Modelling Red Blood Cell Data

Julia Korell

A thesis submitted for the degree of
Doctor of Philosophy
at the University of Otago, Dunedin,
New Zealand.

November 2012

In memory of:

Rolf Korell

14. January 1949 – 9. March 1998

&

Balu

22. March 1999 – 26. November 2011

ABSTRACT

The overarching aim of this work was to utilise modelling and simulation methodology to obtain a better understanding of clinical data generated from or in red blood cells (RBCs). It focussed on RBC survival and the use of RBCs as a matrix for pharmacokinetic (PK) data.

Firstly, a novel statistical model for RBC survival was developed based on prior knowledge of the underlying physiological mechanisms using a bottom-up model building approach. The model was developed within the statistical framework of survival analysis and uses a highly flexible probability density function to describe a hypothetical RBC lifespan distribution that is able to account for plausible physiological processes of RBC destruction. These mechanisms include death due to old age (senescence), random destruction, and early or delayed failure.

The model was extended to describe *in vivo* RBC survival studies using different RBC labelling techniques and flaws inherent in the most commonly used labelling methods. Using an information theoretical approach, it was determined that full parameter estimation would be possible based on ideal labelling methods, but also based on the currently available, flawed labelling methods under an optimised study design with an intensive sampling strategy.

The model was applied to *in vivo* RBC survival data obtained in patients with chronic kidney disease (CKD) as well as healthy controls (using data obtained from the work of other investigators). RBC survival was found to be significantly reduced in CKD patients compared to controls, and the results suggest that increased random destruction is the likely cause of this reduction rather than accelerated senescence.

Secondly, a catenary compartment model describing the intracellular population PK of methotrexate (MTX) and its polyglutamated metabolites (MTXPGs) in RBCs was developed using a data driven, top-down modelling

approach. Model development was based on data obtained from 48 patients with rheumatoid arthritis (RA) receiving once weekly low-dose MTX. The developed model was used to test different hypotheses related to the mechanism of enzymatic deglutamation of MTXPGs, the loss of MTXPGs from RBCs, and the significance of genotypic and phenotypic covariates.

The final model was able to describe the time profiles of MTX and MTXPGs inside RBCs in all 48 patients, and thus can form the basis of a full pharmacokinetic-pharmacodynamic (PKPD) model for low-dose MTX treatment in RA in future work. Such a PKPD model could be used to assess whether RBC MTX or MTXPG concentrations are suitable biomarkers to monitor low-dose MTX treatment, which is currently debated in literature.

In conclusion, two different approaches were successfully applied in this thesis to develop mathematical models that are able to describe different types of RBC derived clinical data: a novel statistical RBC survival model that is able to provide a deeper insight into physiological processes of RBC destruction in the future, and a compartmental PK model describing the intracellular accumulation of MTX and MTXPGs in RBCs that can form the basis of a full PKPD model in further work.

ACKNOWLEDGMENTS

I owe my deepest gratitude to my supervisor Prof. Stephen Duffull. His encouragement, enthusiasm and dedication not only carried me through this PhD, but made and will always make me strive for reaching the stars. Thanks for being the best mentor I could have wished for and a very dear friend.

I am also deeply grateful to my co-supervisor Assoc. Prof. Lisa Stamp. Her support, guidance and valuable comments were indispensable for the methotrexate project as well as the writing of this thesis.

I would like to thank our collaborators of the Kidney in Health and Disease Research Theme, Dr Frederiek Vos, Prof. Robert Walker, Dr John Schollum, and Dr Carolyn Coulter, as well as the Arthritis Research Theme, Prof. Murray Barclay, Ms Judith Dalrymple, Dr Rebecca Roberts, Dr Mei Zhang, and Ms Jill Drake, for sharing their data and knowledge with me.

It was a great honour for me to co-supervise two students during my PhD: Oskar Alskär and Suchaya (Yim) Sanhajariya. Their work allowed me to investigate additional aspects arising from my own project, and I would like to thank them for their great efforts that made each project a success.

Special thanks go to my colleagues and friends from the Otago Pharmacometrics Group: Abhi, Chakri, Dan, Ernie, Helen, Hesh, Indri, Jill, Lee, Paul, Pavan, Shan and Vittal. I will always keep the fondest memory of our coffee sessions, conference trips and social get-togethers with all the laughter and exhilarating discussions over hot chocolate, beer and whiskey.

I wish to thank the staff at the School of Pharmacy, especially my two facilitators, Dr Clare Strachan and Dr James Green, and the IT specialists Brian Young and Tim Campbell, for their support.

I also like to acknowledge the financial support I received during my PhD candidature: a Prestigious PhD Scholarship from the University of Otago, as well as funding to attend conferences and courses from the School of Pharmacy, the Health Science Division, ASCEPT NZ, PAGANZ, and Pharsight®.

While living in New Zealand I met some wonderful friends, who made me feel home in this lovely country: the Duffull/Peterson whānau, Dan & Meg, Frederiek & Sam, Kerstin & Jan, Nirusha, Lipika, Indri, Hesh, and Rachael. These people will always be in my heart.

Finally, my deepest thanks go to my family who has always encouraged me to follow my dreams and supported me from afar during my time in New Zealand.

PUBLICATIONS

THAT AROSE FROM WORK ASSOCIATED WITH THIS THESIS

International Peer-reviewed Journal Articles

Published:

- Alskär O, **Korell J**, Duffull S (2012) *A pharmacokinetic model for the glycation of albumin*. Journal of Pharmacokinetics and Pharmacodynamics. 39(3):273-282.
- **Korell J**, Vos F, Coulter C, Schollum J, Walker R, Duffull S (2011) *Modelling red blood cell survival data*. Journal of Pharmacokinetics and Pharmacodynamics 38(6):787-801.
- **Korell J**, Coulter C, Duffull S (2011) *Evaluation of red blood cell labelling methods based on a statistical model for red blood cell survival*. Journal of Theoretical Biology 291(0):88-98.
- **Korell J**, Coulter C, Duffull S (2011) *A statistical model for red blood cell survival time*. Journal of Theoretical Biology 268(1):39-49.

In preparation:

- **Korell J**, Stamp L, Duffull S *et al.* (2012) *A population pharmacokinetic model for methotrexate in red blood cells*. (manuscript in preparation).

Poster Presentations

- **Korell J**, Vos F, Coulter C, Schollum J, Walker R, Duffull S (2011) *Modelling red blood cell survival data*. Population Analysis Group Europe (PAGE) 20, Abstract 2050, Athens –Greece.
- **Korell J**, Coulter C, Duffull S (2010) *Design of survival studies for red blood cells*. Australasian Society of Clinical and Experimental Pharmacologists & Toxicologists – New Zealand Chapter (ASCEPT NZ), Christchurch – New Zealand.

Oral Presentations

- **Korell J**, Duffull S (2012) *Red blood cell survival and its influence on clinical biomarkers*. World Conference on Pharmacometrics (WCoP), Seoul – South. ***Invited speaker***.
- **Korell J**, Stamp L, Duffull S (2012) *Pharmacokinetics of methotrexate in red blood cells*. ASCEPT NZ, Queenstown – New Zealand.
- **Korell J**, Stamp L, Duffull S (2012) *Pharmacokinetics of methotrexate in red blood cells*. Population Analysis Group Australia and New Zealand (PAGANZ) 14, Abstract 1288, Melbourne – Australia.
- Alskär O, **Korell J**, Duffull S (2012) *Development and application of a pharmacokinetic model for the glycation of albumin*. PAGANZ 14, Abstract 1306, Melbourne – Australia.
- Alskär O, **Korell J**, Duffull S (2011) *Development and application of a pharmacokinetic model for the glycation of albumin*. ASCEPT NZ, Christchurch – New Zealand.
- **Korell J**, Vos F, Sanhajariya S, Coulter C, Schollum J, Walker R, Duffull S (2011) *Modelling red blood cell data*. PAGANZ 13, Abstract 1155, Auckland – New Zealand.
- **Korell J**, Coulter C, Duffull S (2010) *Design of survival studies for red blood cells*. PAGE 19, Abstract 1701, Berlin – Germany.
- **Korell J**, Coulter C, and Duffull S (2010) *Design of survival studies for red blood cells*. PAGANZ 12, Abstract 1115, Adelaide – Australia.
- **Korell J**, Coulter C, Duffull S (2009) *A statistical model for red blood cell survival time*. ASCEPT NZ, Dunedin – New Zealand.

TABLE OF CONTENTS

LIST OF FIGURES

MAIN TEXT

APPENDICES

LIST OF TABLES

MAIN TEXT

APPENDICES

LIST OF EQUATIONS

MAIN TEXT

APPENDICES

GLOSSARY

ABBREVIATIONS

^{14}C	radioactive carbon-14
^{15}N	heavy nitrogen-15 (stable isotope)
^{51}Cr	radioactive chromium-51
^{59}Fr	radioactive iron-59
-2LL	-2x log likelihood
A	adenine
ABC	ATP-binding cassette
ADME	absorption, distribution, metabolism & excretion
AIC	Akaike's Information Criterion
AIP	approximation to the preposterior information
ATIC	5-aminoimidazole-4-carboxamide ribonucelotide transformylase
ATP	adenosine triphosphate
AUC	area under the curve
BEG principle	Box – Einstein – Gelman principle of relaxed parsimony for model selection
BLQ	below limit of quantification
BMI	body mass index
BSA	body surface area
BSV	between subject variability / variance
C	carbon; cytosine
C_1	"one-carbon group", i.e. a carbon residue containing a single carbon
cdf	cumulative distribution function
CI	confidence interval
CKD	chronic kidney disease
CL	clearance; cohort labelling
C_{max}	peak / maximum plasma concentration
CNV	copy number variation
CO	carbon monoxide
CO_2	carbon dioxide
CV	coefficient of variation
DE	determinant of expectation
$DF^{32}P$	radioactive diisopropylfluorophosphate-32
DHFR	dihydrofolate reductase
DMARD	disease modifying anti-rheumatic drug
DNA	deoxyribonucleic acid
DV	dependent variable
EBE	empirical Bayes estimate
ED	expectation of the determinant
eGFR	estimate glomerular filtration rate
ELS	extended least squares
EM	expectation maximisation
EPO	erythropoietin
E-step	expectation step
F	females

FO	first-order estimation
FOCE	first-order conditional estimation
FOCE INT	first-order conditional estimation with interaction
FPGS	folylpolyglutamate synthetase
FW	flexible Weibull distribution
γGH	γ-glutamyl hydrolase
G	guanine
GFR	glomerular filtration rate
GI	gastro intestine; gastrointestinal
Hb	haemoglobin
[Hb]	haemoglobin concentration
HbA_{1c}	glycated haemoglobin A_{1c}
HClnD	hypercube log D-optimal design criterion
HD	haemodialysis
HT	height
i.v.	intravenous
ICSH	International Committee of Standardization in Hematology
IPP method	individual PK parameters = method for PKPD modelling using empirical Bayes estimates
IPRED	individual prediction
IRWLS	iteratively reweighted least squares
K_m	Michaelis constant in enzyme kinetics (substrate concentration at half-maximum reaction rate)
L2	level 2 data item in NONMEM®
LBW	lean body weight
LIDR models	lifespan indirect response models
LL	log likelihood
LOQ	limit of quantification
LRT	likelihood ratio test
LS	lifespan
M	males
M3	Stuart Beal's method 3 for handling BLQ data
M6	Stuart Beal's method 6 for handling BLQ data
MCMC	Markov Chain Monte Carlo
MCV	Mean corpuscular volume
MDRD formula	Modification in Diet in Renal Disease formula
MDRT	multi drug resistance transporter
MMD	maximin D-optimal design
M-step	maximisation step
MTHFR	methylenetetrahydrofolate reductase
MTX	methotrexate (drug in general)
MTXGlu1	parent molecule of MTX containing 1 glutamate moiety
MTXGluX	specific methotrexate (poly)glutamate having X moieties of glutamate attached, e.g. MTXGlu1 … MTXGlu5
MTXPGs	polyglutamated MTX metabolites (in general)
$Na_2{}^{51}CrO_4$	radioactive sodium chromate-51
$NADPH/NADP^+$	nicotinamide adenine dinucleotide phosphate
NLME models	non-linear mixed effects models

ns	statistically non-significant
O_2	oxygen
OLS	ordinary least squares
OFV	objective function value
PBMC	peripheral blood mononuclear cells
PBPK	physiologically based pharmacokinetics
PD	pharmacodynamic(s); peritoneal dialysis
pdf	probability density function
PK	pharmacokinetic(s)
PKPD	pharmacokinetic(s)-pharmacodynamic(s)
PM	parent-metabolite
pvcVPC	prediction and variability corrected visual predictive check
R^2	coefficient of determination
RA	rheumatoid arthritis
RAW	reduced additive Weibull distribution
RBC	red blood cell
RES	reticuloendothelial system
RFC	reduced folate carrier
rHuEPO	human recombinant erythropoietin
RL	random labelling
RUV	residual unexplained variability
SAEM	stochastic approximation expectation minimisation
sc	subcutaneous
SD	standard deviation
SE	standard error
SEX	indicator variable for sex: 1 = male, 0 = female
SHMT	serine hydroxymethyltransferase
SNP	single nucleotide polymorphism
T	thymine
TC	transit compartment
TDM	therapeutic drug monitoring
TNF-α	tumor necrosis factor α
TYMS	thymidylate synthase
V	apparent volume of distribution
VPC	visual predictive check
WT	total body weight

VARIABLES & SYMBOLS

a	indicator variable for a genotypic covariate in the MTX RBC PK model , i.e. a = (0, 1)
A	amount of drug in a compartment
A_{Hb}	total amount of haemoglobin in the body
b	indicator variable for a genotypic covariate in the MTX RBC PK model , i.e. b = (0, 1)
b	day of birth of the b^{th} cohort of RBCs in the RBC survival model
c	constant hazard rate in the RBC lifespan/survival model
C	concentration
C_e	concentration in the effect compartment
CL	clearance
CL_1	clearance of MTX in the MTX plasma PK model
CL_{GluX}	clearance of MTXGluX from RBCs in the MTX RBC PK model
cov	indicator variable for a covariate in the RBC survival model, i.e. cov = (0, 1)
COV	value of a continuous covariate
CV_{prop}	coefficient of variation of a proportional error
CV^2_{prop}	squared coefficient of variation of a proportional error
D	dose
E	effect
EC_{50}	drug concentration resulting in half-maximum effect
Eff	D-efficiency
E_{max}	maximum effect
E_0	baseline effect
f	fraction of renally cleared MTX
F	bioavailability in the MTX PK model
F	bioavailability
$F(t)$	cumulative distribution function
$f(\cdot)$	mathematical function; structural model; probability density function
$f_{FW}(t)$	flexible Weibull distribution (pdf)
$f_{MIX}(t)$	combined pdf of the RBC lifespan model
$f_{RAW}(t)$	reduced additive Weibull distribution (pdf)
$g(\cdot)$	additional function required for rejection sampling
$h(t)$	hazard function
$H(t)$	cumulative hazard function
i	index for individual (human or RBC), i.e. i = (1, … N)
I	intermediate in a turnover PKPD model
I	identity matrix
j	index for observation, i.e. j = (1, … n_i) for the i^{th} individual
J	Jacobian matrix

k	index for iteration
k_a	first-order absorption rate constant
$\mathsf{k_a}$	first-order absorption rate constant in the MTX plasma PK model
k_{el}	first-order elimination rate constant
$\mathsf{k_{el}}$	first-order elimination rate constant in the MTX plasma PK model
k_{eq}	first-order equilibrium rate constant in an effect compartment PKPD model
$\mathsf{k_{FPGSX}}$	first-order rate constant of polyglutamation converting MTXGluX to MTXGlu(X+1) in the MTX RBC PK model
k_{in}	first-order rate constant of input in a TC model
$\mathsf{k_{in}}$	first-order rate constant of MTX uptake into RBCs in the MTX RBC PK model
k_{out}	first-order elimination rate
$\mathsf{k_{out,GluX}}$	first-order rate constant of MTXGluX loss from RBCs in the MTX RBC PK model
k_p	production rate constant of RBCs in the RBC lifespan/survival model
k_p^*	production rate of RBCs in the RBC survival model corrected for reuse of a cohort label
k_{RD}	first-order rate constant of random destruction in a TC model
k_{TR}	first-order rate constant of transit in a TC model
k_v	vesiculation rate constant in the RBC survival model
k_{xy}	first-order intercompartmental transfer rate constant from compartment x to compartment y
$\mathsf{k_{xy}}$	first-order intercompartmental transfer rate constant from compartment x to compartment y in the MTX plasma PK model
$\mathsf{k_{\gamma GHX}}$	first-order rate constant of deglutamation converting MTXGluX to MTXGlu(X-1) in the MTX RBC PK model
$\mathsf{k_{\gamma GHXY}}$	first-order rate constant of deglutamation converting MTXGluX to MTXGluY in the MTX RBC PK model
L	likelihood
$\overline{L}$	calculated mean lifespan
log	natural logarithm
$loss_b$	loss of label between day b-1 and b in the RBC survival model
LS	mean RBC lifespan in a TC model
LS_{app}	apparent mean RBC lifespan
LS_i	lifespan of the i^{th} RBC in the RBC lifespan model
m	mixing parameter in the RBC lifespan/survival model
$\mathbf{M_F}$	Fisher Information matrix
N	total number of individuals (humans or RBCs)
NC	number of transit compartments
$N_{CL}(t)$	total number of labelled RBCs at time t after loss due to vesiculation and including reuse in the RBC survival model for a cohort labelling method
$N_d(t)$	total number of surviving RBCs at time t after loss due to decay in the RBC survival model

$N_e(t)$	total number of surviving RBCs at time t after loss due to elution in the RBC survival model
n	total number of observations
n_i	total number of observations in the i^{th} individual
n_p	total number of parameters
nrCL_1	non-renal clearance of MTX in the MTX plasma PK model
$N_{RBCs}(t)$	number of living RBCs at time t in the RBC lifespan/survival model
$N_{RL}(t)$	total number of labelled RBCs at time t after loss due to vesiculation, decay and elution in the RBC survival model for a random labelling method
NSI	normalised sensitivity index
$N_v(t)$	total number of surviving RBCs at time t after loss due to vesiculation in the RBC survival model
N_0	total number of labelled RBCs at time 0
p	index for parameter, i.e. $p = (1, \ldots n_p)$
P	probability
Q	conditional expectation of the log likelihood
r	ratio $f(x)/g(x)$ calculated for rejection sampling
r_1 r_2	} reduced lifespan parameters in the RBC lifespan/survival model
rCL_1	renal clearance of MTX in the MTX plasma PK model
$reuse_b$	number of reused label on day b in the RBC survival model
rf	fraction of lost label that is reused on the next day in the RBC survival model
R_{in}	rate of production / input
R_{out}	rate of elimination / loss
$\mathbb{R}^+$	set of all positive real numbers
s	standard deviation
s_1 s_2	} senescence parameters in the RBC lifespan/survival model
S	sensitivity
$S(t)$	survival function
t	time
t_{Bi}	day of "birth" of the i^{th} RBC in the RBC lifespan model
t_{Bmax}	time until "birth" of RBCs occurs in the RBC lifespan model
tlag	lag time in the MTX PK model
t_L	time point of labelling in the RBC survival model
$t_{1/2app}$	apparent half-life
$t_{1/2Cr}$	half-life of ^{51}Cr = 27.7025 days
$t_{1/2el}$	half-life of elution of ^{51}Cr = 70 days
T	transpose
T	survival time, time of event
$\mathbf{T}$	time space

v	control factor for rate of vesiculation with v = (1, 2)
V	volume of distribution
$\mathbf{V}$	population covariance matrix
V_{blood}	total blood volume
V_{GluX}	apparent volume of distribution of MTXGluX inside RBCs in the MTX RBC PK model
V_{RBCs}	total volume of RBCs
V_1	apparent volume of distribution of the central compartment in the MTX plasma PK model
w	weight
x	independent variable
$\mathbf{X}$	n-by-n_p matrix of independent variables
y	observation / dependent variable
$\mathbf{Y}$	n-by-1 vector of dependent variables
z	variable of integration over time in LIDR models
Z	indicator variable for living & non-living RBCs in the RBC lifespan model
α	weighting parameter in the HClnD criterion
β	covariate coefficient
δ	level of change in the parameter estimate for a sensitivity analysis
Δ	difference
ε	residual unexplained variability
ϕ	any parameter in a model (fixed or random effect)
$\bar{\phi}$	mean of any parameter obtained in bootstrap analysis (fixed or random effect)
$\hat{\phi}$	estimate of any parameter in a model (fixed or random effect)
$\boldsymbol{\phi}$	vector of all parameters (fixed or random effects)
$\boldsymbol{\Phi}$	parameter space
γ	Hill coefficient
η	random effect (scalar)
$\boldsymbol{\eta}$	random effects (vector or matrix)
λ	constant hazard rate
λ_1	first exponent in a sum of exponentials
λ_2	second exponent in a sum of exponentials
π	RBC lifespan distribution
θ	fixed effect parameter (scalar)
$\bar{\theta}$	population mean fixed effect parameter (scalar)
$\hat{\theta}$	fixed effect parameter estimate (scalar)
$\hat{\bar{\theta}}$	population mean fixed effect parameter estimate (scalar)
$\boldsymbol{\theta}$	n_p-by-1 vector of fixed effect parameter values
$\bar{\boldsymbol{\theta}}$	n_p-by-1 vector of (calculated) population mean fixed effect parameters
$\hat{\boldsymbol{\theta}}$	n_p-by-1 vector of fixed effect parameter estimates
$\hat{\bar{\boldsymbol{\theta}}}$	n_p-by-1 vector of population mean estimates for fixed effect parameters

$\boldsymbol{\Theta}$	parameter space of fixed effects
σ^2_{add}	variance of an additive error
σ_{add}	standard deviation of an additive error
$\boldsymbol{\Sigma}$	error matrix
τ	variable of integration over time in the RBC survival model
ϑ	step size in the SAEM algorithm
ϖ^2	element of the variance-covariance matrix (scalar)
Ω	between subject variance (scalar)
$\boldsymbol{\Omega}$	variance-covariance matrix (dimensions n_p-by-n_p)
Ψ_D	D-optimal design criterion
Ψ_D^*	local D-optimal design

AIM & STRUCTURE OF THE THESIS

The overarching aim of this work was to utilise modelling and simulation methodology to obtain a better understanding of clinical data generated from or in red blood cells (RBCs) with respect to:

- RBC survival and physiological mechanisms of RBC destruction.
- RBCs as matrix for biomarker data with an application to the pharmacokinetics of low-dose methotrexate (MTX) treatment in rheumatoid arthritis.

For this purpose, the thesis has been divided into five parts (Table P.1). First, an introduction to the field of modelling and simulation and the methods used in this thesis is given in Part I. Background information on RBC survival and physiology, as well as an overview of the pharmacokinetics and pharmacodynamics of MTX relevant to this thesis are also provided in this part.

Second, the development, assessment and application of a novel statistical model for RBC survival based on plausible physiological mechanisms of RBC destruction in the human body is described in Part II which comprises three chapters. A so-called bottom-up approach was used here, where model development is based on an understanding of the underlying physiological mechanisms and not dependent on data.

Part III consists of two chapters and covers modelling of clinical data to describe the pharmacokinetics of low-dose MTX based on drug concentrations measured in RBCs. In this part, model development is primarily data driven, also known as a top-down approach. The developed model is then used to test different hypotheses related to the underlying mechanisms of intracellular MTX accumulation as well as to assess the significance of genotypic and phenotypic covariates.

Part VI integrates both approaches by concluding this thesis with a discussion of the findings and future prospects.

Lastly, the appendices in Part V contain additional material including MATLAB® and NONMEM® codes related to the individual chapters as well as a list of all references.

