^{-1}\{F(s)\} = \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} \exp(st)F(s)ds$$

where γ is a real number so that the contour path of integration is in the region of convergence of $F(s)$.

Proposition A.0.4 (Uniqueness - Lerch's theorem)
If the Laplace transforms of f and g exists and if $\mathcal{L}\{f(t)\}(s) = \mathcal{L}\{g(t)\}(s)$ for all s with sufficient large real part, then

$$f(t) = g(t)$$

in every t, where both functions are continuous.

Theorem A.0.5 (Heaviside's theorem)
Let

$$p(s) = \sum_{i=1}^{m} b_i s^i \quad \text{and} \quad q(s) = \sum_{i=1}^{n} a_i s^i$$

be polynomials with $s \in \mathbb{C}$. Further let $m < n$ and $q(s)$ has distinct roots λ_i for $i = 1, ..., n$. Then

$$\mathcal{L}^{-1}\left\{\frac{p(s)}{q(s)}\right\} = \sum_{i=1}^{n} \frac{p(\lambda_i)}{q'(\lambda_i)} \exp(\lambda_i t)\,.$$

Proposition A.0.6 (Transform table)
a.) $\mathcal{L}\{f'(t)\} = s \cdot \mathcal{L}\{f(t)\} - f(0)$

Appendix B

Pharmacokinetic parameters of anticancer drugs

Parameter	Unit	Drug A2 120mg/kg	Drug B 100 mg/kg	Drug A1 120 mg/kg	Drug A1 180 mg/kg	Drug C 100 mg/kg	Drug C 150 mg/kg
		Value (CV%)	Value (CV%)	Value (CV%)	Value	Value (CV%)	Value
k_a	[1/h]	5.54 (40.4)	104 (>100)	4.42 (45.8)	4.42	84.9 (>100)	84.9
k	[1/h]	0.155 (30.0)	0.105 (36.5)				
V_1	[L/kg]	1.52 (10.7)	2.79 (>100)				
A_{oral}	[mg/L]			77.2 (100)	116	16.9 (25.0)	25.3
B_{oral}	[mg/L]			7.45 (41.3)	11.2	26.9 (>100)	40.4
α	[1/h]			3.11 (59.3)	3.11	0.170 (39.0)	0.170
β	[1/h]			0.663 (21.2)	0.663	4.96 (>100)	4.96

Table B.1: *Pharmacokinetic parameters in macro constant parameterization for different anticancer drugs. Drug A2 and B was fitted with a one-compartment model* $c(t) = \frac{k_a dose}{(k_a - k)V_1}(\exp(-kt) - \exp(k_a t))$ *and drug A1 and C with a two-compartment model* $c(t) = A_{oral}\exp(-\alpha t) + B_{oral}\exp(-\beta t) - (A_{oral} + B_{oral})\exp(-kt)$. *Drug A1 180 mg/kg was predicted from 120 mg/kg and drug C 150 mg/kg was predicted from 100 mg/kg.*

Bibliography

[AGZB+91] J. M. Alvaro-Gracia, N. J. Zvaifler, C. B. Brown, K. Kaushansky, and G. S. Firestein. Cytokines in chronic inflammatory arthritis. vi. analysis of the synovial cells involved in granulocyte-macrophage colony-stimulating factor production and gene expression in rheumatoid arthritis and its regulation by il-1 and tumor necrosis factor-alpha. *The Journal of Immunology*, 146(10):3365–3371, 1991.

[Ama95] Herbert Amann. *Gewöhnliche Differentialgleichungen*. Walter de Gruyter, 1995.

[Ame01] Takeshi Amemiya. *Advanced Econometrics*, volume 9. Harvard University Press, 2001.

[And83] David H. Anderson. *Compartmental Modeling and Tracer Kinetics*. Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1983.

[Bar74] Yonathan Bard. *Nonlinear Parameter Estimation*. Academic Press New York and London, 1974.

[BD11] Peter L. Bonate and Danny R. Doward. *Pharmacokinetics in Drug Development: Advances and Applications*, volume 3. Springer, 2011.

[BFK+09] Paul G. Barash, Cullen Bruce F., Stoelting Robert K., Cahalan Michael, and Stock Christine M. *Handbook of Clinical Anesthesia*. Lippincott Williams & Wilkins, 2009.

[Boh06] E. Bohl. *Mathematik in der Biologie*. Springer-Lehrbuch. Springer, 2006.

[Bon06] P. L. Bonate. *Pharmacokinetic-Pharmacodynamic Modeling and Simulation*. Springer, 2006.