Table P.1: *Structure of this thesis.*

Part I	**Introduction**
	Chapter 1 – Introduction
Part II	**Red blood cell survival model**
	Chapter 2 – A statistical model for red blood cell survival time
	Chapter 3 – Evaluation of red blood cell labelling methods
	Chapter 4 – Modelling red blood cell survival data
Part III	**Pharmacokinetics of methotrexate in red blood cells**
	Chapter 5 – A population pharmacokinetic model for methotrexate measured in red blood cells
	Chapter 6 – Hypotheses testing for methotrexate pharmacokinetics in red blood cells
Part IV	**Discussion & Future Prospects**
	Chapter 7 – Discussion & Future Prospects
Part V	**Appendices**
	Appendix 1 – Appendix to Chapter 1
	Appendix 2 – Appendix to Chapter 2
	Appendix 3 – Appendix to Chapter 3
	Appendix 4 – Appendix to Chapter 4
	Appendix 5 – Appendix to Chapter 5
	Appendix 6 – Appendix to Chapter 6
	References

AUTHOR DECLARATION

This thesis is a direct continuation of preliminary work which I did prior to starting the PhD candidature. This preliminary work counted towards the qualification Postgraduate Certificate in Pharmacy (PGCertPharm) at the University of Otago in 2009.

The basic RBC lifespan model developed during this preliminary work is described in the Introduction to this thesis (Chapter 1, Section 1.2.4.3). However, this work required significant further improvement, especially with respect to the underlying numerical code, and general expansion. This additional work was conducted during my PhD and builds the foundation of Chapter 2 and the corresponding publication, which was written as part of my PhD.

All other parts of this thesis are based on work done entirely during my PhD candidature.

PART I

INTRODUCTION

"*If it were not for the great variability among individuals*

medicine might as well be a science and not an art."

Sir William Osler

Chapter 1: Introduction

This chapter is partially based on the following peer-reviewed publications:

Korell J, Coulter C, Duffull S (2011) *A statistical model for red blood cell survival*. Journal of Theoretical Biology 268(1):39-49.

Korell J, Coulter C, Duffull S (2011) *Evaluation of red blood cell labelling methods based on a statistical model for red blood cell survival*. Journal of Theoretical Biology 291(0):88-98.

1.1. Model based analysis of clinical data

This introduction provides an overview of model based analysis of clinical data, with emphasis on the methodology applied in modelling and simulation in the area of population pharmacokinetic-pharmacodynamic (PKPD) modelling. For general references to this section refer to [1-4]; specific references will be indicated as appropriate.

1.1.1. Models & Model development

1.1.1.1. Models

In this thesis, the term "model" refers to a mathematical description of a system given by a function *f*, that relates the input (a *n*-by-n_p matrix of independent variables **X**) with an outcome (a *n*-by-1 vector of dependent variables **Y**) via the parameters $\boldsymbol{\theta}$ (n_p-by-1 vector of estimable regression parameters) and accounts for a residual error $\boldsymbol{\varepsilon}$ (vector of the same size as **Y**). Note, that in this thesis matrices and vectors will be denoted by bold symbols.

$$\mathbf{Y} = f(\boldsymbol{\theta}, \mathbf{X}) + \boldsymbol{\varepsilon}$$

Equation 1.1: *General form of a mathematical model.*

Models are simplified descriptions of reality. They can be used to describe data arising from a system, to learn about the system itself, to generate and test hypotheses and to predict future outcomes based on what-if scenarios. However, a single model does not necessarily have to fulfil all of these criteria, and it is crucial to take into account the purpose of the model, i.e. its intended application, when choosing between competing models during the model development process.

1.1.1.2. Model development

The purpose of model development is to find the function *f* that describes the relationship between $\mathbf{X}$ and $\mathbf{Y}$ and to determine the corresponding parameters $\boldsymbol{\theta}$ that describe this relationship quantitatively.

Models can be developed in two different ways:

- using a *top-down approach*, or
- using a *bottom-up approach*.

1.1.1.2.1. Top-down approach

A top-down approach starts with data. Subsequently, an empirical model is developed that best describes the data. Usually, this is done with emphasis on the principle of parsimony, meaning the simplest model describing the data sufficiently well is preferred. Common empirical models are based on a sum of exponentials or on polynomials, using the lowest possible number of exponentials or polynomials. However, empirical models do not directly relate to the underlying mechanisms that gave rise to the data and therefore do not allow for making inferences on these mechanisms.

Knowledge of the underlying system can however help making decisions during a top-down model building process. Incorporating such prior knowledge leads to semi-empirical models that allow for some mechanistic interpretation. Typical examples of semi-empirical models are compartmental pharmacokinetic (PK) models. These models do not relate in detail to all physiological mechanisms involved in absorption, distribution and elimination of a drug in the body, but provide a parsimonious, simplified mathematical description of the time course of the drug concentrations in the body while still retaining some mechanistic aspects. Compartmental PK models themselves will be introduced in more detail later on (Section 1.1.2.1.2).

1.1.1.2.2. Bottom-up approach

A bottom-up approach on the other hand uses prior knowledge about a system to develop a model. It does not require data *a priori* and is not necessarily bound to the principle of parsimony. The resulting models are often called "system models" or "mechanistic models", and are usually more complex than empirical models.

However, prior knowledge of all underlying mechanisms might not be available during model development, or the complexity of a full mechanistic model might escalate beyond reason with respect to the intended purpose of the model. Under these circumstances, semi-mechanistic models are often developed which also include empirical aspects. These models are simpler than fully mechanistic models, but incorporate the key mechanisms that are of interest in the modelling analysis and therefore allow making direct inference about these mechanisms from the model.

Ultimately, mechanistic as well as semi-mechanistic models need to be tested, i.e. applied to data, to evaluate their credibility. This data needs to be informative for the processes described in the model. Such information rich data can be difficult to obtain clinically or simply be unavailable.

In this thesis, both types of model building approaches have been applied. In the first chapter of Part II (Chapter 2), a semi-mechanistic statistical model for red blood cell (RBC) survival was developed using a bottom-up approach. Prior knowledge of RBC destruction in the human body was used to develop a statistical survival model independent from data. The theoretical aspects of the model were explored in Chapter 3 and the model was finally tested for its ability to describe clinical data in Chapter 4. Part III describes the top-down development of a semi-empirical population PK model for methotrexate (MTX) and its metabolites based on data measured inside RBCs and using a classical compartmental analysis.

1.1.1.2.3. BEG principle

Selecting an appropriate model is essential in the model developing process to ensure that the final model can fulfil its intended purpose. However, it needs to be noted that there will never be a "true" model as George E. P. Box has famously noted [5]:

"Essentially, all models are wrong, but some are useful."

A useful guidance for a more relaxed parsimonious selection of empirical models is the *BEG principle*, which comprises three quotes:

1. Box's quote mentioned above
2. Einstein's modification of Ockham's razor on parsimony [year unknown]:

"A scientific theory should be as simple as possible, but not simpler."

3. Gelman's call for appropriateness [6]:

"Do the model's deficiencies have a noticeable effect on the substantive inferences?"

In other words, a model should be as simple as possible while still being able to fulfil its purpose without leading to biased inferences and conclusions.

1.1.2. Pharmacokinetics and Pharmacodynamics

Here, a brief introduction to the basic concepts of pharmacokinetics (PK) and pharmacodynamics (PD) as well as common models used to describe these will be given.

The description of PK data analysis methods will focus on the compartment model approach. This approach is applied in the third part of this thesis for the development of the population PK model for RBC MTX concentrations.

PD data analysis is not part of this thesis, but the ultimate goal for future work is to extend the RBC MTX PK model towards a full PKPD model. Thus, a brief overview of PD and PKPD is given in Section 1.1.2.2 and Section 1.1.2.3, respectively, for the sake of completeness.

1.1.2.1. Pharmacokinetics

In short: "*Pharmacokinetics is what the body does to the drug*" [7]. More appropriately, PK describes the relationship between the dose of a drug and its concentration in the body as a function of time.

1.1.2.1.1. Time course of drug concentration

The time course of drug concentrations in the body after administration of a drug is governed by four processes, known as the ADME scheme:

- *Absorption*: uptake of the drug into the systemic circulation after administration, passive via diffusion across membranes or active via transport mechanisms.
- *Distribution*: dispersion of the drug in the body tissues after absorption.
- *Metabolism*: transformation of the drug via chemical reactions.
- *Excretion*: removal of the drug (and its metabolites) from the body.

Metabolism and excretion together result in *elimination* of the drug.

For a PK analysis, drug concentrations in the body can be measured in different matrices, but most commonly (due to convenience) concentrations in plasma are obtained.

In mathematical terms, PK then describes the plasma concentration C of a drug as function of the administered dose over time t dependent on the parameters $\boldsymbol{\theta}^{\{\mathbf{PK}\}}$:

$$C(t) = f_{PK}\left(dose, t, \boldsymbol{\theta}^{\{\mathbf{PK}\}}\right)$$

Equation 1.2: *Basic equation for pharmacokinetic models.*

A PK modelling analysis aims to find the function f and the corresponding estimates of the PK parameters $\boldsymbol{\theta}^{\{\mathbf{PK}\}}$ that describe the relationship between the dose of a particular drug and its plasma concentration over time. Common PK parameters of interest are clearance (CL), the apparent volume of distribution (V), and the elimination rate constant k_{el}, which is equal

to CL/V. In addition, the absorption rate constant (k_a) and the oral bioavailability (F) are usually of interest for drugs that can be administered orally.

A standard approach is to use compartment models for this type of modelling analysis. These models will be described in the following section using several examples. In these examples it is assumed that the drug follows linear PK, which means that Equation 1.2 is linear with respect to dose. Although this is true for the majority of drugs, non-linear PK can occur as well, e.g. if the absorption, distribution, elimination and/or metabolism of the drug involves active and therefore saturable transport mechanisms or enzymatic processes, or if the drug itself induces or inhibits its metabolism and/or elimination in a concentration dependent manner.

1.1.2.1.2. Compartment models for PK analysis

Compartment PK models are semi-empirical models where the body is described by a finite number of disposition compartments, commonly not more than three. These compartments do not necessarily reflect true body tissues although associations can be made. The simplest model is a one-compartment model, consisting of a central compartment, e.g. the plasma and other rapidly equilibrating tissues, from which samples are taken (Figure 1.1). Note that only the disposition compartments are shown in the following schemata, while absorption compartments are omitted for simplification purposes.

For a one-compartment model it is assumed that the drug distributes in all body tissues proportionally to its distribution in the central compartment. The resulting PK profile for a drug with first-order elimination after an intravenous (i.v.) bolus dose is monoexponential (Equation 1.3, Figure 1.2 blue line).

The PK profile after oral administration can be described by Equation 1.4 and is shown in Figure 1.2 as red line.

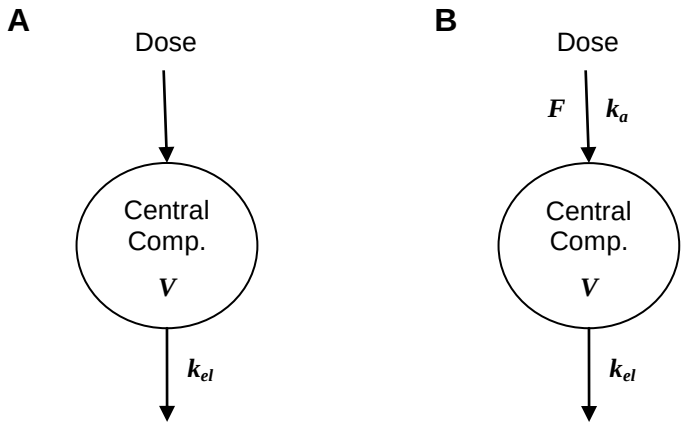


Figure 1.1: *Schematic of a one-compartment PK model with first-order elimination for an i.v. bolus dose* (**A**)*, and for oral administration with first-order absorption* (**B**)*. Parameters: k_{el} = elimination rate constant, V = apparent volume of distribution in the central compartment, k_a = absorption rate constant, F = oral bioavailability.*

$$C(t) = \frac{D}{V} \cdot \exp(-k_{el} \cdot t) = \frac{D}{V} \cdot \exp\left(-\frac{CL}{V} \cdot t\right)$$

Equation 1.3: *One-compartment PK model for i.v. bolus administration with first-order elimination. C(t) = plasma concentration over time, D = dose, V = apparent volume of distribution, k_{el} = CL/V = elimination rate constant, CL = clearance.*

$$C(t) = \frac{D \cdot F \cdot k_a}{V \cdot (k_a - k_{el})} \cdot (\exp(-k_{el} \cdot t) - \exp(-k_a \cdot t))$$

Equation 1.4: *One-compartment PK model for oral administration with first-order absorption and elimination. C(t) = plasma concentration over time, D = dose, F = bioavailability, V = apparent volume of distribution, k_{el} = CL/V = elimination rate constant, CL = clearance.*

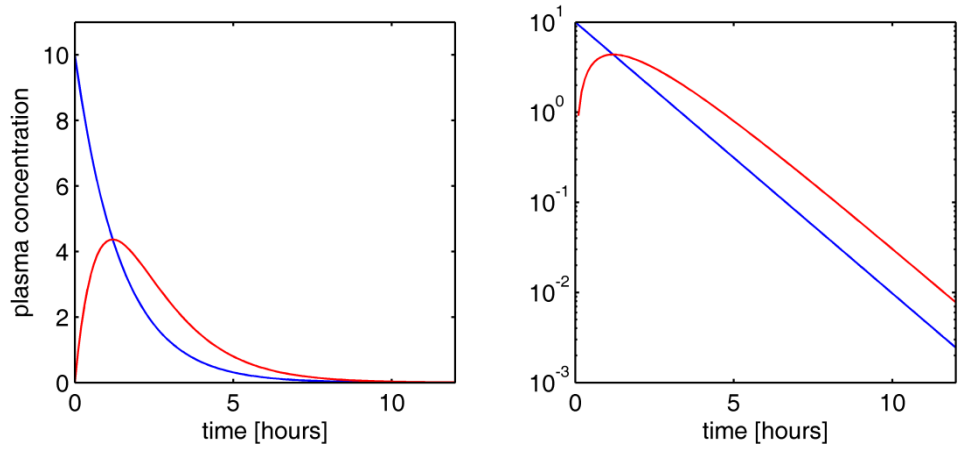


Figure 1.2: *Plasma concentrations for a one-compartment PK model with first-order elimination for an i.v. bolus dose (blue line) and for oral administration with first-order absorption (red line) with D = 10 units, CL = ln(2) hr^{-1}, V = 1 L, F = 1 and k_a = 1 hr^{-1}.*

Two-compartment models (Figure 1.3) have an additional peripheral compartment into which the drug distributes at a different rate than in the plasma, resulting in a biexponential PK profile after i.v. bolus application (Equation 1.5, Equation 1.6, Figure 1.4).

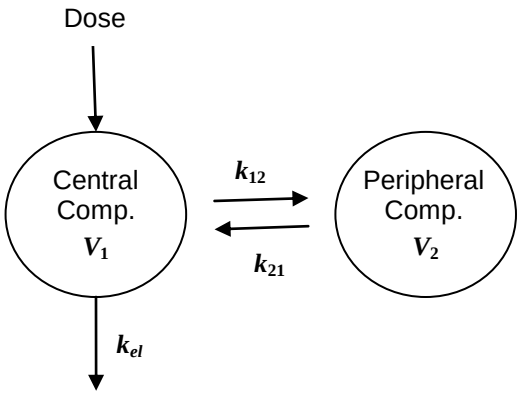


Figure 1.3: *Schematic of a two-compartment PK model with first-order elimination for an i.v. bolus dose. Parameters: k_{el} = elimination rate constant, k_{12}, k_{21} = first-order transfer rate constant from the central to the peripheral compartment and vice versa. V_1 = apparent volume of distribution in the central compartment, V_2 = apparent volume of distribution in the peripheral compartment.*

$$\frac{dA_1}{dt} = -k_{12} \cdot A_1 + k_{21} \cdot A_2 - k_{el} \cdot A_1$$
$$\frac{dA_2}{dt} = k_{12} \cdot A_1 - k_{21} \cdot A_2$$

$$A_1(0) = dose \qquad A_2(0) = 0$$

Equation 1.5: *Two-compartment PK model for i.v. bolus administration with first-order elimination expressed as ordinary differential equations with initial conditions.* A_1 *= amount in central compartment,* A_2 *= amount in peripheral compartment.*

$$C(t) = C_1 \cdot exp(-\lambda_1 \cdot t) + C_2 \cdot exp(-\lambda_2 \cdot t) \qquad C(0) = C_1 + C_2$$

$$k_{12} = \lambda_1 + \lambda_2 - k_{21} - k_{el} \qquad k_{21} = \frac{C_1 \cdot \lambda_2 + C_2 \cdot \lambda_1}{C_1 + C_2} \qquad k_{el} = \frac{\lambda_1 \cdot \lambda_2}{k_{21}}$$

Equation 1.6: *Closed form solution of a two-compartment PK model for i.v. bolus administration with first-order elimination.*

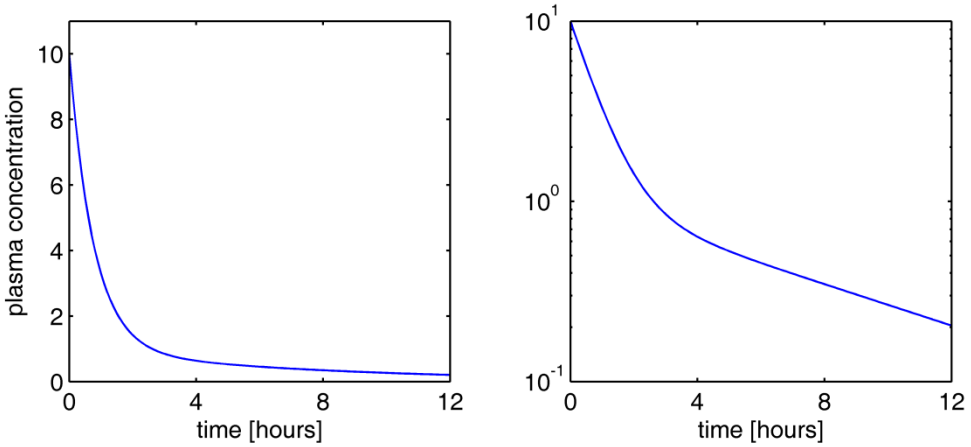


Figure 1.4: *Plasma concentrations for a two-compartment PK model with first-order elimination for an i.v. bolus dose D = 10 units,* k_{el} *= 1* hr^{-1}*,* V_1 *= 1 L,* k_{12} *= 0.5* hr^{-1}*,* k_{21} *= 0.25* hr^{-1}*, and* V_2 *= 5 L.*

For multi-compartment models two different types of models need to be distinguished: *mammillary models* and *catenary models*. In a mammillary model all peripheral compartments are directly linked to the central compartment, but not interlinked. In a catenary model the compartments are linked in sequence (Figure 1.5). In mathematical terms, these models are usually expressed as series of ordinary differential equations equivalent to Equation 1.5.

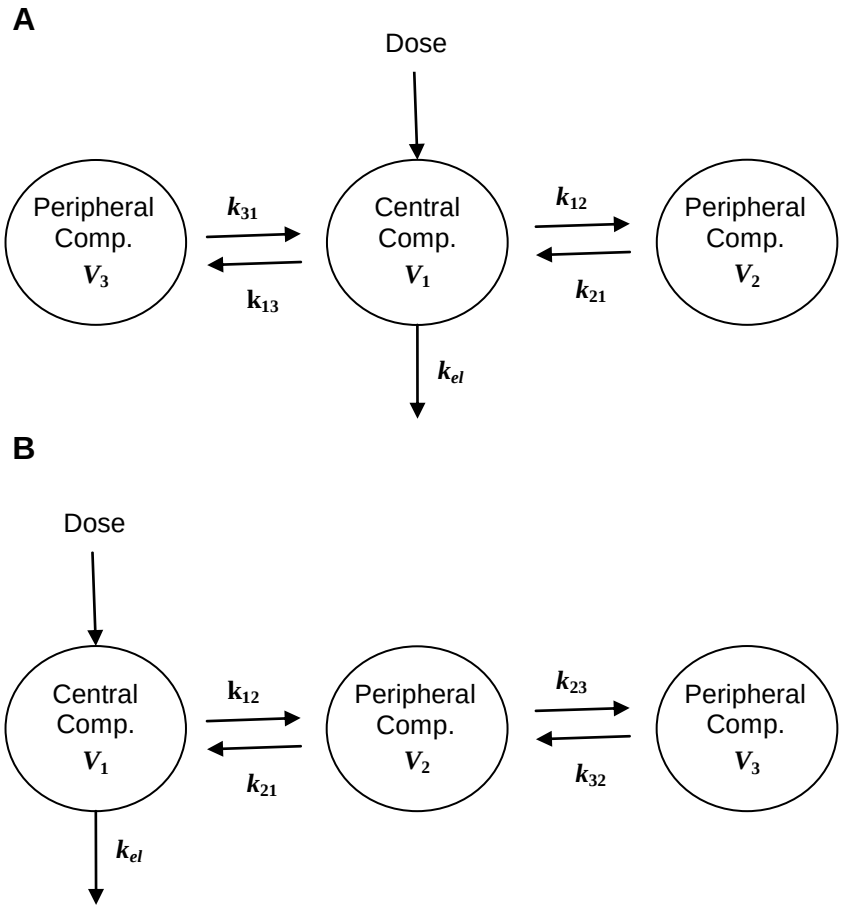


Figure 1.5: *Schematic of a mammillary* (**A**) *and a catenary* (**B**) *three-compartment model with first-order elimination for an i.v. bolus dose. Parameters:* k_{el} *= elimination rate constant,* k_{12}, k_{21}, k_{13}, k_{31}, k_{23}, k_{32}, *= first-order transfer rate constants between the compartments,* V_1 *= apparent volume of distribution in the central compartment,* V_2 *= apparent volume of distribution in the first peripheral compartment,* V_3 *= apparent volume of distribution in the second peripheral compartment.*

Mammillary models are prevailing in PK analysis, as it is commonly assumed that all tissue compartments are directly linked with the central plasma compartment rather than with each other. On the other hand, catenary compartment models are useful to describe transition processes, e.g. the successive formation of multiple metabolites [8] or a time delay in drug absorption after oral administration [9,10].

1.1.2.2. Pharmacodynamics

In short: "*Pharmacodynamics is what the drug does to the body*" [7]. More appropriately, PD describes the relationship between drug concentrations and effect.

1.1.2.2.1. Concentration – effect relationship

After absorption into the systemic circulation and distribution to its site of action, a drug causes a physiological effect by binding to a receptor. Here, "effect" can stand for the desired clinical effect of the drug but also undesired adverse effects.

In mathematical terms, PD describes the effect E of a drug as a function of its plasma concentration C and the PD parameters $\boldsymbol{\theta}^{\{\mathbf{PD}\}}$:

$$E(C) = f_{PD}\left(C, \boldsymbol{\theta}^{\{\mathbf{PD}\}}\right)$$

Equation 1.7: *Basic equation for pharmacodynamic models.*

1.1.2.2.2. PD models

The magnitude of a pharmacodynamic effect E depends on the concentration of the drug. However, the relationship between concentration and effect is generally non-linear, i.e. doubling the concentration does not result in a two fold increase in effect. Also, the effect asymptotes to a maximum effect (E_{max}) with increasing concentrations, resulting in a reduced increase in effect with the increase in concentration ("law of diminishing returns").

Based on the receptor binding theory for drug action, a mathematical model was developed to describe the relationship between drug concentration and effect, which is the same as the Michaelis-Menten equation that is used to describe saturable enzyme kinetics. This basic PD model is commonly referred to as *E_{max} model* (Equation 1.8).

$$E(C) = \frac{E_{max} \cdot C}{EC_{50} + C}$$

Equation 1.8: *E_{max} model with E_{max} = maximum effect of the drug, C = drug concentration, and EC_{50} = drug concentration resulting in half-maximum effect.*

An extension to the basic E_{max} model is the so-called *sigmoidal E_{max} model* which introduces the Hill coefficient γ to describe s-shaped asymptotic behaviour in the model:

$$E(C) = \frac{E_{max} \cdot C^{\gamma}}{EC_{50}{}^{\gamma} + C^{\gamma}}$$

Equation 1.9: *Sigmoidal E_{max} model with γ = Hill coefficient.*

Strictly speaking γ should only be referred to as the "Hill coefficient" if it is an integer as it was mechanistically derived based on the theory of allosteric binding of multiple ligands to the same receptor [11]. However, in a modelling analysis it can be estimated as an empirical exponent that is allowed to take any positive value (Figure 1.6).

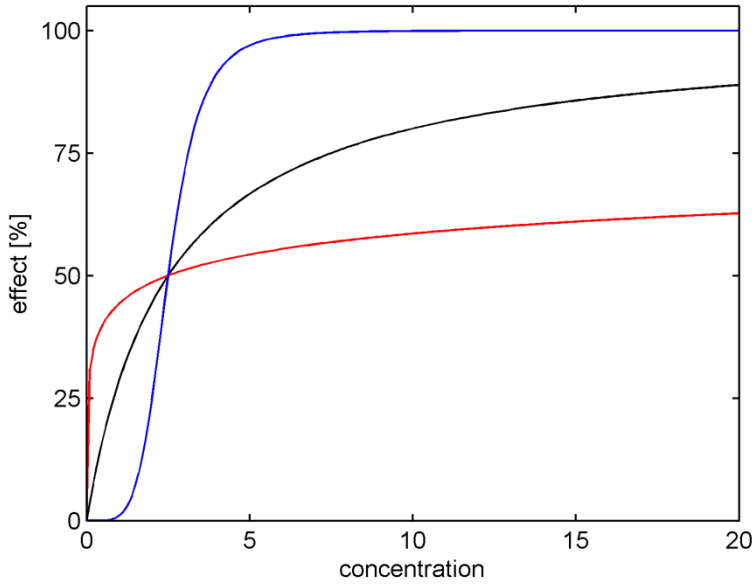


Figure 1.6: *Concentration-effect relationship for an E_{max} model with E_{max} = 100%, EC_{50} = 2.5 units/L, and different values for γ: black line γ = 1, red line γ = 0.25, and blue line γ = 5.*

The E_{max} model can furthermore be extended to include a constant baseline effect E_0 in the absence of the drug (Equation 1.10). If the baseline is changing over time, e.g. due to progress of the disease, E_0 can be replaced by a time varying function in Equation 1.10, which is generally called a disease progression model. More details on disease progression models can be found in Chan and Holford, 2001 [12].

$$E(C) = E_0 + \frac{E_{max} \cdot C}{EC_{50} + C}$$

Equation 1.10: *E_{max} model with constant baseline effect E_0.*

1.1.2.3. Pharmacokinetics-Pharmacodynamics

By combining PK and PD the time course, i.e. onset and duration, of a drug effect can be described.

1.1.2.3.1. Time course of drug effect

Mathematically, we can substitute Equation 1.2 in Equation 1.7 and obtain:

$$E(t) = f_{PD}\left(f_{PK}\left(dose, t, \boldsymbol{\theta}^{\{\mathbf{PK}\}}\right), \boldsymbol{\theta}^{\{\mathbf{PD}\}}\right)$$

Equation 1.11: *Basic equation for pharmacokinetic-pharmacodynamic models.*

From Equation 1.11 it can be seen that the drug effect *E* is now expressed as a function of time *t* depending on dose, the PK parameters $\boldsymbol{\theta}^{\{\mathbf{PK}\}}$, and the PD parameters $\boldsymbol{\theta}^{\{\mathbf{PD}\}}$, effectively describing the time course of the drug effect.

1.1.2.3.2. PKPD models

PKPD models are conventionally divided into two categories, depending on the nature of the link function $f_{PD}(\cdot)$ that is used to combine PK and PD:

- *immediate effect PKPD models*
- *delayed effect PKPD models*

An *immediate effect PKPD model* assumes no time delay between PK and PD. This means that the effect follows the same time course as the plasma concentration of the drug (Figure 1.7), e.g. the maximum effect is achieved at the same time as maximum plasma concentrations are observed.

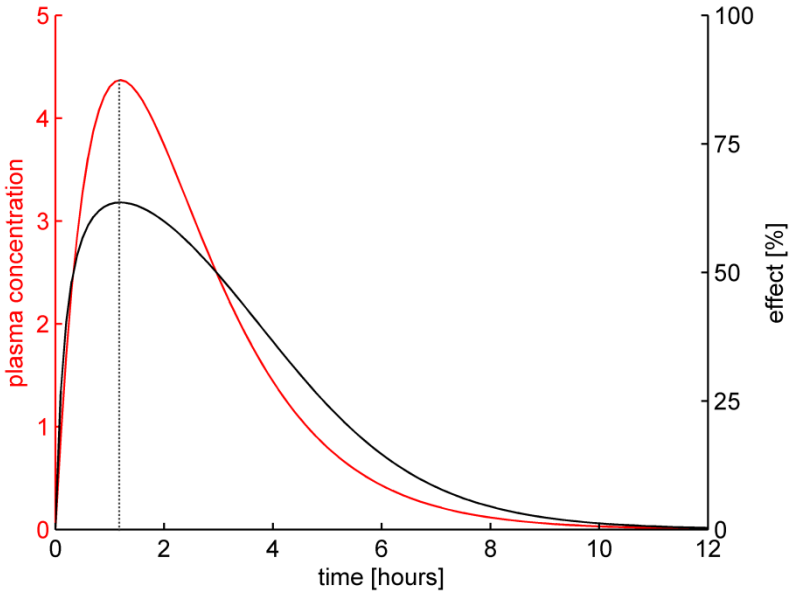


Figure 1.7: *Time course of drug effects for an immediate effect PKPD model. Black line: PD effect with E_{max} = 100%, EC_{50} = 2.5 units/L, and γ = 1. Red line: PK profile for a one-compartment PK model with oral administration, first-order absorption and first-order elimination where D = 10 units, CL = ln(2) hr^{-1}, V = 1 L, F = 1 and k_a = 1 hr^{-1}.*

We can substitute the concentration $C(t)$ in Equation 1.8 (alternatively Equation 1.9 or Equation 1.10) with the plasma concentration described by the PK model (Equation 1.2) to obtain:

$$E(t) = \frac{E_{max} \cdot C(t)}{EC_{50} + C(t)}$$

Equation 1.12: *Immediate effect PKPD model.*

However, only very few drugs do not exhibit a time delay between PK and PD. More often, the PD effect of the drug does not follow the same time course as its PK, e.g. the time of the maximum effect does not correspond to the time of the highest plasma concentration as illustrated in Figure 1.8. This delay can be attributed to various factors, such as a delay in the distribution of the drug from plasma to the site of action, a slow binding of the drug to the receptors, or the drug can stimulate or inhibit physiological processes that have their own time course and which in turn cause the observed effect *E*.

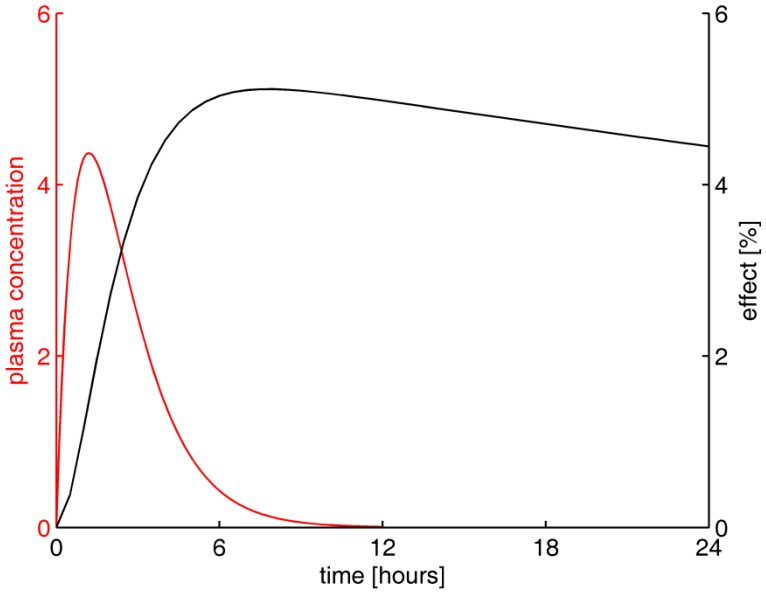


Figure 1.8: *Time course of drug effects for a delayed effect PKPD model. Black line: PD effect with E_{max} = 100%, EC_{50} = 2.5 units/L, and γ = 1. Red line: PK profile for a one-compartment PK model with oral administration, first-order absorption and first-order elimination where D = 10 units, CL = ln(2) hr^{-1}, V = 1 L, F = 1 and k_a = 1 hr^{-1}.*

Two different types of *delayed effect PKPD models* have been developed: the *effect compartment PKPD model* and *turnover PKPD models*.

The *effect compartment model* assumes the presence of an additional compartment in which the drug is assumed to exhibit its effect. The distribution into this hypothetical compartment results in the delay between the PK profile in the plasma (described by the central compartment) and the PD effect. Alternatively, the delayed distribution into the effect compartment can also be interpreted as slow binding to the receptor. This model assumes that only a marginal amount of drug distributes into the effect compartment so that its influence on the mass balance of the PK model is negligible. The volume of distribution of the effect compartment is unidentifiable. However, under the assumption that the steady state concentrations in the central compartment and in the effect compartment are equal, the equilibrium rate constant k_{eq} for the distribution out of the effect compartment can be estimated. Equation 1.13 describes the rate of change of the concentration in the effect compartment C_e

and the PKPD model resulting from substituting C_e into Equation 1.8 (alternatively Equation 1.9 or Equation 1.10):

$$\frac{dC_e}{dt} = k_{eq} \cdot (C - C_e)$$
$$E(t) = \frac{E_{max} \cdot C_e(t)}{EC_{50} + C_e(t)}$$

Equation 1.13: *Effect compartment PKPD model with C_e = drug concentration in the hypothetical effect compartment, k_{eq} = equilibration rate constant.*

Note that the effect in Figure 1.7 was plotted based on Equation 1.13 using k_{eq} = 0.01 hr^{-1}.

Turnover models on the other hand assume that the drug directly affects physiological processes such as the production or elimination of an intermediate *I* and the changes in these processes result in the delayed effect. As the drug can either stimulate or inhibit these processes four turnover PKPD models exist (Equation 1.14) [13]:

$$(\mathbf{A}) \qquad \frac{dI}{dt} = R_{in} \cdot \left(1 - \frac{E_{\max} \cdot C(t)}{EC_{50} + C(t)}\right) - k_{out} \cdot I$$
$$(\mathbf{B}) \qquad \frac{dI}{dt} = R_{in} \cdot \left(1 + \frac{E_{\max} \cdot C(t)}{EC_{50} + C(t)}\right) - k_{out} \cdot I$$
$$(\mathbf{C}) \qquad \frac{dI}{dt} = R_{in} - k_{out} \cdot \left(1 - \frac{E_{\max} \cdot C(t)}{EC_{50} + C(t)}\right) \cdot I$$
$$(\mathbf{D}) \qquad \frac{dI}{dt} = R_{in} - k_{out} \cdot \left(1 + \frac{E_{\max} \cdot C(t)}{EC_{50} + C(t)}\right) \cdot I$$

Equation 1.14: *Turnover PKPD models with I = intermediate, R_{in} = production rate of I, k_{out} = elimination rate constant.* (**A**) *inhibition of production,* (**B**) *stimulation of production,* (**C**) *inhibition of elimination,* (**D**) *stimulation of elimination.*

While the idea of an effect compartment is to some extent adopted in the third part of this thesis, turnover PKPD models are not applied in this work and hence are not elaborated on further. Excellent review articles on PD models as well as PKPD models have been published and provide more details [14-17].

1.1.3. Survival analysis

Survival analysis (also known as time-to-event analysis) quantifies the time until a certain event occurs [18]. It is similar to a PK data analysis in the sense that the independent variable is time. However, while the dependent variable in a PK analysis (plasma concentration) is continuous, survival analysis is concerned with binary data as dependent variable: either event (1) or no event (0).

Five mathematically equivalent functions are used to describe the survival time. These functions and their relationships are introduced briefly in the following section and a simple example is given in Section 1.1.3.2.

1.1.3.1. Functions of survival time

1.1.3.1.1. Survival function

The survival function $S(t)$ defines the probability P of the survival time T of an individual being longer than time t.

$$S(t) = P(T > t) \quad \text{with } S(t) = \begin{cases} 1 & \text{for } t = 0 \\ 0 & \text{for } t = \infty \end{cases}$$

Equation 1.15: *Survival function.*

1.1.3.1.2. Cumulative distribution function

Similarly, the cumulative distribution function (cdf) $F(t)$ describes the inverse probability for T being less than t:

$$F(t) = P(T < t) \quad \text{with } F(t) = \begin{cases} 0 & \text{for } t = 0 \\ 1 & \text{for } t = \infty \end{cases}$$

$$F(t) = 1 - S(t)$$

Equation 1.16: *Cumulative distribution function.*

1.1.3.1.3. Probability density function

The probability density function (pdf) $f(t)$ describes the random distribution of the survival time T and is given as the derivative of the cdf with respect to time t:

$$f(t) = \frac{dF}{dt}$$

Equation 1.17: *Probability density function.*

Note, that with respect to survival analysis $f(t)$ specifically denotes a pdf, whereas in this thesis $f(t)$ is also used to indicate a structural model.

1.1.3.1.4. Hazard function

The instantaneous risk of an event occurring at time t is described by the hazard function $h(t)$, which can be derived as the ratio of pdf and survival function:

$$h(t) = \frac{f(t)}{S(t)}$$

Equation 1.18: *Hazard function.*

1.1.3.1.5. Cumulative hazard function

The cumulative hazard function $H(t)$ describes the cumulative risk of an event occurring at time T and is given as the integral of $h(t)$ over time. It can also be described as the negative natural logarithm of the survival function $S(t)$.

$$H(t) = \int_0^T h(t)dt = -\ln S(t)$$

Equation 1.19: *Cumulative hazard function.*

1.1.3.2. Example

This example illustrates the different functions of survival time for a constant hazard model (Equation 1.20), where the instantaneous risk of an event occurring, i.e. $h(t)$, does not change over time (Figure 1.9).

$h(t) = \lambda$	hazard function
$H(t) = \lambda \cdot t$	cumulative hazard
$f(t) = \lambda \cdot \exp(-\lambda \cdot t)$	pdf
$F(t) = 1 - \exp(-\lambda \cdot t)$	cdf
$S(t) = \exp(-\lambda \cdot t)$	survival function

Equation 1.20: *Functions of survival time for a constant hazard model.*

It can be seen from these equations that a survival analysis based on a constant hazard model is equivalent to a PK analysis using a one-compartment model with i.v. bolus dosing as the survival function in Equation 1.20 equals to Equation 1.3 with a unit dose and a volume of distribution V of 1. The hazard rate constant λ is therefore equivalent to the elimination rate constant k_{el} in a PK analysis, which means that k_{el} can also be interpreted as the instantaneous risk of a single drug molecule to be removed from the circulation.

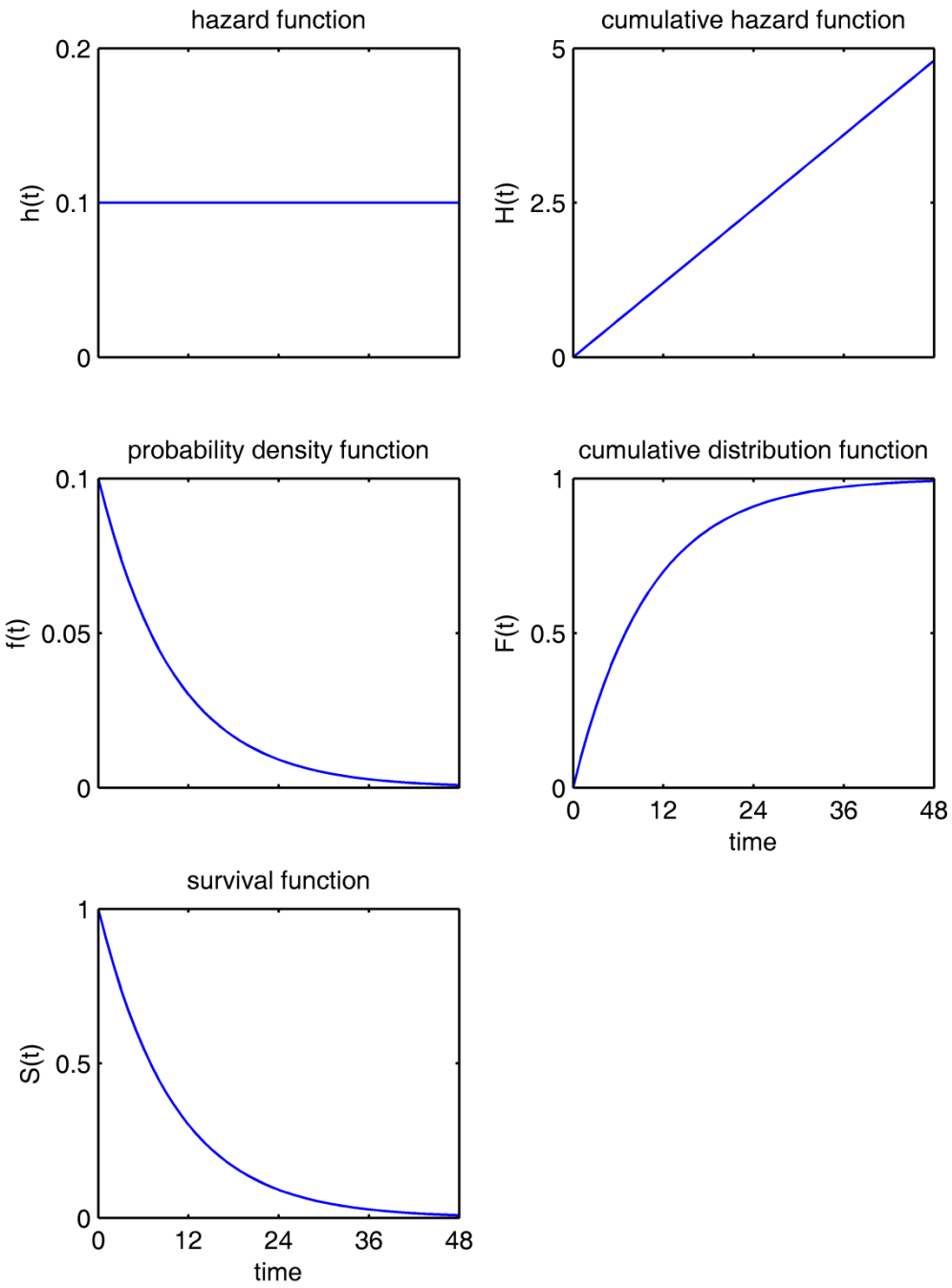


Figure 1.9: *Functions of survival time for a constant hazard model with λ = 0.1 per unit of time.*

1.1.4. Population analysis

Clinical data commonly arises from more than one individual while multiple samples are often taken from each individual, and variability not only between the different people but also within the data from the same individual is observed. With respect to the opening quote of this thesis by Sir William Osler (page 3), understanding this variability is paramount to replacing the "art" in medicine with science.

A population analysis can be used to achieve this goal. Two types of population analysis will be introduced here that can quantify the variability in the data: the *two-stage approach* and the *full population approach*. For comparison, the *naïve pooled approach* for data analysis will also be considered, although this approach does not allow for the quantification of variability in the data. More information on population analysis methods can be found in [19-22].

1.1.4.1. Naïve pooled approach

Using the naïve pooled approach, the data is analysed as if it had arisen from just a single individual (or as if each observation was obtained from a different individual). The whole data set is pooled together and just one set of parameter values is estimated for the corresponding model. In terms of computational effort, this method is the least complex approach. Equation 1.21 describes the use of ordinary least squares (*OLS*) to calculate the objective function value (*OFV*) for the naïve pooled approach:

$$OFV_{OLS}(\boldsymbol{\theta}) = \sum_{i=1}^{N}\left(\sum_{j=1}^{n_i}\left(y_{ij} - f\left(x_{ij}, \boldsymbol{\theta}\right)\right)^2\right)$$

Equation 1.21: *Naïve pooled approach using ordinary least squares* (*OLS*) *to calculate the objective function value* (*OFV*).

Here, $\boldsymbol{\theta}$ denotes the n_p-by-1 parameter vector of the structural model $f(\cdot)$ as in Equation 1.1, i denotes the i^{th} individual, N is the total number of individuals such that $i = (1,\ldots N)$, y_{ij} denotes the j^{th} observation in the i^{th} individual with n_i being the total number of observations for this individual,

and x_{ij} is the j^{th} independent variable for the i^{th} individual. Note, that this notation will be used throughout this thesis.

The naïve pooled analysis aims to find the set of parameters $\hat{\boldsymbol{\theta}}$ that minimise Equation 1.21, i.e. that provides the best fit of function *f* to the data. No between subject variability (BSV) can be estimated from this approach. This results in biased parameter estimates as well as an inflated residual unexplained variability (RUV) [19]. It also needs to be noted that using *OLS* as objective function for a naïve pooled analysis also does not allow estimating the RUV variance parameters. Due to these drawbacks the naïve pooled approach is commonly only used to obtain initial parameter estimates for a more complex population method.

1.1.4.2. Two-stage approach

In a two-stage approach, the data for each individual are analysed separately, e.g. using *OLS* as objective function, resulting in *N* sets of individual parameter estimates $\hat{\boldsymbol{\theta}}_i$ (n_p-by-1 vector). The expectation of the population mean value for the p^{th} parameter $\bar{\theta}_p$ can then be calculated as either the arithmetic or geometric mean of these parameter sets, while the BSV of the p^{th} parameter denoted by Ω_p can be quantified as the variance of the individual estimates of the p^{th} parameter:

$$
\begin{aligned}
OFV_i(\boldsymbol{\theta}_i) &= \sum_{j=1}^{n_i} \left(y_{ij} - f\left(x_{ij}, \boldsymbol{\theta}_i\right)\right)^2 \\
\hat{\boldsymbol{\theta}}_i &= \underset{\boldsymbol{\theta}_i \in \boldsymbol{\Theta}}{argmax}\left(OFV_i(\boldsymbol{\theta}_i)\right) \\
\bar{\theta}_p &= \frac{1}{N} \cdot \sum_{i=1}^{N} \hat{\theta}_{ip} \\
\Omega_p &= \frac{1}{N-1} \cdot \sum_{i=1}^{N} \left(\hat{\theta}_{ip} - \bar{\theta}_p\right)^2
\end{aligned}
$$

Equation 1.22: *Two-stage approach using ordinary least squares as objective function.*

The two-stage approach is a relatively simple method to obtain an estimate of the BSV in addition to the population mean parameters. However, this approach requires rich data for each individual and the obtained BSV tend to be inflated compared to the true variability in the parameters between the individuals as it does not take into account the uncertainty in the individual parameter estimates.

1.1.4.3. Full population approach

A full population approach provides the most accurate and precise quantification of the population mean parameter estimates as well as RUV and BSV for a given data set. In addition, this approach is able to handle sparse data. However it is also the most complex analysis method of these three approaches.

A hierarchical model structure is assumed for a full population approach. The first level in the model hierarchy describes the structural model on an individual level together with the statistical model for RUV. Equation 1.23 gives an example where the RUV is described by an additive error $\boldsymbol{\varepsilon}$ which is assumed to be normally distributed with variance σ^2.

$$y_{ij} = f\left(x_{ij}, \hat{\boldsymbol{\theta}}_i\right) + \varepsilon_{ij}$$

$$\text{with } \boldsymbol{\varepsilon} \sim N\left(0, \sigma^2\right)$$

Equation 1.23: *Structural model for the i^{th} individual in a full population approach. y_{ij} = j^{th} observations, x_{ij} = j^{th} independent variable, $\hat{\boldsymbol{\theta}}_i$ = vector of individual parameter estimates for the i^{th} individual, and σ^2 = variance of the additive error $\boldsymbol{\varepsilon}$.*

The second hierarchical level describes the model for the BSV in the individual parameter estimates (Equation 1.24).

$$\hat{\theta}_{ip} = \hat{\bar{\theta}}_p \cdot e^{\eta_{ip}}$$

Equation 1.24: *Between subject variability of the individual estimate of the p^{th} parameter for the i^{th} individual $\hat{\theta}_{ip}$ in a full population approach, where $\hat{\bar{\theta}}_p$ denotes the population mean parameter estimate and η_{ip} denotes the random effect for the i^{th} individual.*

It is assumed that the BSV terms $\boldsymbol{\eta}_p$ are normally distributed with variance ϖ_{pp}^2, where $\boldsymbol{\Omega}$ is the variance-covariance matrix for all elements of $\boldsymbol{\eta}$ according to:

$$\boldsymbol{\Omega} = \begin{bmatrix} \varpi_{11}^2 & \varpi_{21}^2 & \cdots & \varpi_{n_p 1}^2 \\ \varpi_{12}^2 & \varpi_{22}^2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \varpi_{1n_p}^2 & \varpi_{2n_p}^2 & \cdots & \varpi_{n_p n_p}^2 \end{bmatrix}$$

Equation 1.25: *Variance-covariance matrix for the between subject variability.*

In contrast to a two-stage analysis, where the expectation of the population mean parameter values $\bar{\boldsymbol{\theta}}$ are calculated post hoc as the mean or median of the individual parameter estimates $\hat{\boldsymbol{\theta}}_i$, a full population approach allows estimating the population mean parameter values from the data. Throughout this thesis these population mean estimates will be denoted as $\hat{\bar{\boldsymbol{\theta}}}$ (n_p-by-1 vector).