[BVT10] Lisette Bevaart, Margriet J. Vervoordeldonk, and Paul P. Tak. Evaluation of therapeutic targets in animal models of arthritis: How does it relate to rheumatoid arthritis? *Arthritis and Rheumatism*, 62(8):2192–2205, 2010.

[BZHC00] Bischof, Zafiropoulos, Hamilton, and Campbell. Exacerbation of acute inflammatory arthritis by the colony-stimulating factors csf-1 and granulocyte macrophage (gm)-csf: evidence of macrophage infiltration and local proliferation. *Clinical & Experimental Immunology*, 119(2):361–367, 2000.

[CBC+01] Andrew Cook, Emma Braine, Ian Campbell, Melissa Rich, and John Hamilton. Blockade of collagen-induced arthritis post-onset by antibody to granulocyte-macrophage colony-stimulating factor (gm-csf): requirement for gm-csf in the effector phase of disease. *Arthritis Res*, 3(5):293–298, 2001.

[CJ02] Abhijit Chakraborty and William J. Jusko. Pharmacodynamic interaction of recombinant human interleukin-10 and prednisolone using in vitro whole blood lymphocyte proliferation. *Journal of Pharmaceutical Sciences*, 91(5):1334–1342, 2002.

[CPR+98] P. B. Carrieri, V. Provitera, T. De Rosa, G. Tartaglia, F. Gorga, and O. Perrella. Profile of cerebrospinal fluid and serum cytokines in patients with relapsing-remitting multiple sclerosis. a correlation with clinical activity. *Immunopharmacology and Immunotoxicology*, 20(3):373–382, 1998.

[dEU08] Rat der Europäischen Union. *Mitteilung an die Presse - 2876. Tagung des Rates*. 2008.

[DG95] M. Davidian and D. M. Giltinan. *Nonlinear models for repeated measurement data*. Monographs on statistics and applied probability. Chapman & Hall, 1995.

[DGJ93] Natalie L. Dayneka, Varun Garg, and William J. Jusko. Comparison of four basic models of indirect pharmacodynamic responses. *Journal of Pharmacokinetics and Pharmacodynamics*, 21:457–478, 1993.

[DHG03] Joseph A. DiMasi, Ronald W. Hansen, and Henry G. Grabowski. The price of innovation: new estimates of drug development costs. *Journal of Health Economics*, 22, 2003.

[DM93] Russell Davidson and James G. Mackinnon. *Estimation and Inference in Econometrics*. Oxford University Press, 1993.

[DM09] Aristides Dokoumetzidis and Panos Macheras. Fractional kinetics in drug absorption and disposition processes. *Journal of Pharmacokinetics and Pharmacodynamics*, 36:165–178, 2009.

[DMM10] Aristides Dokoumetzidis, Richard Magin, and Panos Macheras. Fractional kinetics in multi-compartmental systems. *Journal of Pharmacokinetics and Pharmacodynamics*, 37:507–524, 2010.

[Doe76] G. Doetsch. *Einführung in Theorie und Anwendung der Laplace-Transformation*. Birkhäuser Verlag, 1976.

[Dos53] F. H. Dost. *Der Blutspiegel: Kinetik der Konzentrationsabläufe in der Kreislaufflüssigkeit*. Thieme, 1953.

[Dos68] F. H. Dost. *Grundlagen der Pharmakokinetik*. Georg Thieme Verlag, Stuttgart, 1968.

[Dri77] R. D. Driver. *Ordinary and Delay Differential Equations*, volume Applied Mathematical Science 20. Springer-Verlag, New York, Heidelberg, Berlin, 1977.

[DSW09] D. Z. D'Argenio, A. Schumitzky, and X. Wang. Adapt 5 user's guide: Pharmacokinetic/pharmacodynamic systems analysis software. *Biomedical Simulations Resource, Los Angeles*, 2009.

[EDM+08a] J. C. Earp, D. C. Dubois, D. S. Molano, N. A. Pyszczynski, R. R. Almon, and Jusko W. J. Modeling corticosteroid effects in a rat model of rheumatoid arthritis ii: mechanistic pharmacodynamic model for dexamethasone effects in lewis rats with collagen-induced arthritis. *The Journal of Pharmacology and Experimental Therapeutics*, 326(2):546–554, 2008.

[EDM+08b] J. C. Earp, D. C. Dubois, D. S. Molano, N. A. Pyszczynski, C. E. Keller, R. R. Almon, and W. J. Jukso. Modeling corticosteroid effects in a rat model of rheumatoid arthritis i: mechanistic disease progression model for the time course of collagen-induced arthritis in lewis rats. *The Journal of Pharmacology and Experimental Therapeutics*, 326(2):532–545, 2008.