Using an exponential model to describe BSV as shown in Equation 1.24 is of convenience in a PKPD analysis as it results in a lognormal distribution of $\hat{\boldsymbol{\theta}}_i$ that restricts all parameter estimates to be greater than zero. Given that physiological parameters have a natural lower boundary of zero, this is more plausible than a normal distribution, which would allow for negative parameter estimates as well.

1.1.4.3.1. Non-linear mixed effect modelling

The hierarchical structure of a population analysis model requires the estimation of two types of effects: *fixed effects* given by the structural parameters $\bar{\boldsymbol{\theta}}$, and *random effects* describing the remaining unexplained differences seen in the data set via the variance terms $\boldsymbol{\Omega}$ and σ^2. These types of models are also known as *non-linear mixed effect models* (NLME models). Here, non-linear refers to the non-linearity of the model in the random effect parameters.

Due to the additional complexity of these models, parameter estimation using *OLS* is no longer feasible. Instead, a likelihood based approach is required to allow for the estimation of $\bar{\boldsymbol{\theta}}$, $\boldsymbol{\Omega}$ and σ^2. The likelihood L is expressed as the probability P of the data $\mathbf{Y}$ arising under the structural model f given the parameters $\bar{\boldsymbol{\theta}}$, $\boldsymbol{\Omega}$ and σ^2. Maximum likelihood estimation then aims to find those parameter values that maximise L.

$$L = P\left(\mathbf{Y} \mid \bar{\boldsymbol{\theta}}, \boldsymbol{\Omega}, \sigma^2\right)$$

Equation 1.26: *Likelihood for a NMLE model.*

The most commonly used software package for NMLE modelling used for population PKPD analyses is NONMEM®, originally developed by Lewis Sheiner and Stuart Beal [19,23]. More recently, MONOLIX® was introduced as an alternative NMLE modelling software [24]. An overview of the common estimation methods used to maximise Equation 1.26 with both programs is given next. Note, that more NMLE modelling programs are available and are used in the area of population PKPD. However, here focus is given only on those programs and algorithms that are utilised in this thesis.

1.1.4.3.2. Estimation methods in NONMEM®

No analytical solution is available for NMLE models that allows estimating $\bar{\boldsymbol{\theta}}$, $\boldsymbol{\Omega}$ and σ^2 simultaneously using a likelihood based approach. Instead, when using the most common estimation methods FO (first-order) and FOCE (first-order conditional estimation), NONMEM® minimizes -2x log likelihood (-2*LL*) as objective function based on a numerical solution where $f\left(x_{ij}, \bar{\boldsymbol{\theta}}, \boldsymbol{\eta}_i\right)$ is linearised using a first-order Taylor series expansion around $\boldsymbol{\eta}$. Under the assumption of a normally distributed residual error -2*LL* is equivalent to the extended least squares (*ELS*) given by Equation 1.27 [25]:

$$OFV_{ELS} = \sum_{i=1}^{N}\left(\left(\mathbf{Y}_i - f\left(\mathbf{X}_i, \bar{\boldsymbol{\theta}}\right)\right)\mathbf{V}_i^{-1}\left(\mathbf{Y}_i - f\left(\mathbf{X}_i, \bar{\boldsymbol{\theta}}\right)\right)^{\mathrm{T}} + \log\left(\left|\, \mathbf{V}_i \,\right|\right)\right)$$

Equation 1.27: *Objective function used in NONMEM®* [19].

Here, $|\cdot|$ denotes the determinant, while $\mathbf{V}_i$ denotes the population covariance matrix of the data in the i^{th} individual which depends on $\mathbf{X}_i$, $\bar{\boldsymbol{\theta}}$, $\boldsymbol{\Omega}$ and σ^2. Under the assumption of a single residual error variance σ^2, $\mathbf{V}_i$ is approximated as:

$$\mathbf{V}_i = \mathbf{J}\boldsymbol{\Omega}\mathbf{J}^{\mathrm{T}} + \sigma^2\mathbf{I}_{n_i}$$

Equation 1.28: *Linear approximation of the population covariance matrix* [19].

where $\mathbf{J}$ is the Jacobian matrix of the first partial derivatives of f with respect to the random effects $\boldsymbol{\eta}$, and $\mathbf{I}_{n_i}$ is an n_i-by-n_i identity matrix.

This expansion is evaluated for all $\eta_{ip} = 0$ in the FO method or at a conditional estimate for all η_{ip} denoted $\hat{\eta}_{ip}$ in the FOCE method, and estimates for $\bar{\boldsymbol{\theta}}$ and σ^2 are obtained at each iteration. Subsequently, $\boldsymbol{\eta}_i$ are estimated from the empirical Bayes estimates (EBEs) after each iteration conditioned on the current estimates of the population parameters $\bar{\boldsymbol{\theta}}$. EBEs are the individual parameter estimates $\hat{\boldsymbol{\theta}}_i$ obtained by using $\bar{\boldsymbol{\theta}}$ as prior information. It needs to be noted that the FO method performs well only for models that are close to

linear, and FOCE is therefore the method of choice for NLME analysis in NONMEM®. In addition, the interaction option should be used in the presence of a heteroscedastic error model, so that the algorithm accounts for the relationship between $\boldsymbol{\varepsilon}$ and $\boldsymbol{\eta}$ when calculating σ^2.

Besides these long-standing and commonly used estimation methods, recent versions of NONMEM® include a range of alternative estimation methods such as the Laplacian method (which uses a second-order Taylor series approximation) or various stochastic methods. More detail on the FO and FOCE estimation as well as the alternative methods can be found in [1,23].

1.1.4.3.3. SAEM algorithm in MONOLIX®

With the FO and FOCE method in NONMEM® an exact solution for the linear approximation of the likelihood is obtained. In contrast to that, the SAEM (stochastic approximation expectation maximisation) algorithm is a stochastic method which provides an approximate solution to the exact likelihood. The MATLAB® based implementation of the SAEM algorithm used in MONOLIX® 1.1 [24] was used in Chapter 4 of this thesis as the model used in this chapter was also implemented in MATLAB®. Nonetheless, this algorithm is now also available in NONMEM® (version 7.1. and above).

The SAEM algorithm is an extension of the EM algorithm that is used for linear mixed effect models. For each iteration *k* the EM algorithm first computes the conditional expectation of the log likelihood *Q* (E-step) which is then maximised with respect to all estimated parameters $\boldsymbol{\phi} = \left(\bar{\boldsymbol{\theta}}, \boldsymbol{\Omega}, \sigma^2\right)$ (M-step):

$$\begin{aligned} &\text{E-step:}\quad Q^{\{k\}}(\boldsymbol{\phi}) = E\left(\log P(\mathbf{Y}, \mathbf{X}; \boldsymbol{\phi}) \mid \mathrm{Y}, \boldsymbol{\phi}^{\{k-1\}}\right) \\ &\text{M-step:}\quad \boldsymbol{\phi}^{\{k+1\}} = \underset{\boldsymbol{\phi}\in\boldsymbol{\Phi}}{argmax}\left(Q^{\{k\}}(\boldsymbol{\phi})\right) \end{aligned}$$

Equation 1.29: *EM algorithm for linear mixed effect models.*

For NLME models, *Q* cannot be computed directly. However, the E-step in the EM algorithm can be replaced by a stochastic approximation for these models, resulting in the SAEM algorithm:

$$Q^{\{k\}}(\phi) = Q^{\{k-1\}}(\phi) + \vartheta^{\{k\}}\left(\log P\left(\mathbf{Y}, \mathbf{X}^{\{k\}}; \phi\right) - Q^{\{k-1\}}(\phi)\right)$$

Equation 1.30: *Stochastic approximation of the EM algorithm for NMLE models.*

where $\mathbf{X}^{\{k\}}$ is randomly drawn from the conditional distribution $P\left(\cdot \middle| \mathbf{Y}, \phi^{\{k\}}\right)$ and the step size $\vartheta^{\{k\}}$ is decreasing with each iteration in a predefined manner. For NMLE models, a Markov Chain Monte Carlo (MCMC) procedure is used to obtain the simulations for $\mathbf{X}^{\{k\}}$ required in the expectation step of each iteration. Using the SAEM algorithm in MONOLIX®, the E- and M-steps are repeated for a user-defined number of iterations, usually ≥500. Convergence is assessed in this implementation of the algorithm by the user based on automatically generated graphical output where the trajectory of $\phi^{\{k\}}$ is plotted against *k*.

More information on the SAEM algorithm and MONOLIX® can be found in [24,26,27].

1.1.5. Optimal design

Within the setting of design of experiments, optimal design theory aims to optimise a study design to gain maximum information from a given design within a set of design constraints. The type of information sought depends on the question asked. For example, the aim of discriminating between competing models requires different information than the aim of parameter estimation for a given model. Different statistical criteria have been developed to optimise for different questions. In this thesis, optimal design theory will be applied for parameter estimation and the D-optimality criterion will be introduced here. More detailed information on optimal design can be found in Foo and Duffull [28] and Atkinson and Donev [29].

1.1.5.1. Theory

For clinical studies the design variables to be optimised are most often the number and timing of blood samples with the aim of gaining as much information as possible on the underlying parameters in the model. A sample taken at a time point where the response is most sensitive to change in a particular parameter value provides the most information about that parameter value and thus is optimal for estimating the parameter.

The sensitivities of a function *f* with respect to changes in the parameters $\boldsymbol{\theta}$ can be expressed as the partial derivatives of the function with respect to the parameters. This can be expressed as a sensitivity matrix of partial derivatives over all model parameters $\boldsymbol{\theta} = \left(\theta_1, \ldots \theta_{n_p}\right)^{\mathrm{T}}$ at all time points $\mathbf{t} = (t_1, \ldots t_n)$ which is denoted as the Jacobian $\mathbf{J}$ (dimension n_p-by-n):

$$\mathbf{J} = \begin{bmatrix} \frac{\partial f(t_1)}{\partial \theta_1} & \cdots & \frac{\partial f(t_n)}{\partial \theta_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial f(t_1)}{\partial \theta_{n_p}} & \cdots & \frac{\partial f(t_n)}{\partial \theta_{n_p}} \end{bmatrix}$$

Equation 1.31: *Jacobian matrix.*

The Fisher Information matrix ($\mathbf{M_F}$) can then be calculated according to:

$$\mathbf{M_F} = \mathbf{J}\boldsymbol{\Sigma}^{-1}\mathbf{J}^{\mathrm{T}}$$

Equation 1.32: *Fisher Information matrix.*

Here, $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_n$, where σ^2 is the variance of the residual error assuming an additive error model and $\mathbf{I}_n$ is an *n*-by-*n* identity matrix, yielding a diagonal error matrix $\boldsymbol{\Sigma}$. Note, that the parameter vector $\boldsymbol{\theta}$ can informally be assessed to be locally identifiable when $\mathbf{M_F}$ is positive definite for a given design.

The lower bound of the standard error of the parameter vector $\boldsymbol{\theta}$ is given by the square root of the diagonal entries of the inverse of $\mathbf{M_F}$. Therefore, maximizing $\mathbf{M_F}$ is equivalent to minimising the standard error of the parameter estimates.

1.1.5.1.1. Local D-optimality criterion

It is common to summarise the $\mathbf{M_F}$ by taking its determinant $|\mathbf{M_F}|$ yielding the local D-optimality criterion Ψ_D. The local D-optimal design Ψ_D^* is gained by maximising Equation 1.33 over the design space **T**, where $\mathbf{T} \in \mathbb{R}^+$.

$$\Psi_D^* = \underset{\mathbf{t}\in\mathbf{T}}{argmax} \left(\left| \mathbf{M_F}(\boldsymbol{\theta}, \mathbf{t}) \right| \right)$$

Equation 1.33: *Local D-optimality criterion.*

As the $\mathbf{M_F}$ is a function of the parameters $\boldsymbol{\theta}$, the local D-optimal design depends on the parameter values as well, thus these would need to be known *a priori* which is usually not the case.

1.1.5.1.2. Robust D-optimality criteria

Uncertainty in the parameter values can be incorporated by assuming a distribution of parameter uncertainty and applying a robust D-optimality criterion. Various robust criteria have been developed, such as ED (expectation of the determinant), DE (determinant of the expectation) MMD (maximin D-optimal design) and AIP (approximation to the preposterior information, equivalent to the negative log of the expectation of the determinant) [30-32]. In this thesis, a hypercube log D-optimal design (HClnD) will be applied as robust optimality criterion [33], and the introduction here will focus on this criterion only.

In contrast to other robust criteria, HClnD does not require the distribution of the uncertainty in the parameter values to be specified *a priori* but rather a lower and upper point estimate. 2^{n_P} parameter sets are formed representing all possible combinations of the 2.5th and 97.5th percentiles of the prior parameter distribution. The logarithm of the determinant of the $\mathbf{M_F}$ is then calculated for each of the parameter sets $\boldsymbol{\theta}^{\{k\}}$ and the design is optimized based on their summation according to:

$$\Psi_D^* = \underset{\mathbf{t}\in\mathbf{T}}{argmax}\left(\sum_{k=1}^{2^{n_p}} \frac{\alpha}{n_p} \ln\left(\left| \mathbf{M_F}\left(\boldsymbol{\theta}^{\{k\}}, \mathbf{t}\right) \right| \right) \right)$$

Equation 1.34: *HClnD optimal design criterion.*

In this work, the weighting of each model (α) is fixed to 1 and the number of parameters that are to be estimated n_p is the same for all models.

1.1.5.1.3. D-efficiency

The D-efficiency *Eff* of any given design Ψ_D in comparison to the D-optimal design Ψ_D^* is then calculated according to

$$Eff\,(\%) = \left(\frac{\Psi_D}{\Psi_D^*} \right)^{1/n_p} \cdot 100\%$$

Equation 1.35: *D-Efficiency.*

The efficiency is related to the experimental effort, whereby a value of, for example, 50% indicates that the design in the numerator is half as efficient as the design in the denominator and doubling the experimental effort of the former would yield equivalently informative designs.

1.2. Red blood cells

RBCs, or erythrocytes, are the main cellular component of human blood and a very unique type of cell. Their main function is to transport oxygen (O_2) from the lungs to the body tissues and carbon dioxide (CO_2) back from the tissues to the lungs. Because of this specialised function, RBCs are almost entirely filled with haemoglobin (Hb), the protein that binds O_2 and CO_2 and that gives blood its red colour [34].

1.2.1. RBC production

RBCs are produced in the stroma of the bone marrow, a process known as erythropoiesis, where pluripotent haematopoietic stem cells differentiate over several precursor cell types into reticulocytes [34]. This differentiation is mediated by the hormone erythropoietin (EPO), which is produced in the kidneys in response to reduced O_2 content in the blood. Thus, a feedback loop between RBC production in the bone marrow, O_2 content in the blood (an indicator for the number of RBCs in the circulation) and EPO production in the kidneys exists.

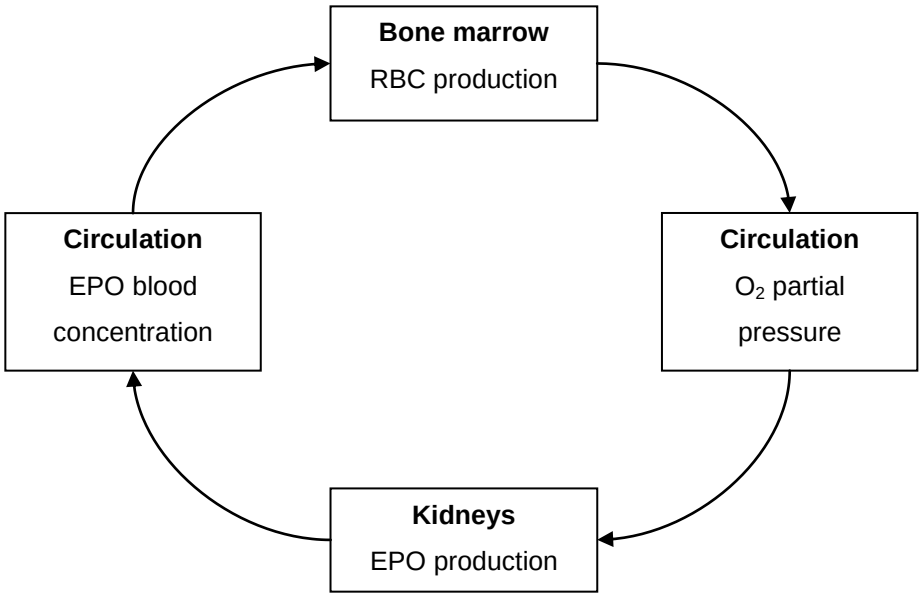


Figure 1.10: *Erythropoietic feedback circle.*

During the differentiation process in the bone marrow, the cell's nucleus degenerates and is extruded, sacrificing the ability of the cell to synthesise proteins, perform self-renewal and mitosis for a greater Hb carrying capacity. The resulting reticulocytes migrate into the circulation and subsequently differentiate into mature RBCs within one to two days by degeneration and loss of the remaining organelles [34].

1.2.2. RBC lifespan & destruction

Without the ability of self-renewal and mitosis, mature RBCs are committed to die after a certain period of time. This period of time is commonly referred to as the *lifespan* of RBCs and is generally accepted to be approximately 120 days, although this value lacks clear scientific evidence. Lifespan values reported in the literature differ depending on the method used to determine the lifespan, and mean lifespan values between 100 and 125 days have been reported [34-36].

1.2.2.1. Lifespan concept & destruction mechanisms

Four general processes are involved in the physiological destruction of RBCs:

- Senescence (death from old age)
- Mid-life destruction of misshapen RBCs
- Random destruction and loss during the circulation
- Early death of unviable RBCs and due to neocytolysis

It has been suggested, that the potential lifespan of a RBC is an innate characteristic of the cell itself. It is assumed to resemble the ability of the cell to resist the stress the cells are exposed to during the circulation in the body, e.g. shear forces in the capillaries [37]. Mature RBCs are not able to repair any structural damage or metabolic failures by *de novo* synthesis of proteins or lipids [34]. Therefore, their ability to resist destructive processes is limited and eventually determines the cell's lifespan [37]. This limited ability decreases gradually during the lifetime of the cell, finally resulting in death of the cell

due to old age which is termed senescence. Furthermore, misshapen RBCs have an inherent reduced ability of resistance and, therefore, are removed from the circulation earlier than normal RBCs.

RBCs are also subject to age-independent random destruction, at least to a small fraction in healthy humans and to a greater extent in certain types of anaemia [38]. Furthermore, erythrokinetic studies have shown that erythropoiesis is partly ineffective resulting in unviable RBCs [39], which either never reach circulation or are destroyed shortly after their release [40]. Additionally, selective reticuloendothelial sequestration of the youngest RBCs termed neocytolysis has been described [41]. Neocytolysis is a fast and fine tuned mechanism to control the number of circulating RBCs [41], and is part of the pathophysiology of the anaemia of chronic kidney disease (CKD) [42].

1.2.2.1.1. Effect of pathological conditions on RBC lifespan and destruction

The influence of pathological conditions on the RBC lifespan and destruction mechanisms is not fully understood. In CKD for example, RBC survival is decreased in addition to the diminished RBC production that results from a decreased production of EPO by the kidneys [43,44]. However, it is unclear whether this decreased survival is due to an accelerated senescence, i.e. the cells have a generally decreased survival capacity, or whether the uremic environment results in an increased destruction of the cells irrespective of their age. Deeper insight into the underlying mechanisms of RBC destruction would be desirable to obtain a better understanding of the anaemia of CKD, but also other diseases affecting RBC survival.

1.2.2.2. Destruction sites & stimuli

Damaged and senescent RBCs are removed by the reticuloendothelial system (RES), predominantly by the spleen, but also by the liver [34,45].

RBCs are sequestered within the vascular spaces of the RES, where they are haemolysed after phagocytosis by macrophages. Various factors are involved in triggering the removal of RBCs in the RES, most of which are related to the ability of the cell to deform normally or the cell's membrane

characteristics. Cell deformation is crucial for the RBCs to pass through the small gaps in the filtration beds in the RES, while altered membrane characteristics can trigger phagocytosis by macrophages. The following characteristics have been reported to trigger RBC destruction:

- decreased activity of glycolytic enzymes with subsequent metabolic failure [46]
- increased intracellular calcium concentrations leading to a higher cell density and decreased deformability [47]
- oxidative injury and polymerisation of cytoskeletal proteins resulting in increased cell stiffness [34]
- changes in the cell membrane affecting deformability and surface charge [34,45]
- increased binding of autologous antibodies to antigens in the RBC membrane [34]

In addition, changes in EPO concentrations not only mediate RBC production but also influences RBC destruction in the RES. Neocytolysis was shown to be induced by rapidly decreasing EPO concentrations which result in wider gaps in the filtration beds of the RES and also in direct changes on the RBC membrane [48-50]. These mechanisms allow for a faster decrease in RBC numbers when shifting to lower altitude than would be possible by removal of old and damaged RBCs alone [41]. Neocytolysis may also result from rapid and large fluctuations in EPO concentrations which can occur with i.v. EPO therapy [42,51].

1.2.3. Methods to determine RBC survival

To estimate the lifespan of RBCs it is generally considered necessary to follow the cells over their entire life. This is usually achieved by using a label. Two different types of labelling methods have been established: *cohort labelling* and *random labelling methods* [34,52].

1.2.3.1. Existing RBC labelling methods

An overview of existing labelling methods for RBCs is given in Table 1.1. Using cohort labelling methods newly formed RBCs are labelled as a group, while random labelling involves labelling RBCs of all ages that are present at one point in time.

Table 1.1: *Labelling methods for RBCs.*

Label	Mechanism
Cohort labelling methods	
Radioactive iron (^{59}Fe)	Incorporation into the heme group during the production of haemoglobin [35,53]
Glycine tagged with isotopes (^{15}N or ^{14}C)	Incorporation into the protoporphyrin part of the heme group [35,38,54]
Random labelling methods	
Differential agglutination method	Transfusion of donor RBCs into a recipient with a different blood group but without the ability of producing antibodies against the blood group of the donor [55]
Radioactive chromium (^{51}Cr)	Complexation with haemoglobin [56-58]
Radioactive diisopropyl-fluorophosphate (DF^{32}P)	Binding to intracellular cholinesterases [59-61]
Biotin	Covalent binding to membrane proteins [36]

Unfortunately, all labelling methods have significant limitations. In the case of cohort labelling methods the label is usually incorporated into Hb during RBC production over several days. Therefore, the labelled cells do not have exactly the same age. Other problems associated with both types of labelling techniques include the reuse of a label released after the breakdown of a cell and random loss of label from the cells during their circulation in the body. For an accurate estimation of the RBC lifespan it would be necessary to account for all of those problems occurring with the label used.

Of the methods mentioned in Table 1.1, the biotin technique would be expected to provide the most accurate estimation of the RBC lifespan. Biotin is permanently bound to the RBC membrane in a covalent manner and RBCs carrying the biotin label can be detected with high sensitivity using flow-cytometry. However, this method is not widely used for clinical purposes due to the expensive equipment required and the complex procedure. Therefore, only limited information on this method exists in comparison to the more commonly applied techniques, especially the radioactive chromium (^{51}Cr) method.

1.2.3.2. Other methods

An alternative method that has been used to determine RBC survival and that does not rely on the concept of labelling is the CO exhalation method [62-65]. This method measures the CO content in the exhaled breath and therefore quantifies heme turnover as one molecule of CO is produced per molecule of destroyed heme. The measured CO content is then converted mathematically into a RBC lifespan value where the contribution of the breakdown of other heme containing enzymes and proteins such as myoglobin is accounted for empirically.