[El'73] L. E. El'sgol'ts. *Introduction to the theory and application of differential equations with deviating arguments*. Academic Press, New York, 1973.

[FCBI11] N. Frances, L. Claret, R. Bruno, and A. Iliadis. Tumor growth modeling from clinical trials reveals synergistic anticancer effect of the capecitabine and docetaxel combination in metastatic breast cancer. *Cancer Chemotherapy and Pharmacology*, 68:1413–1419, 2011.

[FHM+02] Lena E. Friberg, Anja Henningsson, Hugo Maas, Laurent Nguyen, and Mats O. Karlsson. Model of chemotherapy-induced myelosuppression with parameter consistency across drugs. *Journal of Clinical Oncology*, 20(24):4713–4721, 2002.

[FM98] M. Feldmann and R. Maini. Rheumatoid arthritis: introduction. *Springer Seminars in Immunopathology*, 20:3–4, 1998.

[Fox97] J. Fox. *Applied regression analysis, linear models, and related methods*. Sage Publications, 1997.

[GEGU+10] Kosalaram Goteti, C. Edwin Garner, Lucas Utley, Jing Dai, Susan Ashwell, Demetri Moustakas, Mithat Gönen, Gary Schwartz, Steven Kern, Sonya Zabludoff, and Patrick Brassil. Preclinical pharmacokinetic/pharmacodynamic models to predict synergistic effects of co-administered anti-cancer agents. *Cancer Chemotherapy and Pharmacology*, 66:245–254, 2010.

[GI83] Zvi Griliches and Michael D. Intriligator. *Handbook of Econometrics*, volume 1. North-Holland Publishing Company, 1983.

[GM01] Jogarao V. S. Gobburu and Patrick J. Marroum. Utilisation of pharmacokinetic - pharmacodynamic modelling and simulation in regulatory decision-making. *Clinical Pharmacokinetics*, 40, 2001.

[GP82] Milo Gibaldi and Donald Perrier. *Pharmacokinetics Second Edition Revised and Expanded*. Marcel Dekker, Inc., 1982.

[Gre12] W. H. Greene. *Econometric analysis*. Prentice Hall, 2012.

[GW06] Johan Gabrielsson and Dan Weiner. *Pharmacokinetic and Pharmacodynamic Data Analysis: Concepts and Applications*. Swedish Pharmaceutical Press, 2006.

[GWJ03] B. U. Gajewska, R. E. Wiley, and M. Jordana. Gm-csf and dendritic cells in allergic airway inflammation: basic mechanisms and prospects for therapeutic intervention. *Current Drug Targets - Inflammation Allergy*, 2:279–292, 2003.

[HBBY02] Rikard Holmdahl, Robert Bockermann, Johan Baecklund, and Hisakata Yamada. The molecular pathogenesis of collagen-induced arthritis in mice - a model for rheumatoid arthritis. *Ageing Research Reviews*, 1(1):135 – 147, 2002.

[Hil04] S. A. Hill. Pharmacokinetics of drug infusions. *Continuing Education in Anaesthesia, Critical Care & Pain*, 4(3):76–80, 2004.

[HK96] Jack K. Hale and Hüseyin Kocak. *Dynamics and Bifurcations*. Springer New York, 1996.

[Jon06] Stephen W. Jones. Are rate constants constant? *The Journal of Physiology*, 571(3):502, 2006.

[KD03] H. C. Kimko and S. B. Duffull. *Simulation for Designing Clinical Trials: A Pharmacokinetic-Pharmacodynamic Modeling Perspective*. Drugs and the Pharmaceutical Sciences. Marcel Dekker, 2003.

[KOM02] William K. Kelly, Owen A. O'Connor, and Paul A. Marks. Histone deacetylase inhibitors: from target to clinical trials. *Expert Opinion on Investigational Drugs*, 11(12):1695–1713, 2002.

[KPR12] Wojciech Krzyzanski and Juan Perez Ruixo. Lifespan based indirect response models. *Journal of Pharmacokinetics and Pharmacodynamics*, 39:109–123, 2012.

[KRJ99] Wojciech Krzyzanski, Rohini Ramakrishnan, and William Jusko. Basic pharmacodynamic models for agents that alter production of natural cells. *Journal of Pharmacokinetics and Pharmacodynamics*, 27:467–489, 1999.

[Krz11] Wojciech Krzyzanski. Interpretation of transit compartments pharmacodynamic models as lifespan based indirect response models. *Journal of Pharmacokinetics and Pharmacodynamics*, 38:179–204, 2011.

[KS12] Gilbert Koch and Johannes Schropp. General relationship between transit compartments and lifespan models. *Journal of Pharmacokinetics and Pharmacodynamics*, 2012. Accepted for publication.

[KWJ06] Wojciech Krzyzanski, Sukyung Woo, and William Jusko. Pharmacodynamic models for agents that alter production of natural cells with various distributions of lifespans. *Journal of Pharmacokinetics and Pharmacodynamics*, 33:125–166, 2006.

[KWLS09] Gilbert Koch, Antje Walz, Gezim Lahu, and Johannes Schropp. Modeling of tumor growth and anticancer effects of combination therapy. *Journal of Pharmacokinetics and Pharmacodynamics*, 36:179–197, 2009.

[Kwo01] Y. Kwon. *Handbook of essential pharmacokinetics, pharmacodynamics, and drug metabolism for industrial scientists*. Kluwer Academic/Plenum Publishers, 2001.

[KWPZ+12] Gilbert Koch, Thomas Wagner, Christine Plater-Zyberk, Gezim Lahu, and Johannes Schropp. A multi-response model for rheumatoid arthritis based on delay differential equations in collagen-induced arthritic mice treated with an anti-gm-csf antibody. *Journal of Pharmacokinetics and Pharmacodynamics*, 39:55–65, 2012.

[Lai64] Anna Kane Laird. Dynamics of tumor growth. *British Journal of Cancer*, 18(3), 1964.

[LB02] Evelyn D. Lobo and Joseph P. Balthasar. Pharmacodynamic modeling of chemotherapeutic effects: Application of a transit compartment model to charaterize methorexate effects in vitro. *AAPS PharmSci*, 4:212–222, 2002.

[LC98] E.L. Lehmann and G. Casella. *Theory of point estimation*. Springer texts in statistics. Springer, 1998.

[MCM+91] Joyce Mordenti, Sharon A. Chen, Jerome A. Moore, Bobbe L. Ferraiolo, and James D. Green. Interspecies scaling of clearance and volume of distribution data for five therapeutic proteins. *Pharmaceutical Research*, 8:1351–1359, 1991.

[MI06] Panos Macheras and Athanassios Iliadis. *Modeling in Biopharmaceutics, Pharmacokinetics, and Pharmacodynamics*. Springer, Heidelberg, 2006.

[MMN+04] Lois E. McCarthy, Paolo Mannelli, Michelle Niculescu, Kevin Gingrich, Ellen M. Unterwald, and Michelle E. Ehrlich. The distribution of cocaine in mice differs by age and strain. *Neurotoxicology and Teratology*, 26(6):839 – 848, 2004.

[MSP+06] P. Magni, M. Simeoni, I. Poggesi, M. Rocchetti, and G. De Nicolao. A mathematical model to study the effects of drugs administration on tumor growth dynamics. *Mathematical Biosciences*, 200(2):127 – 151, 2006.

[Mur89] J. D. Murray. *Mathematical Biology*, volume 19 of *Biomathematics*. Springer, 1989.

[Pet84] L. R. Petzold. Automatic selection of methods for solving stiff and nonstiff systems of ordinary differential equations. *Journal of Scientific and Statistical Computing*, 15, 1984.

[PRKC+05] Juan Perez-Ruixo, Hui Kimko, Andrew Chow, Vladimir Piotrovsky, Wojciech Krzyzanski, and William Jusko. Population cell life span models for effects of drugs following indirect mechanisms of action. *Journal of Pharmacokinetics and Pharmacodynamics*, 32:767–793, 2005.

[PZJH+07] C. Plater-Zyberk, L. A. B. Joosten, M. M. A. Helsen, J. Hepp, P. A. Baeuerle, and W. B. van den Berg. Gm-csf neutralisation suppresses inflammation and protects cartilage in acute streptococcal cell wall arthritis of mice. 66(4):452–457, 2007.

[Rao87] B.L.S. P. Rao. *Asymptotic theory of statistical inference*. Wiley series in probability and mathematical statistics: Probability and mathematical statistics. Wiley, 1987.

[Rao97] B.L.S. Prakasa Rao. Weighted least squares and nonlinear regression. *Ind. Soc. Ag. Statistics*, 50:192–191, 1997.