While CO exhalation avoids reuse and label loss, it is confounded by environmental issues and only gives an indirect measure of RBC turnover at a certain time point. This method does not follow RBCs longitudinally and therefore does not provide any information on the RBC lifespan distribution. Additionally, the equations used to derive an RBC lifespan value from the CO content in the breath are potentially confounded. The fraction of CO attributed to RBC breakdown in these equations is usually determined based on the assumption of an average RBC lifespan of 120 days, and not all alternative physiological mechanisms that result in CO production other than RBC breakdown are accounted for in most of these equations.

1.2.3.3. Reference method to determine RBC survival

The International Committee for Standardization in Hematology (ICSH) recommends random labelling with ^{51}Cr to determine RBC survival [66]. This method is therefore most often used in clinical practice.

Although ^{51}Cr is firmly bound to intracellular Hb *in vitro* [56,57], a permanent loss of label is observed *in vivo* [57,67]. This loss seems to be partly due to the dissociation of the chromium-haemoglobin complex. Previous authors refer to this loss as "elution" [57,67], and correcting factors were established to account for it [66]. However, the extent of elution varies depending on the labelling method and is additionally increased in haemolytic disorders [67]. Furthermore, it was shown more recently that Hb is lost from RBCs due to vesiculation [68]. This process further increases the amount of lost label. Yet, the loss of label due to vesiculation of Hb has not been considered in previous calculations of the RBC lifespan based on ^{51}Cr labelling experiments which casts doubt on the accuracy of the RBC lifespan values determined in this way.

1.2.4. Previously proposed lifespan models

Mathematical models have been used to describe RBC survival since the first labelling methods for RBCs were established in the 1940s and 1950s. Most of these models assumed the lifespan to be uniform and finite for all RBCs in an individual (i.e. a point distribution); however random destruction and simple lifespan distributions, such as normal or lognormal distributions, were considered as well [38,54,69]. An extensive review of the early work can be found in Berlin *et al*. [35].

Here, focus will be given to models developed more recently in the area of pharmacometrics.

1.2.4.1. Lifespan based indirect response models

Over the past 20 years, lifespan based indirect response (LIDR) models have been developed and used to describe various haematopoietic cell lines, including erythropoiesis in rats [70]. A recent review by Krzyzanski and Perez Ruixo excellently describes the fundamentals, applications and limitations of LIDR models [71].

LIDR models are extensions of simple turnover models (Figure 1.11, Equation 1.36) where the cell loss rate $R_{out}(t)$ is described as a function of time dependent on the lifespan distribution, denoted by the pdf, $f(t,z)$, of the cells (Equation 1.37). The general notation used in Equation 1.37 allows the pdf to be time variant, i.e. cohorts of cells produced on different days are allowed to have different lifespan distributions.

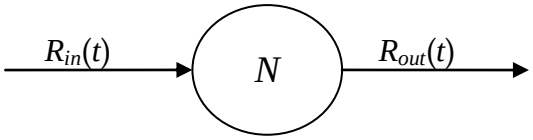


Figure 1.11: *Schematic of a cell turnover model. $R_{in}(t)$ = cell production rate, $R_{out}(t)$ = cell loss rate, N = total number of cells in the system.*

$$\frac{dN}{dt} = R_{in}(t) - R_{out}(t)$$

Equation 1.36: *Mass balance equation for a cell turnover model.*

$$R_{out}(t) = \int_0^{\infty} R_{in}(t-z) \cdot f(t-z, z)dz$$

Equation 1.37: *Time-dependent cell loss rate R_{out} as a function of the lifespan distribution of the cells denote by the corresponding pdf $f(t,z)$.*

Equation 1.37 is equivalent to the convolution $R_{in}*f(t)$ and can be substituted in Equation 1.36, which then becomes a delayed differential equation.

Basic LIDR models use a point density as the pdf where all cells have the same lifespan. These distributions are either assumed to be time invariant in models where only a change in cell production is of interest [72], or time variant where the location of the point density is allowed to shift over time, e.g. due to the influence of agents that affect the elimination of the cells [73,74]. However, the main simplifying assumption that cells produced on the same day will all have the same lifespan holds true even for time varying basic LIDR models.

Despite being more realistic, continuous pdfs have only been used sporadically as lifespan distributions in LIDR models [75,76], mainly because solving the corresponding delayed differential equation is very difficult.

It should be noted that all LIDR models proposed so far assume cell destruction to occur due to one mechanism only, namely senescence. Only the use of a complex pdf would allow for different destruction mechanisms in a LIDR model.

1.2.4.2. Transit compartment models

Transit compartment (TC) models describe the survival of a cell population based on a series of catenary compartments (Figure 1.12) where cell death is assumed to occur only by senescence and is equivalent to the removal of the cells from the last compartment.

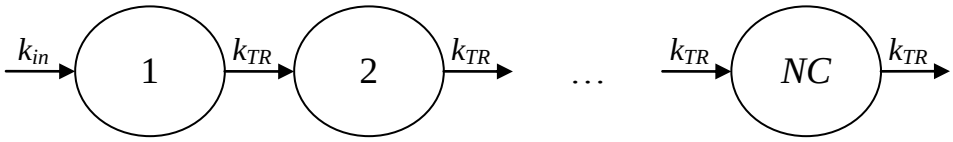


Figure 1.12: *Schematic of a transit compartment model. NC = number of compartments, k_{TR} = transfer rate constant, k_{in} = production rate constant.*

TC models can be regarded as a special case of LIDR models, where the underlying pdf is described by a gamma function [77]. However, they are easier to solve mathematically as they do not require delayed differential equations; and a closed form solution is available [9,10]. Note, that if the

number of compartments approaches infinity the TC model becomes equivalent to the LIDR model as the gamma function will approach a point distribution.

TC models with different numbers of compartments have been used to describe the survival for RBCs in literature [78-80]. Of particular interest is the work of Kalicki *et al.* who extended the basic TC model for RBC survival by including random destruction as a first-order loss process from all 29 compartments in their model [80]. This was the first attempt to include two different types of RBC destruction into one of the recent models for RBC survival.

1.2.4.3. A proposed RBC lifespan model

In contrast to the more parsimonious objectives of the models described before, I have previously developed a model for RBC survival based on physiological destruction mechanisms of RBCs in a bottom-up model building approach [81,82].

Development of this model was based on the theory of aging and longevity and survival analysis methods were applied to propose a pdf that accounts for four plausible destruction mechanisms of RBCs in the human body:

- early removal of unviable RBCs from the circulation
- age-independent random destruction
- premature removal of misshapen RBCs
- normal cell death due to senescence

1.2.4.3.1. Proposed RBC lifespan distribution

For this purpose, a bathtub-shaped hazard function was used as this type of function generally includes three components: 1) an instantaneous elevated risk of system failure followed by rapidly decreasing risk rate in the earliest phase of existence (analogous to infant mortality), 2) a constant hazard rate that spans the whole existence and that describes random loss (analogous to

random death) and 3) a final phase of increased system failure due to senescence of the system (analogous to an age-dependent increasing mortality rate). This results in a typically u-shaped hazard curve, thus the name *bathtub function*.

Bebbington *et al.* proposed a pdf for human mortality showing a bathtub-shaped hazard curve for the population of Indonesia [83]. They combined a modified flexible Weibull (FW) distribution, denoted by the parameters $s_1 > 0$, $s_2 > 0$ and $c \geq 0$, and a reduced additive Weibull (RAW) distribution, denoted by the parameters $r_1 > 0$ and $r_2 > 1$. The resulting pdf for lifetime $t > 0$ (equivalent with age at time of death) is denoted by the following equation:

$$f_{MIX}(t,\boldsymbol{\theta}) = m \cdot f_{FW}(t) + (1-m) \cdot f_{RAW}(t)$$

$$f_{MIX}(t,\boldsymbol{\theta}) = m \cdot \left(\exp\left(-\exp\left(s_1 \cdot t - s_2/t\right) - c \cdot t\right) \cdot \left(\left(s_1 + s_2/t^2\right) \cdot \exp\left(s_1 \cdot t - s_2/t\right) + c\right)\right) + (1-m) \cdot \exp\left(-\left(r_1 \cdot t\right)^{r_2} - \left(r_1 \cdot t\right)^{1/r_2}\right) \cdot \left(\left(r_1 \cdot r_2\right)\left(r_1 \cdot t\right)^{r_2-1} + \left(r_1/r_2\right)\left(r_1 \cdot t\right)^{(1/r_2)-1}\right)$$

$$\boldsymbol{\theta} = \left[c, m, r_1, r_2, s_1, s_2\right]^{\mathrm{T}}$$

Equation 1.38: *Probability density function of the proposed RBC lifespan model.*

Here, the proportional contribution of each distribution is determined by the unit-less mixing parameter m ($0 \leq m \leq 1$). The mean parameter values for the Indonesian population were estimated as r_1 = 0.0260 days^{-1}, r_2 = 5.7936, s_1 = 0.0434 days^{-1}, s_2 = 234.94 days, c = 0.0029 days^{-1} and m = 0.7843 [83].

The underlying modified flexible Weibull distribution can be regarded as the pdf of the healthy subpopulation, shown in Figure 1.13A.

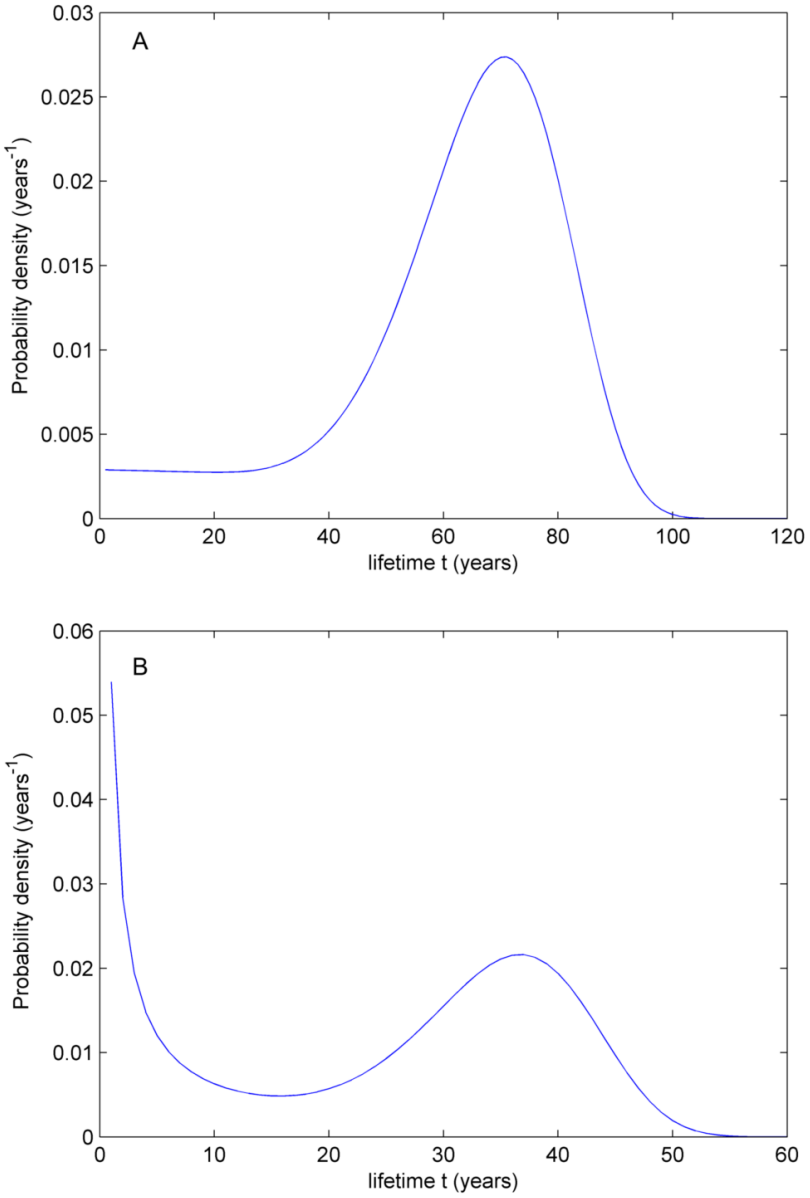


Figure 1.13: **A** *Probability density function $f_{FW}(t)$ of the flexible Weibull distribution accounting for death due to senescence and random loss in the healthy population.* **B** *Probability density function $f_{RAW}(t)$ of the reduced additive Weibull distribution showing early infant mortality and reduced life expectancy of individuals with congenital disorders.*

Here, death occurs either due to senescence (component 3) or age-independently in a random fashion (component 2), e.g. fatal accidents or highly virulent infections that kill irrespective of age. Initially, the curve is elevated from the x-axis by the influence of the constant hazard rate *c*, while the maximum at the age of 70 years accounts for death due to senescence. On the other hand component 1, infant mortality together with a reduced life expectancy is described by the reduced additive Weibull distribution (Figure 1.13B). Infant mortality is rapidly decreasing during the first five years and the mortality rate reaches a minimum at the age of 15 years. The saddle at the age of approximately 37 years is thought to account for a reduced life expectancy of those in the population who have congenital disorders that were not associated with immediate mortality (e.g. cystic fibrosis).

Combining these two distributions according to the mixing parameter m = 0.7843 (from the Indonesian population) results in the overall lifespan distribution for the Indonesian population (Figure 1.14A, Equation 1.38). The main maximum at the age of 70 years equates the maximum of the modified flexible Weibull distribution, while the saddle at the age of approximately 37 years equates the maximum of the reduced additive Weibull distribution. This results in the saddle between the age of 35 and 50 years of the corresponding bathtub-shaped hazard function (Figure 1.14B), which represents the instantaneous risk of death at any point in time.

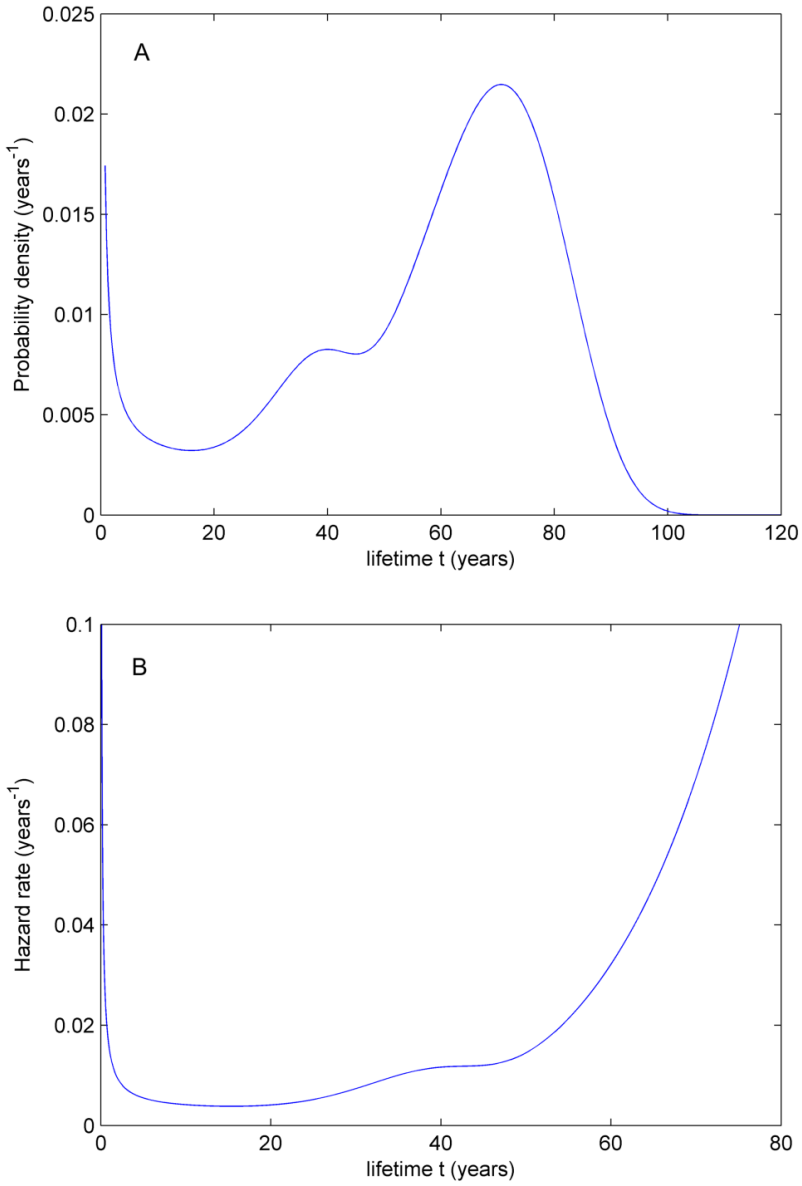


Figure 1.14: *Combined model for the Indonesian population:* **A** *Probability density function* $f_{MIX}(t)$. **B** *Corresponding bathtub-shaped hazard curve* $h_{MIX}(t)$.

The Indonesian model was chosen as a template for the lifespans of RBCs due to characteristics that compare to a potential bathtub-shaped hazard function for RBC survival. The characteristics of the Indonesian model are mapped to the postulated features for the RBC model in Table 1.2.

Table 1.2: *Indonesian model versus RBC model.*

Indonesian model	RBC model
Early infant mortality	Early death of unviable RBCs or neocytolysis
Constant hazard rate during the main lifetime	Random destruction of RBCs during their circulation in the body
Reduced life expectancy of individuals with congenital disorders	Reduced lifespan of misshapen RBCs
Exponentially increasing hazard rate with increasing age	Removal of normal RBCs from circulation due to senescence

1.2.4.3.2. Implementation of the model in MATLAB®

The model was implemented in the programming software MATLAB® (The MathWorks™ Inc., Natick, USA). The overarching premise was to assign each simulated RBCs an individual lifespan, randomly sampled from the lifespan distribution described by the proposed pdf.

Several difficulties arose from this idea:

1) The complex structure of the pdf does not allow direct sampling of lifespans from the corresponding cdf as the inverse of the cdf is not available in closed form. To circumvent this problem, a rejection sampling method was used to generate lifespans with the same distribution (see Appendix A.1.1 for a description of this method) [84,85].

2) The sampled lifespans had to be rescaled to the dimension of RBC survival. This was achieved by multiplying each lifespan by the ratio of assumed median RBC age (in days) and median human age (in years). Applying this converting factor preserves the shape of the original distribution, while rescaling it to the dimension of RBC survival. A median RBC age of 115

days was assumed, according to the mean lifespan of pooled RBC survival data determined using ^{51}Cr and the historically frequently used method of differential agglutination [34].

3) This version of the model suffered from very long run times as the initial MATLAB® code was very inefficient relying on multiple loops and excessive interim data storage.

1.2.4.3.3. Relation to this thesis

The work presented in the second part of this thesis is based on the newly proposed RBC lifespan model. In Chapter 2, the model is improved to overcome the above mentioned difficulties. It is furthermore enhanced to describe different labelling techniques commonly used to determine RBC survival *in vivo*.

The ability of estimating all parameters in the final model from RBC survival studies using hypothetical, ideal labelling methods as well as currently available, non-ideal methods is assessed based on an information theoretic approach. The model is then applied to *in vivo* data to compare RBC survival in patients with CKD and healthy controls.

1.2.5. RBCs and clinical data

Besides measuring the lifespan of RBCs, various other data derived from RBCs can be used in clinical practice. It reaches from standard laboratory measurements of haematological parameters such as Hb concentrations or the mean corpuscular volume (MCV) to the most common marker of glycaemic control in diabetic patients, glycated haemoglobin HbA_{1c}.

In addition, RBCs can also serve as matrix to measure PK data for drugs that accumulate intracellularly. An example of this will be explored in Part III of this thesis where a population PK model for intracellular RBC concentrations of methotrexate (MTX) is developed.

1.3. Methotrexate

MTX is a folate analogue that is widely used as an immune suppressant and anti-cancer drug because of its anti-inflammatory and anti-proliferative properties [86]. The structures of MTX as well as folic acid are shown in Figure 1.15.

In the following, a brief overview of MTX PK and PD will be given with emphasis on its use in rheumatoid arthritis (RA).

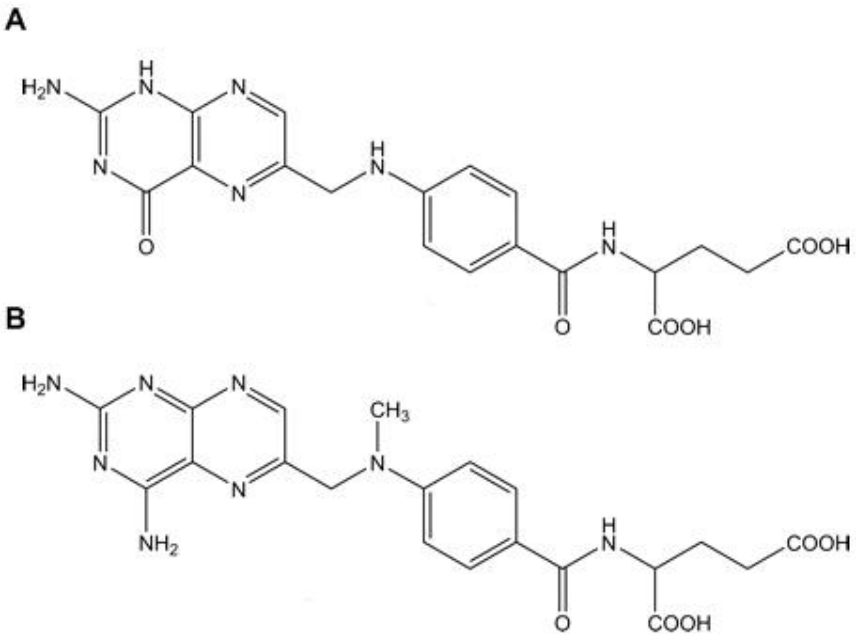


Figure 1.15: *Chemical structures of folic acid* (**A**) *and methotrexate* (**B**).

1.3.1. Pharmacokinetics

1.3.1.1. General pharmacokinetics of the parent drug

MTX can be administered orally or parenterally. Absorption from the gastrointestinal (GI) tract is by means of an active transport mechanism which is saturable at high doses. The bioavailability of oral MTX doses less than 30 mg/week is approximately 70% [86-88].

A biphasic PK profile is observed after i.v. dosing of MTX that is best described by a two-compartment model [89-92]. *In vivo* metabolism of MTX is minimal; the main hepatic metabolite after high-dose administration of MTX is 7-hydroxy-MTX which results from hydroxylation of the pteridin structure in the molecule. 81% of the administered dose is excreted unchanged via the kidneys within 48 hours [86]. Glomerular filtration and tubular secretion both play a role in the renal clearance of MTX [86].

1.3.1.2. Intracellular metabolism via polyglutamation

Similar to folate, MTX is also taken up into cells via the reduced folate carrier (RFC) [93]. Inside the cells, the enzyme folylpolyglutamate synthetase (FPGS) adds glutamate moieties to the parent molecule in a stepwise manner, resulting in the formation of polyglutamated MTX metabolites (MTXPGs) [94-96]. The MTX parent molecule contains one moiety of glutamate onto which the additional glutamates are added via γ-linkage forming a polyglutamate chain. A second enzyme, γ-glutamyl hydrolase (γGH) catalyses the reverse reaction and cleaves one and/or two terminal moieties of glutamate [97]. MTXPGs accumulate intracellularly, as only the parent drug can leave the cells via multi drug resistance transporters (MDRT) of the ABC (ATP binding cassette) family [95,98,99].

1.3.2. Pharmacodynamics

1.3.2.1. Postulated mechanisms of action

As a folate analogue MTX inhibits a key enzyme in the folate pathway, dihydrofolate reductase (DHFR) (Figure 1.16) [100]. DHFR converts folate to its active form tetrahydrofolate which acts as an acceptor and donor of methyl groups in other enzymatic reactions. Inhibition of DHFR by MTX therefore affects all cellular processes that require C_1-transfer.

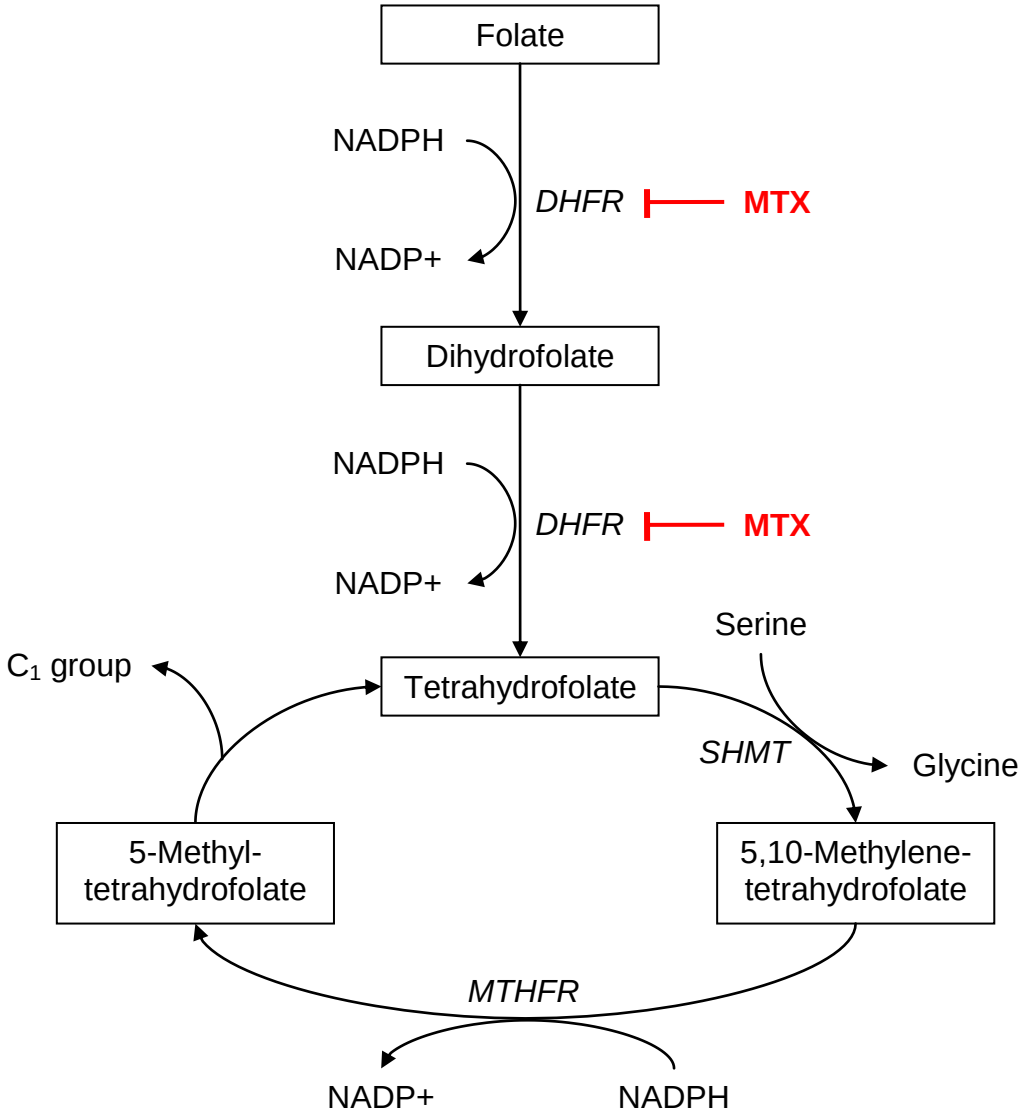


Figure 1.16: *Folate pathway and mechanism of action of MTX. NADPH / NADP⁺ = nicotinamide adenine dinucleotide phosphate, SHMT = serine hydroxymethyltransferase, MTHFR = methylenetetrahydrofolate reductase.*

MTX also interferes with DNA synthesis by direct inhibition of folate dependent enzymes such as 5-aminoimidazole-4-carboxamide ribonucleotide transformylase (ATIC) and thymidylate synthase (TYMS). These enzymes are essential for the *de novo* synthesis of purines and pyrimidines (Figure 1.17) [101-103], and their inhibition is mainly responsible for the anti-proliferative action of MTX. In addition, MTX also has anti-inflammatory and immunosuppressant properties, which can be attributed to a reduction in the production of pro-inflammatory cytokines such as interferon-γ, TNF-α and interleukin 1β as a result of an increased intracellular accumulation and secretion of adenosine (Figure 1.17) [104,105].

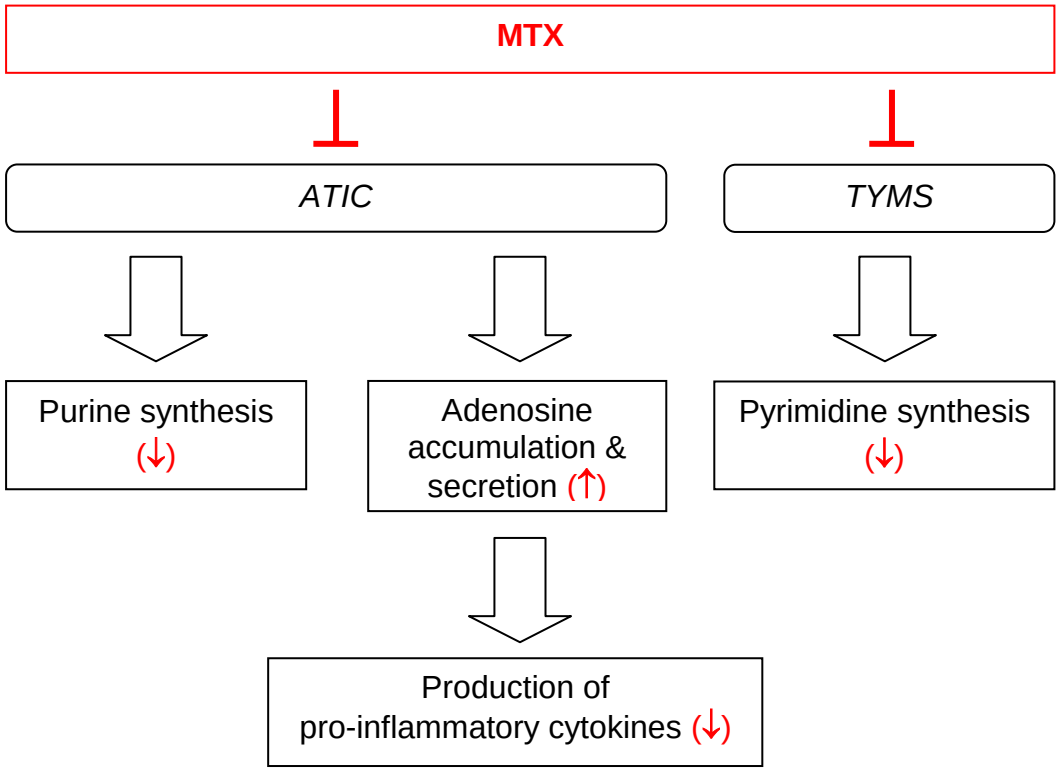


Figure 1.17: *Anti-proliferative and anti-inflammatory mechanism of action of MTX. ATIC = 5-aminoimidazole-4-carboxamide ribonucleotide transformylase, TYMS = thymidylate synthase [106].*

A more detailed illustration of the postulated pathway of action of MTX is provided in Appendix A.1.2, App Fig. 1.2.1.

MTXPGs are active metabolites that have been shown to exhibit an even higher affinity for folate dependent enzymes than MTX, and a similar affinity to DHFR [86,98]. At least part, if not the majority, of the activity of MTX can therefore be attributed to these active metabolites, and their intracellular accumulation results in a prolongation of the PD effects [86,95,107].

Nevertheless, the exact mechanism of action of MTX in RA is still unclear, and it also remains unknown whether MTX has a local effect in the joint as well as systemic effects.

1.3.2.2. Adverse effects

The adverse effects of MTX are mostly associated directly with its anti-proliferative mechanisms of action. It affects fast dividing cells such as hair follicles, mucosal cells of the GI tract, and the bone marrow, resulting in alopecia, mild to severe GI irritations, and myelosuppression, respectively [86]. In addition, adverse effects related to the central nervous system such as fatigue and loss of concentration can occur. Most severely, hepatotoxicity and pneumonitis may occur, albeit rarely [86,108]. Severe adverse effects require discontinuation of MTX therapy, while mild to moderate GI irritations after oral administration usually can be overcome by changing to parenteral administration [108].

1.3.3. Use of MTX in rheumatoid arthritis

Low-dose MTX (usually ≤25 mg once a week) exhibits anti-inflammatory activity and is used in the therapy of several autoimmune conditions, such as inflammatory bowel diseases and RA [86,109].

In RA, MTX is the gold standard disease modifying anti-rheumatic drug (DMARD) and is commonly given in doses of 10 – 25 mg once a week, either orally or via subcutaneous administration [108]. Rapid and effective disease control is desirable in RA to prevent irreversible joint damage [108]. However, the required dose is difficult to predict as it varies widely among patients.

In clinical practice, MTX treatment is usually started at a low dose (7.5 - 10 mg/week) which is subsequently escalated according to clinical response [110]. Starting MTX treatment at maximum dose (25 mg/week) is generally avoided as it may result in GI adverse effects which are poorly tolerated resulting in patients discontinuing MTX treatment [111].

Finding biomarkers that can be used to predict clinical outcomes, guide dose escalation and monitor MTX treatment is therefore of great clinical interest [108].

Note, that in this thesis the term "biomarker" will be used for any biological measure that has a predictive, but not necessarily a causative relationship with disease or treatment outcomes.

1.3.3.1. Monitoring of MTX treatment in RA

Most commonly, plasma concentration measurements such as steady state, peak or trough concentrations are used to monitor drug therapy. However, plasma concentrations of MTX are unsuitable for monitoring weekly, low-dose MTX treatment. MTX plasma concentrations fall below the limit of quantification (LOQ) within 24 hours after application and no steady state concentrations can be measured [112,113]. Currently, the literature provides no evidence for a correlation of clinical outcomes or adverse effects with alterative PK measures such as peak concentrations. In addition, the active metabolites MTXPGs contribute significantly to the activity of MTX, which may not be adequately reflected by measuring common biomarkers related to the parent drug only. Therefore, an alternative biomarker for monitoring of low-dose MTX treatment is required that correlates with clinical outcomes and/or adverse effects in RA and that can be measured easily.

1.3.3.1.1. RBC MTXPG concentrations

Potential candidates as alternative biomarkers to monitor MTX treatment are MTXPG concentrations measured inside RBCs. As in other cells, MTXPGs accumulate inside RBCs. The simple access to RBCs and their abundance make them easy candidates for measuring intracellular MTXPG concentrations.

However, RBC MTXPG concentrations have neither a known causal relationship with RA nor with the clinical outcomes of MTX treatment. RBCs are not involved in the pathology of RA and are also not located on the postulated pathway of action of MTX. Yet despite this lack of causality, RBC MTXPG concentrations could be used as a clinically reasonable biomarker to assess MTX treatment if there is a predictable correlation with the outcomes such as disease control and/or adverse effects. Such a correlation is debated in literature. Several studies have shown a correlation between disease control and RBC MTXPG concentrations [114-116], but these findings could not be verified in a more recent cross-sectional study [117]. A population PKPD model could prove useful to determine whether a predictable correlation between MTXPG concentrations in RBCs and clinical outcome of MTX therapy truly exists.

In the third part of this thesis, a population PK model for RBC MTX and MTXPGs is developed using a top-down modelling approach. Such a model provides a better understanding of the time course of MTXPG accumulation inside RBCs and will facilitate future work assessing the correlation between RBC MTXPGs and disease outcomes and/or adverse effects of low-dose MTX therapy in RA in a full population PKPD model.

PART II

RED BLOOD CELL SURVIVAL MODEL

Chapter 2: A statistical model for red blood cell survival time

This chapter is partially based on the following peer-reviewed publication:

Korell J, Coulter C, Duffull S (2011) *A statistical model for red blood cell survival*. Journal of Theoretical Biology 268(1):39-49.

2.1. Synopsis of the Chapter

In this chapter, the development of the proposed statistical model for the lifespan of red blood cells (RBCs) (introduced in the Introduction Section 1.2.4.3) is continued. Furthermore, sensitivity of the model to changes in the parameter values is explored and general (*a priori*) model identifiability is assessed using information theoretic principles. Finally, possible applications of this model to describe various pathological conditions are discussed.

2.2. Introduction

2.2.1. Previously proposed models for red blood cell survival

Lifespan based indirect response (LIDR) models have been described successfully for natural cell populations and used to simulate the influence of various agents on the production of different cell types [72,118], e.g. the stimulating effect of erythropoietin (EPO) on reticulocyte count [72]. In these compartment models, the loss of cells from a population is not controlled by a first-order process. Instead, the cells leave the compartments either due to death by senescence or by differentiation into another cell type [72]. The transition time between the compartments is assumed to be equal for all cells in the population corresponding to a point mass distribution of lifespans which follows the simple assumption of a fixed lifespan of all cells [72].

A more complex and arguably realistic approach is the implementation of a continuous lifespan distribution into the model. The influence of EPO on the reticulocyte population has been simulated by Krzyzanski *et al.* [75] using three different probability density functions (pdf): a gamma distribution, a lognormal distribution, and a pdf suggested by Dornhorst [69] that accounts for random destruction of the cells until they reach their finite lifespan and die due to senescence. Although more mechanistically focussed none of these lifespan distributions provided a better fit to the reticulocyte counts after administration of EPO than the fixed lifespan model.

Freise *et al.* [76] used a truncated three-parameter Weibull distribution to simulate a distribution of reticulocyte lifespans in sheep. Applying a time variant location parameter depending on the stimulation of RBC production, this distribution yields better results with respect to the underlying physiology than a time variant point mass distribution, which assumes a fixed lifespan for all cells stimulated at the same point in time.

Recently, Kalicki *et al.* [80] compared the goodness of fit of several transit compartment (TC) models to data of RBC survival. They found a 29-TC model including random destruction to be superior to simpler TC models without random destruction.

In contrast to these mostly parsimonious and application focused previous approaches, the underlying idea of the model proposed in the Introduction to this thesis (Section 1.2.4.3) was to obtain a mathematical description of the RBC lifespan distribution that is focused on the plausible processes of RBC destruction, with the intention of providing a new perspective on modelling the lifespan of RBCs, that might facilitate future work in this area.

2.2.2. Red blood cell destruction mechanisms

Four types of physiological destruction mechanisms of RBCs are plausible:

- early removal of unviable RBCs from the circulation
- age-independent random destruction
- premature removal of misshapen RBCs
- normal cell death due to senescence

These complex mechanisms involved in RBC destruction contradict the assumption of a fixed lifespan for all RBCs. The multitude of mechanisms can only be described by a lifespan distribution and cohort labelling studies provide evidence for such a distribution [54,119]. Yet, the true shape of this distribution remains largely elusive, and none of the so far proposed

distributions accounts for all these possible physiological mechanisms of RBC destruction.

2.3. Objectives

Although not all of these processes of RBC destruction might be present simultaneously in an individual at any point in time, the objective here was to postulate a continuous lifespan distribution that is specific for the survival time of RBCs and that is in accordance with the plausible processes of RBC destruction in humans. It is intended to prove useful in future to inform RBC survival studies in healthy volunteers as well as in patients with various types of haemolytic disorders as the extent to which each of these processes contribute to the overall lifespan distribution can be varied in this model, which will be discussed later in this chapter.

2.4. Model development

2.4.1. Previous version of the model

The underlying theory of the proposed model has been described in detail in the Introduction (Section 1.2.4.3). The model was implemented using the programming software MATLAB® (The MathsWorks™ Inc., Natick, USA).

In brief, a previously published survival model for humans [83] was rescaled to the time scale of RBC survival based on matching characteristics between human mortality and RBC destruction mechanisms (see Table 1.2 in Section 1.2.4.3.1 of the Introduction) [81,82]. The corresponding pdf over time t constitutes of a mixture of two Weibull distributions and is given by:

$$
\begin{aligned}
f_{MIX}(t) &= m \cdot f_{FW}(t) + (1-m) \cdot f_{RAW}(t) \\
&= m \cdot \left(exp\left(-exp\left(s_1 \cdot t - s_2 / t \right) - c \cdot t \right) \cdot \left(\left(s_1 + s_2 / t^2 \right) \cdot exp\left(s_1 \cdot t - s_2 / t \right) + c \right) \right) + \\
&\quad (1-m) \cdot exp\left(-(r_1 \cdot t)^{r_2} - (r_1 \cdot t)^{1/r_2} \right) \cdot \left((r_1 \cdot r_2)(r_1 \cdot t)^{r_2 - 1} + (r_1 / r_2)(r_1 \cdot t)^{(1/r_2)-1} \right)
\end{aligned}
$$

Equation 2.1: *Probability density function of the proposed RBC lifespan model.*

Here, *FW* stands for flexible Weibull which is a function of the parameters s_1, s_2 and c, while *RAW* stands for reduced additive Weibull and the corresponding function is described by r_1 and r_2. Both these distributions are combined as fractions based on the mixing parameter m.

As mentioned in the Introduction to this thesis (Section 1.2.4.3.2), the previous version of the proposed model suffered from several drawbacks: Firstly, it was based on a numerically intensive and time consuming code that followed each generated RBC individually over time using multiple loops in MATLAB®. Secondly, the previously published human parameter values were used and individual rescaling of each lifespan based on the ratio of median RBCs survival (assumed to be 115 days) to median human survival was required. Both these issues had to be resolved to render the model suitable for future work. To address the first issue, a more efficient code was implemented in MATLAB®, while the second issue was resolved by developing a specific model for the lifespan of RBCs. Both modifications are described in the following section.

2.4.2. Modifications of the model

2.4.2.1. Improvement of the run times

To overcome the first drawback of very slow run times, the MATLAB® code was improved using vectorisation. Here, RBCs are generated as entries in an array with the dimensions k_p (production rate per day) times t_{Bmax} (time until "birth" occurs). Therefore, the corresponding column represents the day of "birth" for each cell, i.e. the day of release from the bone marrow. Each RBC is then assigned an individual lifespan, randomly drawn with replacement from the lifespan distribution. Subsequently, for every day of a simulation it is

checked whether each individual cell is born and still alive, or not. Living cells per day are summed up and plotted *versus* time *t* according to:

$$N_{RBCs}(t) = \sum_{i=1}^{N} Z_{i,t}$$

$$\text{with} \quad N = k_p \cdot t_{B\max}$$

$$\text{where} \quad \begin{cases} Z = 1 & \text{if } t > t_{Bi} \text{ and } t < t_{Bi} + LS_i \\ Z = 0 & \text{else} \end{cases}$$

$$\text{and} \quad LS \sim \pi(s_1, s_2, c, r_1, r_2, m)$$

Equation 2.2: *Lifespan distribution model for RBC survival.*

Here, N is the total number of RBCs in the simulation, k_p denotes the constant production rate per day, t_{Bmax} is the time until "birth" occurs, t_{Bi} is the individual day of "birth" of the i^{th} RBC and LS_i is a random sample of an individual lifespan, which are distributed according to Equation 1.38 (π). In this context, the day of "birth" is assumed to be the day of release into the circulation. This implies that reticulocytes and mature RBCs in the circulation are not distinguished, thus reticulocytes represent the fraction of the youngest RBCs in this model.

2.4.2.2. Estimation of RBC specific parameters

The requirement of individual rescaling of each RBC lifespan was addressed by estimating the RBC specific parameter values for Equation 1.38. A rescaled sample of 1,770,887 pre-posterior RBC lifespans was obtained based on the human pdf using the rejection sampling method (Appendix A.1.1). A histogram containing 200 bins was created from these samples and normalised based on the total number of sampled lifespans. This normalisation results in an integral (from zero to infinity) of one for the normalised histogram and converts the units of the y-axis from absolute count to relative frequency, which is equivalent to a probability density.

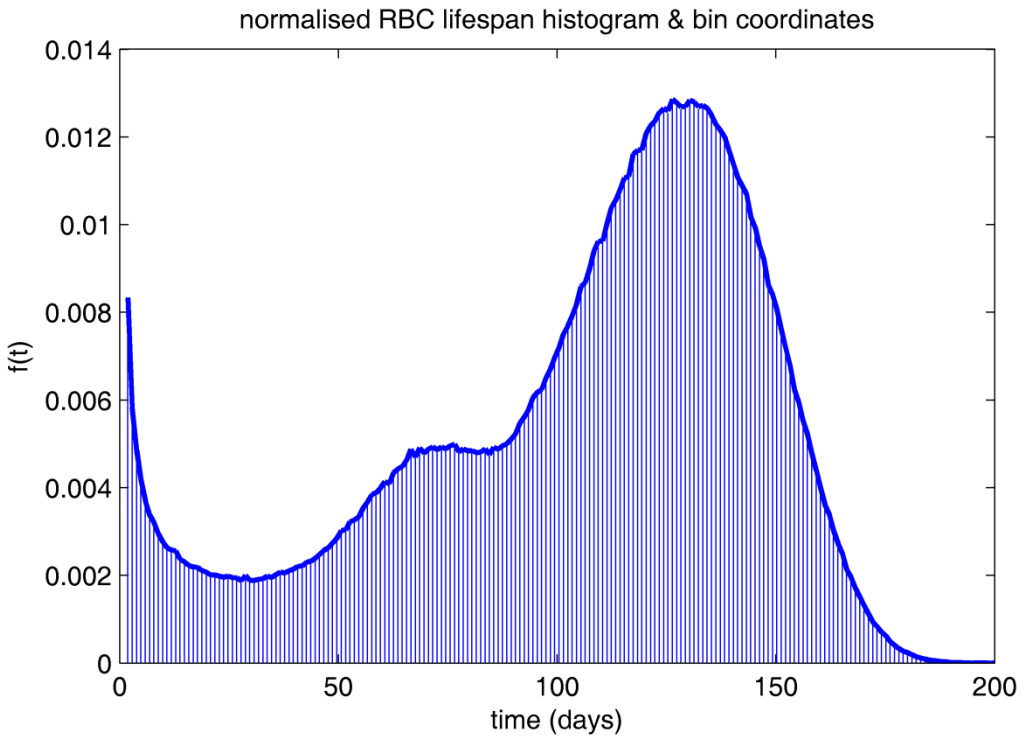


Figure 2.1: *Normalised RBC lifespan histogram* (*blue stems indicating the midpoints of each bin*) *and corresponding normalised coordinates* (*blue line*) *based on 200 bins. x-axis RBC lifespan in days, y-axis relative frequency = probability density f*(*t*).

Fitting the pdf given by Equation 1.38 to the coordinates of the mid points of the normalised bins allows estimating the RBC specific parameters for the model. The fitting was achieved by using a non-linear least square regression method implemented in MATLAB® as function *lsqcurvefit*.

However, when the parameter estimation was based on all 200 normalised coordinates, the fit to the main maximum of the function described by the flexible Weibull distribution was given more weight compared to the fit to the reduced additive Weibull distribution part that accounts for the saddle in the profile. This resulted in a significant difference between the profile of the estimated pdf and the profile of the normalised coordinates. An empirically determined set of parameter values was able to capture the coordinate profile much better without noticeable loss in the fit to the shape of the main maximum (Figure 2.2).