[RBG+09] M. Rocchetti, F. Del Bene, M. Germani, F. Fiorentini, I. Poggesi, E. Pesenti, P. Magni, and G. De Nicolao. Testing additivity of anticancer agents in preclinical studies: A pk/pd modelling approach. *European Journal of Cancer*, 45(18):3336 – 3346, 2009.

[RSP+07] M. Rocchetti, M. Simeoni, E. Pesenti, G. De Nicolao, and I. Poggesi. Predicting the active doses in humans from animal studies: A novel approach in oncology. *European Journal of Cancer*, 43(12):1862 – 1868, 2007.

[SJ98] Yu-Nien Sun and William J. Jusko. Transit compartments versus gamma distribution function to model signal transduction processes in pharmacodynamics. *Journal of Pharmaceutical Sciences*, 87(6):732–737, 1998.

[SJKK07] Radojka Savic, Daniël Jonker, Thomas Kerbusch, and Mats Karlsson. Implementation of a transit compartment model for describing drug absorption in pharmacokinetic studies. *Journal of Pharmacokinetics and Pharmacodynamics*, 34:711–726, 2007.

[SL77] G. A. F. Seber and A. J. Lee. *Linear Regression Analysis*. Wiley Series in Probability and Statistics. John Wiley & Sons, 1977.

[SMC+04] Monica Simeoni, Paolo Magni, Cristiano Cammia, Giuseppe De Nicolao, Valter Croci, Enrico Pesenti, Massimiliano Germani, Italo Poggesi, and Maurizio Rocchetti. Predictive pharmacokinetic-pharmacodynamic modeling of tumor growth kinetics in xenograft models after administration of anticancer agents. *Cancer Research*, 64(3):1094–1101, 2004.

[Smi10] Hal Smith. *An Introduction to Delay Differential Equations with Applications to the Life Sciences*, volume 57 of *Texts in Applied Mathematics*. Springer, 2010.

[SPGB82] Jean-Louis Steimer, Yves Plusquellec, Anne Guillaume, and Jean-Francois Boisvieux. A time-lag model for pharmacokinetics of drugs subject to enterohepatic circulation. *Journal of Pharmaceutical Sciences*, 71(3), 1982.

[SPRVC07] Kim Stuyckens, Juan Jose Perez-Ruixo, An Vermeulen, and Eugene Cox. Modeling drug effects and resistance development on tumor growth dynamics. *PAGE*, 2007.

[SS90] Ashish Sen and Muni Srivastava. *Regression Analysis: Theory, Methods and Applications*. Springer Verlag, 1990.

[SSV+79] L. B. Sheiner, D. R. Stanski, S. Vizeh, R. D. Miller, and J. Ham. Simultaneous modeling of pharmacokinetics and pharmacodynamics: Application to d-tubocurarine. *Clin Pharamcol Ther*, 25:358–371, 1979.

[SW89] G. A. F. Seber and C. J. Wild. *Nonlinear Regression*. Wiley series in probability and mathematical statistics. John Wiley & Sons, Inc., 1989.

[Teo37a] T. Teorell. Kinetics of distribution of substances administered to the body i. the extravascular modes of administration. *Archs. int. Pharmacodyn. Ther*, 1937.

[Teo37b] T. Teorell. Kinetics of distribution of substances administered to the body ii. the intravascular modes of administration. *Archs. int. Pharmacodyn. Ther*, 1937.

[VDGG09] Piet H. Van Der Graaf and Johan Gabrielsson. Pharmacokinetic-pharmacodynamic reasoning in drug discovery and early development. *Future Medicinal Chemistry*, 1, 2009.

[VWB+91] E. G. E. De Vries, P. H. B. Willemse, B. Biesma, A. C. Stern, P. C. Limburg, and E. Vellenga. Flare-up of rheumatoid arthritis during gm-csf treatment after chemotherapy. *The Lancet*, 338(8765):517 – 518, 1991. Originally published as Volume 2, Issue 8765.

[Whe88] T. E. Wheldon. *Mathematical Models in Cancer Research*. Adam Hilger, 1988.

[Wid66] David Vernon Widder. *The Laplace Transform*. Princeton University Press, 1966.

[Wil04] Richard O. Williams. Collagen-induced arthritis as a model for rheumatoid arthritis. In Angelo Corti, Pietro Ghezzi, and John M. Walker, editors, *Tumor Necrosis Factor*, volume 98 of *Methods in Molecular Medicine*, pages 207–216. Humana Press, 2004.

[ZG11] Qingyu Zhou and James Gallo. The pharmacokinetic/pharmacodynamic pipeline: Translating anticancer drug pharmacology to the clinic. *The AAPS Journal*, 13:111–120, 2011.