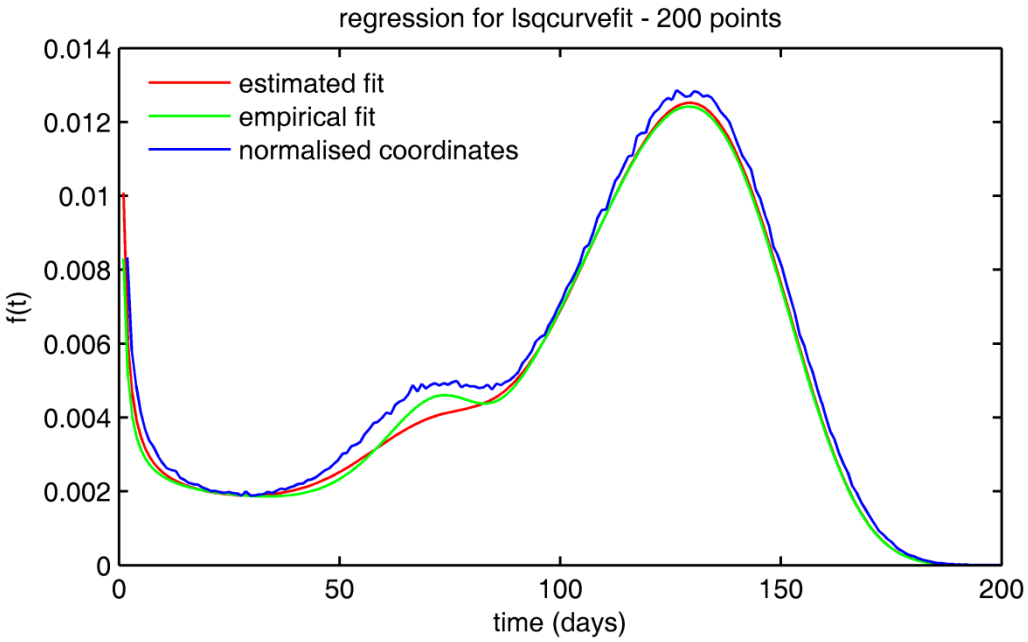


Figure 2.2: *Profile of the 200 normalised histogram coordinates* (*blue*) *compared with the empirically fitted pdf* (*green*) *and the estimated pdf* (*red*) *based on all 200 points.*

To improve the curve fitting procedure, the number of data points used for estimation was reduced to produce five data sets by removing every second coordinate from the original data set, resulting in data sets with 100, 50, 25, twelve and six normalised coordinates only. (Note that the pdf contains six parameters and therefore requires a minimum of six support points for parameter estimation.) Equation 1.38 was then fitted to each new data set. The difference between the original 200 normalised coordinates and the predictions based on each set of parameter estimates (based on the six data sets) was calculated and plotted against time (difference curve). By calculating the area between the difference curve and null (AUC) it was determined which set of parameter estimates provided the best fit to the original data. The fit was assessed based on the AUC value (the lower the better); however the ability to capture the profile of the original data in comparison to the empirical parameter set was also taken into account. The corresponding plots for all six data sets are given in Appendix A.2.1, while a summary of AUC values for the difference plots is given in Table 2.1.

Table 2.1: *AUC of difference plots based on each set of parameter estimates.*

Set of parameter estimates	AUC
empirically determined	0.0682
estimated based on 200 points	0.0644
estimated based on 100 points	0.0616
estimated based on 50 points	0.0600
estimated based on 25 points	0.0584
estimated based on 12 points	0.0588
estimated based on 6 points	0.0665

The data sets containing 25 and twelve normalised coordinates provided comparably low AUC values for the difference plots. However, the data set containing twelve support points was considered to be superior in capturing the original profile. The corresponding plots for this data set are shown in Figure 2.3. The parameter estimates obtained by fitting Equation 1.38 to this data set were used in the following as RBC specific parameter values and are given in Table 2.2.

Table 2.2: *RBC specific parameter values.*

Parameter	Estimated value	Units
s_1	0.0241	$days^{-1}$
s_2	440.78	days
r_1	0.0140	$days^{-1}$
r_2	8.9681	-
c	0.0024	$days^{-1}$
m	0.8941	-

Based on these parameter values a lifespan distribution specific for RBCs can be generated directly using the resampling method [84,85] described in Appendix A.1.1, without the requirement of rescaling.

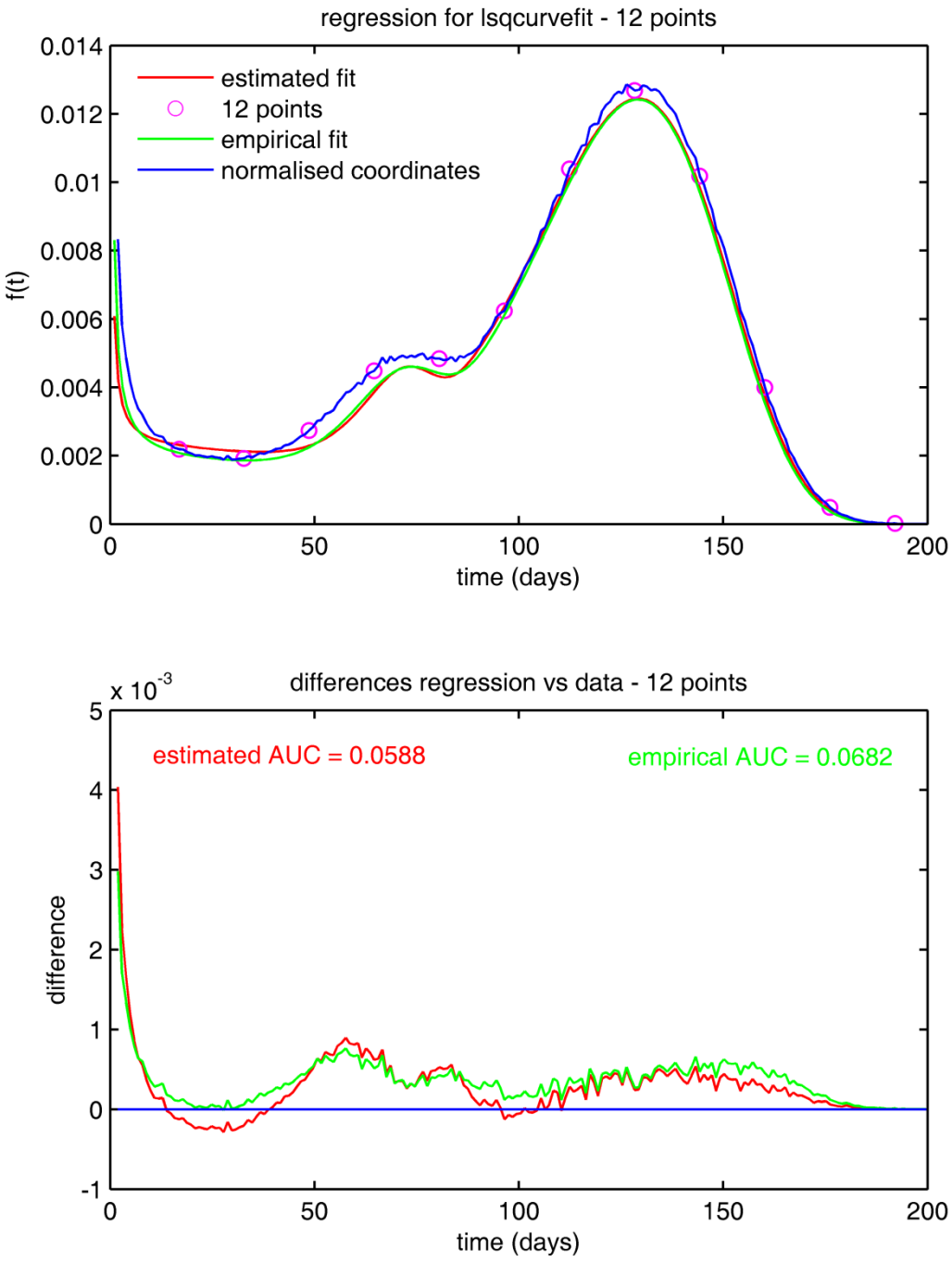


Figure 2.3: *Upper panel: Profile of the 200 normalised histogram coordinates (blue) compared with the empirically fitted pdf (green) and the estimated pdf (red) based on twelve points (pink). Lower panel: Difference plots for empirical and estimated pdf fits with area between difference curve and null (AUC).*

The final new MATLAB® code for this RBC specific model is given in Appendix A.2.2 and was used for the following simulations. Note that similar simulations were also carried out during the preliminary work [81], and equivalence of the new coding method was ensured by recreating these in this chapter.

2.5. Simulations

Three simulations were conducted using the enhanced code for the model: 1) a group of 1000 RBCs all born on day one was simulated; 2) a constant daily production rate of 1000 RBCs during the first ten days was assumed; 3) constant production of 1000 RBCs per day was assumed to occur over 500 days to achieve a steady state of the total number of RBCs during this simulation.

The assumption of a constant daily production rate is based upon the equivalent proposal of Schiodt [120]. This assumption was confirmed by comparing the variability in reticulocyte count over time as a useful metric to determine the production rate of RBCs [121]. Sandberg *et al.* [122] have shown that the variability within a subject is small in comparison to the between subject variability, which was confirmed by us [81,82] based on data of Seip [123]. Therefore, the production rate of RBCs in each individual can be assumed as constant in relation to the variability between different subjects.

2.5.1. Simulation 1

The simulation of a cohort of 1000 RBCs born on the first day results in an s-shaped disappearance curve. In Figure 2.4 the percentage of surviving RBCs is plotted *versus* time as dashed line. The early death of some of the RBCs accounts for the concave beginning of the disappearance curve. Afterwards, a small fraction of RBCs die due to random destruction and with increasing age the death rate increases due to senescence. The tailing at the end of the curve results from those cells that have been assigned a very long survival time much greater than the median.

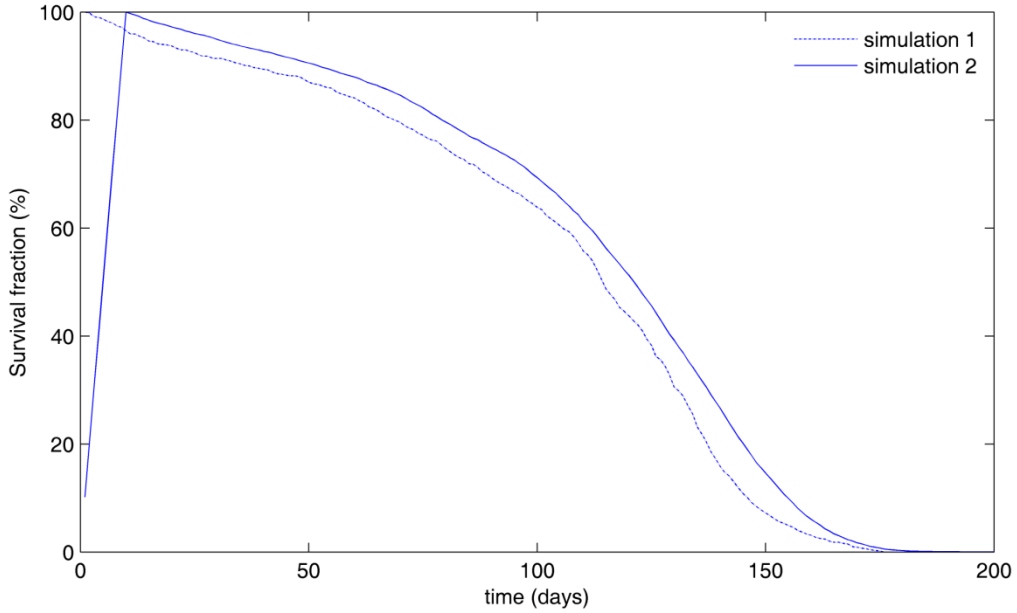


Figure 2.4: *Disappearance curve of the first simulation with 1000 RBCs born on day one* (···)*, and survival curve of the second simulation with a constant daily production rate of 1000 RBCs during the first ten days* (—). *The y-axis is scaled to percentage of RBC survival.*

It needs to be noted that the resampling method used for defining survival induces Monte Carlo error in the apparent survival fraction.

2.5.2. Simulation 2

A similar survival curve is observed simulating a constant production of 1000 RBCs per day during the first ten days. The solid line in Figure 2.4 represents the percentage of surviving RBCs in this simulation. The descending part of the curve shows the same s-shaped pattern with a less pronounced initial concave part and tailing at the end as the disappearance curve in the first simulation. Steady state is not achieved in this simulation due to the short time of production which is less than the typical age. The daily production rate during the first ten days is bigger than the destruction rate on these days.

2.5.3. Simulation 3

Assuming a constant daily production rate of 1000 RBCs over the first 500 days results in a survival curve with steady state occurring after approximately 180 days (Figure 2.5).

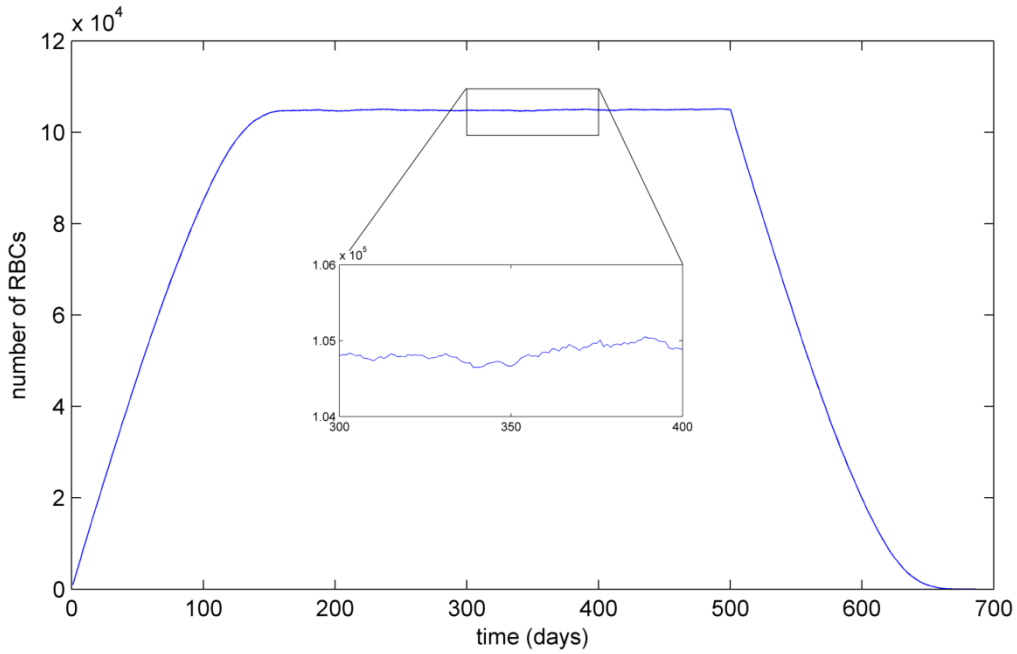


Figure 2.5: *Survival curve of the third simulation assuming constant production of 1000 RBCs per day to occur over 500 days. After approximately 180 days steady state is reached until production stops after 500 days. Insert: Fluctuation in steady state.*

The number of RBCs keeps growing until all RBCs born on the first day have died. At this time steady state is reached. A smooth transition from the ascending part of the curve into the steady state is observed. This is due to the fraction of long-lived RBCs with a survival time of approximately 180 days which were born on the first day. As a result of the short survival time of some cells, the total number of RBCs in the steady state is less than the product of production rate and the median lifespan. In contrast, assuming a fixed lifespan for all RBCs results in a total number of RBCs at steady state that is equal to the product of production rate and lifespan (data not shown). The number of RBCs in this simulation fluctuates around a value of approximately 105,000 RBCs (Figure 2.5 insert). This fluctuation equates to variations of the randomly sampled lifespans (Monte Carlo error) as each single RBC is followed during its

entire lifetime. The descending part of the curve has lost the s-shaped pattern observed during the first two simulations. An initial concave part of the disappearance curve is no longer observed visually. Instead, an almost linear decline with a sharp onset and tailing at the end is present.

2.6. *Model application*

Here, the results of the simulations are compared with previously published RBC survival studies to show that the newly proposed RBC lifespan model is applicable for the interpretation of RBC survival studies. Previously reported survival curves obtained using cohort labelling methods as well as random labelling methods are considered.

2.6.1. Cohort labelling studies

The first simulation represents a hypothetical cohort labelling method using an ideal label, where only cells born on the same day are labelled. The corresponding disappearance curve of the labelled cells is equivalent to the dotted line in Figure 2.4. Unfortunately, such a label for RBCs does not exist and this simulation can therefore only serve as a theoretical example.

The second simulation can be regarded as a simplified model of an *in vivo* survival study conducted by Shemin and Rittenberg [54]. These authors used glycine tagged with heavy nitrogen (^{15}N) as a cohort label for RBC which is incorporated into haemoglobin (Hb) during the production of RBCs in the bone marrow. A comparison of their data and the second simulation is shown in Figure 2.6.

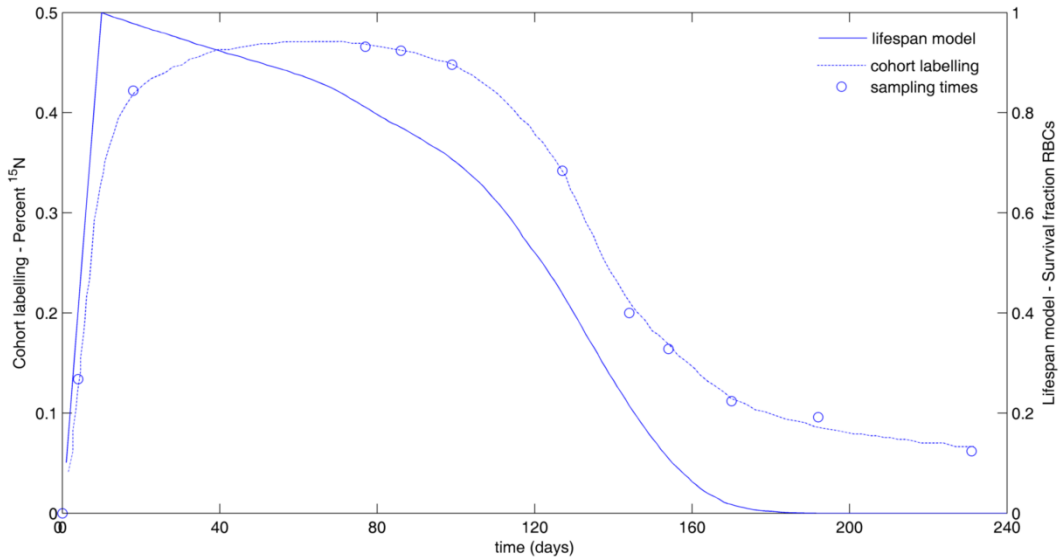


Figure 2.6: *Survival curves for cohort labelling methods. Survival curve obtained by Shemin & Rittenberg using ^{15}N (··· with ○ as the observed data). The graph was digitally extracted and recreated from [54]. Survival curve obtained in the second simulation* (—).

The constant production of RBCs over ten days assumed in this simulation equals to a prolonged incorporation of label during Hb synthesis. In both cases, an s-shaped disappearance curve is observed, although the initial concavity is not seen in the *in vivo* study. Furthermore, Shemin and Rittenberg almost observed a steady state with a smooth transition to a descending curve (Figure 2.6 dotted line), while in the simulation presented here, the early break down of some of the young RBCs results in an immediate start of the disappearance curve with a short initial concave part (Figure 2.6 solid line). There are two explanations for these differences: Firstly, the incorporation of the cohort label might not occur at a constant rate. It is likely, that a pool of labelled glycine is formed in the bone marrow. Subsequently, this labelled glycine is used for the synthesis of Hb during erythropoiesis, decreasing the pool of label over time. Therefore, the fraction of RBCs labelled during erythropoiesis would decrease over time as well. This can explain the plateau as well as the long presence of label in the circulation observed by Shemin and Rittenberg. In contrast, for simplification purposes it is assumed in the model that the production of labelled RBCs is constant over time. The second

explanation for the observed differences is a possible reuse of the label. A fraction of the released ^{15}N-tagged glycine might be used for the formation of new Hb after the breakdown of the initially labelled cells, whereas in the simulation the label is assumed to be ideal without showing any reuse. This would also explain that the *in vivo* survival curve does not return to zero, even after almost 240 days which is twice as long as the commonly accepted lifespan of RBCs.

Furthermore, an early disappearance of a small fraction of young RBCs and the presence of a small rate of random destruction cannot be observed in the study of Shemin and Rittenberg due to the irregular and long blood sampling intervals. Released label during the intervals from short-lived RBCs would already be reused when the next sample is taken. Otherwise, in the case of an even longer but reduced incorporation of the label during erythropoiesis, the continuous formation of tagged RBCs would obscure the loss of a small fraction of labelled cells. Hence, the lifespan distribution that was derived by Shemin and Rittenberg can only reflect those RBCs dying from senescence, whereas the distribution underlying the new model accounts for all plausible processes of RBC destruction. Here, each cell is followed individually in the model, resembling an almost ideal cohort labelling technique without reuse of the label and a constant daily production of labelled cells over a definite and known time period (ten days in the simulation).

An alternative model for cohort labelling studies of RBCs has recently been presented by Kalicki *et al.* [80]. Using previously published data of RBC survival [124], they estimated the parameter values of a TC model (Figure 2.7A) as mean lifespan LS = 107 days, k_{RD} = 0.00527 days^{-1}, and $k_{TR} = NC/LS$ (days^{-1}) with NC = number of compartments [personal communication]. To simulate the fractional survival of a RBC cohort the initial number of RBCs in the first compartment is set to 1, and it is assumed that k_{in} = 0 days^{-1} to stop any further production. The resulting graph of RBC survival (Figure 2.7B dotted line) shows a similar pattern as the first simulation using the newly proposed lifespan distribution model (Figure 2.7B solid line). However, due to the

empirically chosen median RBC age of 115 days in the lifespan distribution model compared to the shorter mean RBC lifespan of 107 days and the higher rate of random destruction in the TC model, the overall survival of the RBC cohort is longer using the lifespan distribution model. This is reflected in Figure 2.7B by the right shift of the survival curve of the proposed lifespan distribution model compared to the curve of the TC model. Additionally, the TC model does not account for the selective early break down of younger RBC whereas with the new model early death can be seen as an initial concave part of the survival curve.

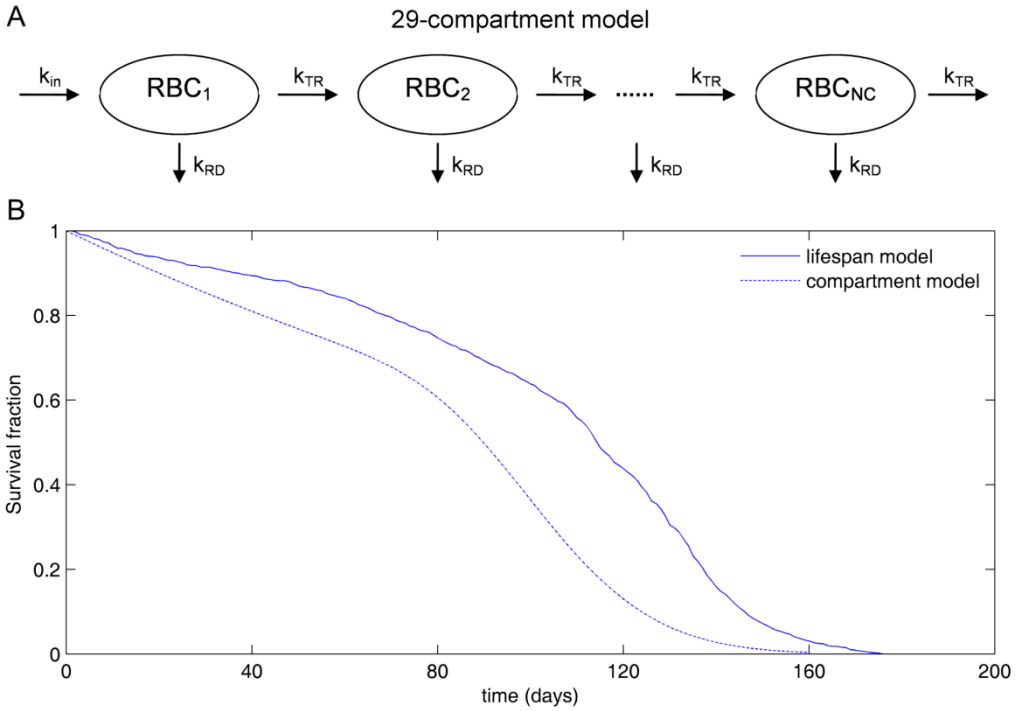


***Figure 2.7:* A** *Compartment model with random destruction and variability of RBC lifespan as proposed by Kalicki et al. [80]. NC = Number of compartments, k_{in} = rate constant of RBC production, k_{RD} = rate constant of random destruction, k_{TR} = rate constant of transit and LS = mean RBC lifespan.* **B** *(····) Survival fraction of an RBC cohort (RBC_1 = 1) simulated with the transit compartment model for NC = 29, LS = 107 days, k_{TR} = NC/LS ($days^{-1}$), k_{in} = 0 $days^{-1}$, k_{RD} = 0.00527 $days^{-1}$. (—) Survival fraction of an RBC cohort simulated with the newly proposed lifespan distribution model assuming a median RBC age of 115 days.*

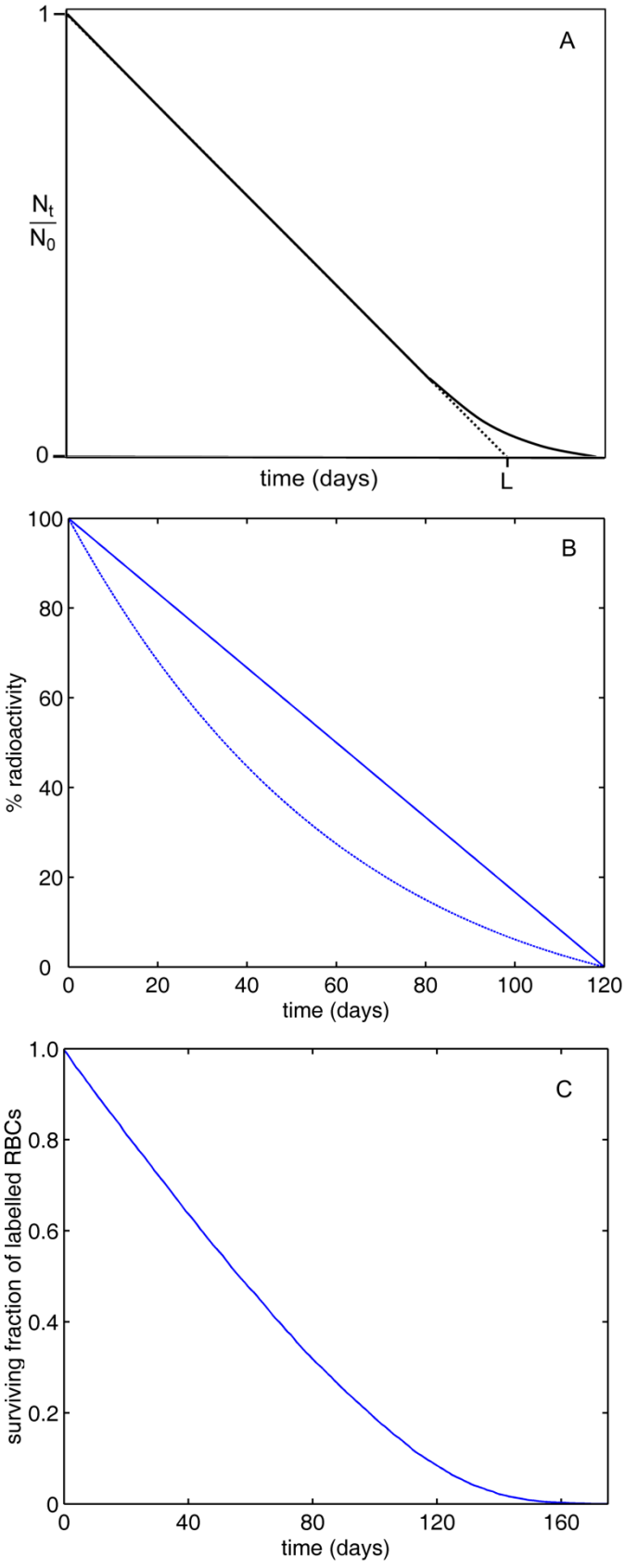
2.6.2. Random labelling studies

In a comprehensive review on the analytical interpretation of RBC survival studies Dornhorst illustrated the effect of normally distributed lifespans on the survival curve of RBCs in healthy individuals where no random destruction and early death of RBCs were assumed to occur [69]. Applying this distribution, the author predicted a straight line with tailing at the end as the disappearance curve for a random labelling technique such as Ashby's agglutination technique [55] (Figure 2.8A). The descending part of the survival curve obtained in the third simulation equates the predicted disappearance of cells assuming labelling to occur on day 500 (Figure 2.8C).

The International Committee of Standardization in Hematology (ICSH) suggests a reference method for conducting RBC survival studies using radioactive chromium (^{51}Cr) [66]. This method should yield a straight line for the disappearance of label from circulation as well; provided that the data is corrected for random loss of the label from the cells and that death of RBCs occurs largely due to senescence (Figure 2.8B). The disappearance curve observed using this method equals the descending part of the survival curve in the third simulation (Figure 2.8C), although the model also accounts for random destruction and early disappearance of a small fraction of cells. The small rate of random loss present in healthy individuals is neglected in the interpretation of the ICSH for the sake of simplification as it is difficult to estimate with the labelling methods currently available. The same holds true for early loss of RBCs.

following page:

Figure 2.8: *Survival curves obtained using random labelling methods.* **A** *Survival curve as predicted by Dornhorst assuming a normal distribution of RBC lifespans with L = mean lifespan. Digitally recreated from [69].* **B** *Disappearance curve proposed for ^{51}Cr labelled cells, uncorrected* (····) *and corrected for random loss* (—). **C** *Disappearance curve obtained in the third simulation using the lifespan model.*


A
1
0
$\frac{N_t}{N_0}$
time (days)
L
B
100
80
60
40
20
0
% radioactivity
0
20
40
60
80
100
120
time (days)
C
1.0
0.8
0.6
0.4
0.2
0
surviving fraction of labelled RBCs
0
40
80
120
160
time (days)

2.7. Ability of parameter estimation

The aim of this section is to investigate whether the parameters of the survival function corresponding to the RBC pdf could be estimated from an *in vivo* RBC survival study and at what times blood samples should be taken to measure RBC counts. The blood sampling times were determined by applying a local D-optimal design criterion [29]. A description of D-optimal design is given in the Introduction to this thesis (Section 1.1.5). The determinant of the Fisher Information matrix was maximised using an exchange algorithm implemented in MATLAB® (Appendix A.2.3).

A hypothetical *in vivo* study with 100 patients using an ideal radioactive RBC cohort label without reuse of the label was assumed. Measurement noise was included by assuming an additive error of 36 counts per minute. This value equals to 12% of the average number of counts per minute during the experiment, which was set to 300. Optimal blood sampling times were determined to be on day 4, 64, 75, 87, 120 and 148 after labelling (Figure 2.9).

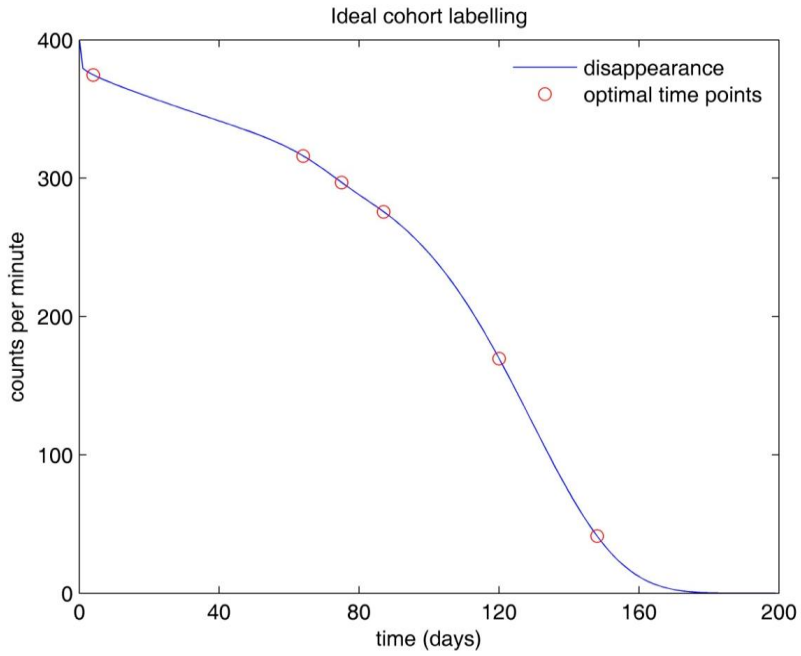


Figure 2.9: *Disappearance of an ideal cohort label with optimal time points for blood sampling at day 4, 64, 75, 87, 120 and 148 after labelling.*

The percentage standard error for the parameter estimates given as the diagonal elements of the inverse Fisher Information matrix are shown in Table 2.3. All six parameter estimates, except r_2, have a standard error of <10% and, therefore, could be estimated from this hypothetical *in vivo* study with good precision. The imprecision of the estimate of r_2 was slightly larger at 33%.

Table 2.3: *Percentage standard error (SE) for the parameter estimates from a hypothetical in vivo RBC survival study with 100 patients using an ideal cohort label.*

Parameter	r_1	r_2	s_1	s_2	c	m
SE (%)	4.25	33.07	2.39	2.81	8.30	1.30

2.8. Influence of parameter changes

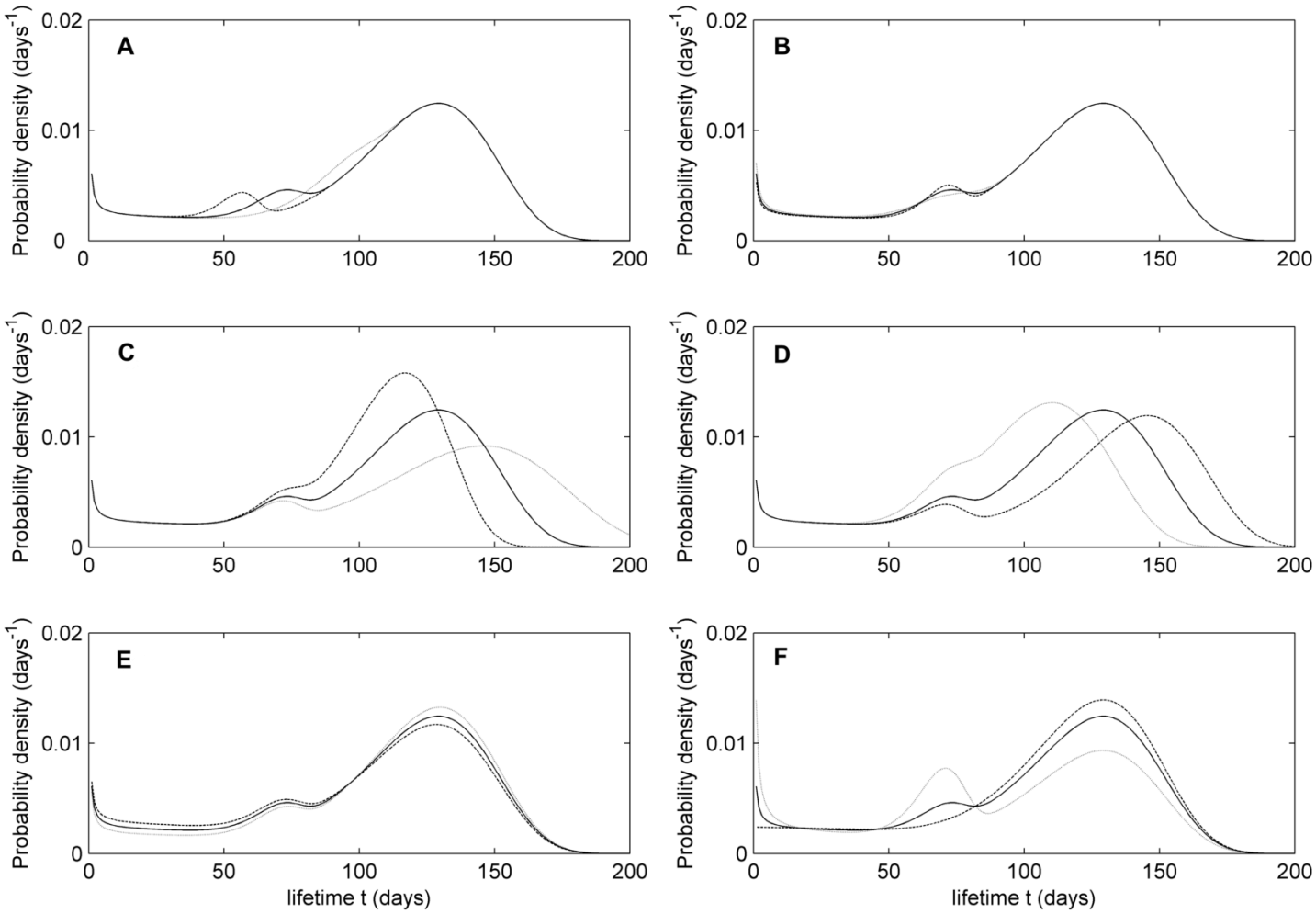
The influence of changes in the RBC specific parameters on the pdf is shown in Figure 2.10.

The location of the maximum of the reduced additive Weibull distribution (equals the saddle of the mixed pdf shown in Figure 2.10), which accounts for the reduced lifespan of misshapen RBCs, is controlled by parameter r_1 (Figure 2.10A). Increasing the value of r_1 results in a decreased lifespan of misshapen RBCs.

Parameter r_2 controls the shape of the reduced additive Weibull distribution (Figure 2.10B). Increasing the value of r_2 results in an increasing number of misshapen RBCs with reduced lifespan without having an effect on the location of the maximum of the reduced additive Weibull distribution.

following page:

Figure 2.10: *Influence of parameter changes on the shape of the RBC pdf. (—) parameter values estimated for RBC pdf. (····) +25% change (respectively m = 1), (—) -25% change in the corresponding parameter value. A detailed description is given in the text on the following pages.*


A
B
C
D
E
F
Probability density (days^{-1})
0.02
0.01
0
0
50
100
150
200
lifetime t (days)

The parameters s_1 and s_2 control the location of the maximum of the flexible Weibull distribution (equals the maximum of the mixed pdf), while s_1 also has influence on the location of the probability region associated with this density of the distribution. Increasing the value of s_1 shifts the maximum to lower values, therefore decreasing the lifespan of normal RBCs dying due to senescence, while the shape of the distribution is additionally narrowed (Figure 2.10C). On the other hand, increasing the value of the true location parameter s_2 increases the lifespan of normal RBCs without an effect on the location of the central density of the distribution (Figure 2.10D). Therefore, s_2 is of particular interest as it can be described as a function depending on environmental factors or drug dosages, which will then allow the estimation of a reduced or prolonged RBC lifespan under the influence of these factors or drugs within any candidate model.

The constant hazard rate accounting for random destruction of RBCs is solely controlled by c. Increasing the value of c increases the number of RBCs dying of random destruction and therefore reduces the number of RBCs dying due to senescence as well (Figure 2.10E).

Finally, the mixing parameter m controls the extent to which each of the underlying distributions contributes to the mixed model. Increasing the influence of the reduced additive Weibull distribution by decreasing the value of m results in a higher early death and number of misshapen RBCs with a reduced lifespan and decreases the number of normal RBCs dying due to senescence (Figure 2.10F). *Vice versa*, increasing m increases the influence of the flexible Weibull distribution, therefore increasing the number of RBCs dying due to senescence and reducing early death and the number of misshapen RBCs. For $m = 1$ (dashed line in Figure 2.10F) the reduced additive Weibull distribution is completely removed from Equation 1.38, leaving only random destruction and death due to senescence.

Future work using this model might include expressing these parameter values as functions of physiological factors, like the production of EPO or various pathological states. This allows changing the extent to which each

process of RBC destruction is represented in the lifespan distribution, and provides an opportunity to adapt the model according to actual clinical situations.

2.9. *Possible application in pathological states*

This section gives three examples to demonstrate the possible application of the model to pathological states that involve altered conditions of RBC destruction: stress erythropoiesis, sickle cell anaemia, and immune-mediated haemolytic anaemia.

2.9.1. Stress erythropoiesis

During stress erythropoiesis RBC production is increased as a result of increased EPO production under hypoxic conditions. There is evidence that the survival of RBCs produced under acute stress conditions is decreased [37].

Both, acute and chronic stress erythropoiesis can be accommodated in the model: The increased RBC production can be described by an increased production rate while the reduced RBC survival in the acute state could be expressed dependent on the underlying mechanism as a shift of the senescence component (maximum of the flexible Weibull distribution) to the left, an increase in random destruction or an alteration in the density associated with misshapen RBCs. These alterations in the model parameter values would either occur over a defined period of time in the case of an acute insult, or be permanent in the presence of chronic conditions.

2.9.2. Sickle cell anaemia

Sickle cell anaemia is characterized by abnormally shaped, rigid RBCs with a shortened survival in comparison to normal shaped cells. The reason is a mutation in the Hb gene resulting in faulty Hb which aggregates in the absence of oxygen in the venous system. This polymerisation leads to a deformation of the RBCs which over time becomes irreversible. The misshapen RBCs cannot pass through narrow capillaries due to their rigidity and increased intra- and

extravascular haemolysis occurs [34]. These processes lead to an increased random destruction of RBCs and a significantly shortened survival of the RBCs in comparison to healthy normal subjects [125]. Although heterozygous carriers of the sickle cell gene are usually assumed to have normal RBC morphology, some indication is given that these patients show a variable fraction of misshapen, sickle-like RBCs [126]. As these cells show morphological features similar to the fully sickled cells in homozygous carriers, they are likely to have a shortened survival in comparison to normal shaped cells.

It is therefore possible to include the homozygous and heterozygous form of the sickle cell disease in the model as follows: The RBC lifespan distribution is assumed to lack the senescence component in homozygous carriers. Instead all cells die at early times either due to a drastically increased random destruction or due to early splenic sequestration as a result of their deformation. The early destruction can be expressed mathematically by a dramatic reduction in the density of the main maxima of the pdf as almost no cells die due to normal senescence while the saddle accounting for early death of misshapen RBCs will increase as virtually all cells are misshapen. Random destruction governed by parameter c will additionally be elevated in comparison to normal conditions. In contrast, in heterozygous carriers of the sickle cell gene the pdf is suspected to express a shape similar to the one presented here including the saddle for the early destruction of the misshapen RBCs in addition to the normal death of RBCs due to senescence. The location and size of the saddle would be dependent on how severely the survival of the deformed cells is affected.

2.9.3. Immune-mediated haemolytic anaemia

Immune-mediated haemolytic anaemia is characterized by the formation of autologous antibodies against RBC membrane proteins [34]. RBCs with bound antibodies are subsequently removed from the circulation by the immune system and in the spleen and liver. This type of haemolytic anaemia can therefore be regarded as a pathological condition with an increased

random destruction of RBCs as there is no indication of an age-dependent factor. The model is able to account for this type of altered RBC destruction by an increase of parameter *c*, while the density of the senescence component of the pdf would be decreased or even missing dependent on the severity of the anaemia and therefore the random destruction.

2.10.Discussion

2.10.1. Model development

The purpose of this model was to describe RBC survival using a lifespan distribution that accounts for the possible physiological mechanisms of RBC destruction. Since there is no reliable information available on the actual distribution of RBC lifespans in humans, it was necessary to empirically choose a pdf for this purpose. The chosen distribution accounts for all potential hazard types resulting in the death of RBCs: senescence, random destruction, death due to initial or delayed failures. This is the first time that early death is incorporated into a model of RBC survival. It accounts for unviable RBCs that are destroyed shortly after their release, and also for neocytolysis, the selective destruction of the youngest RBCs [41].

The previous version of this model relied on rescaling of human lifespans and was implemented in MATLAB® using a time consuming code. Here, these drawbacks were eliminated by vectorising the MATLAB® code and estimating the RBC specific parameter values for the pdf from a normalised histogram of the rescaled RBC lifespan distribution. Based on these RBC specific parameters, a lifespan distribution for RBCs can be generated directly without the need for rescaling. Note that since the inverse cumulative density function for the pdf is not available in closed form a sampling based algorithm is required for generating lifespans. For the simulations with the model, each RBC is assigned a discrete lifespan, randomly drawn with replacement from this RBC lifespan distribution. Each individual RBC is then followed during its entire lifespan.

This results in a more comprehensive model of RBC survival than has previously been developed.

In contrast to other work, where reticulocytes and RBC production in the bone marrow are an essential part of the model, this work focuses on the disappearance of RBCs from the circulation. The day of "birth" in the model is assumed to be the day of the release of a cell from the bone marrow into the circulation. Reticulocytes and mature RBCs were intentionally not distinguished in this context. It is unlikely that there is a difference between both types of red cells in terms of labelling. Cohort labels are incorporated during erythropoiesis in the bone marrow and, therefore, are already present in the cells at the time of release into circulation, whereas the uptake of a random label into reticulocytes should not deviate largely from the uptake into mature RBCs. Therefore, reticulocytes are assumed to represent the fraction of the youngest RBCs in the model.

2.10.2. Simulations

Adequacy of the new version of the model compared to the old code was ensured by recreating three simulations that had already been part of previous work [81]. The results obtained with the new code are equivalent to the earlier ones for all three simulations.

The assumption of a constant daily production rate in the second and third simulation is based on the lower within subject variability of the reticulocyte count compared to the between subject variability, as the reticulocyte count is a useful method to determine erythropoiesis activity [121]. Nevertheless, this assumption holds only as long as normal steady state conditions are present. Any alteration of the erythropoietic feedback circle combining oxygen partial pressure, EPO production and erythropoiesis will ultimately result in a change of the production rate of RBCs.

For the simulations presented here, the value for the production was set to 1000 RBCs per day as there is no need for the true value in the presented

simulations. It should be noted that the value of 1000 is arbitrary and does not affect the generalisability of the model.

2.10.3. Model application

Comparison of the newly developed model to previously observed or proposed RBC survival curves demonstrate the applicability of the model for simulating the survival time of RBCs. Importantly, the model is able to describe survival curves similar to those observed using cohort labelling methods, as well as those proposed for random labelling of RBCs. Therefore, it is possible to use the model for the interpretation of future RBC survival studies.

The first simulation equals the ideal case of labelling a cohort of RBCs all born on the same day. However, this is unlikely to occur in practice as cohort labels are normally incorporated into Hb during erythropoiesis over several days. Therefore, the second simulation is a more realistic model for cohort labelling studies, as RBC production is assumed to occur over several days in this simulation. However, the incorporation of a label during erythropoiesis might not be constant, which is assumed for simplification in the current model. Furthermore, the predicted disappearance curve of the simulation does not account for a possible reuse of the label. In both cases it would be straightforward to incorporate corresponding correcting factors into the model to simulate the survival curves actually observed using these methods. This will be explored in the next chapter of this thesis.

The third simulation can be regarded as a model for random labelling studies. Regarding Ashby's historical method of differential agglutination, transfusing donor RBCs into a recipient is similar to stopping the production and merely observing their following disappearance. In the third simulation the transfused sample equals a fraction of RBCs taken on day 500 containing cells of all ages. Regarding other random labelling methods, the labelled RBCs parallel a blood sample taken on day 500, which is reinjected into the donor's circulation after labelling on the same day. However, this simulation equates labelling of RBCs with an ideal label that is not lost from the cell during

circulation and does not decay. It would be necessary to extend the model for simulations of the actually observed survival curves. For example, to simulate the observed survival curves in ^{51}Cr studies, it is necessary to include factors that account for the radioactive decay, the dissociation of the chromium-haemoglobin complex, and the loss of label due to vesiculation of Hb. Again, the next chapter of this thesis will address these required modifications.

2.10.4. Ability of parameter estimation

It was shown that all six parameters of the corresponding survival function could be estimated with good precision from a hypothetical *in vivo* RBC survival study with 100 patients using an ideal RBC cohort label. However, this is based on the assumption that the label is only incorporated into RBCs born on the same day and that no reuse of the label is present. Further work is required to evaluate the ability of parameter estimation from *in vivo* studies using currently available labelling methods. It is thereby necessary to account for the specific shortcomings that are associated with any given label.

In addition, the local D-optimal design applied here assumes the values of the parameters in the model are known exactly. Yet, this would not be the case when estimating the parameters from a RBC survival study. Parameter uncertainty should therefore be taken into account when designing RBC survival studies using this model, e.g. by applying a robust D-optimal design criterion.

2.10.5. Influence of parameter changes

Given the flexibility in the parameter set, adjustments of the model are possible to account for various haemolytic disorders by modelling the data and allowing one or more parameter to be estimated. Expressing the parameters as functions dependent on physiological factors could be used to control the influence each of the destruction processes has in an individual, as these processes might not be equally present simultaneously in each individual.

Additionally, drug effects can be included into the model by describing specific parameters as a function of the dose of a drug.

The location parameter s_2 is of particular interest regarding these possible variations, as it controls the location of the maximum of the pdf without having influence on the density of the function. For example, haemolytic disorders which lead to a shortened survival of normal RBCs can be expressed by incorporating a model for s_2 which results in a reduction in the value of the location parameter s_2 depending on the severity of the disease. Another possible scenario would be an acute stress erythropoiesis, where the production of RBCs is increases over a short period of time due to elevated endogenous EPO production or the administration of exogenous EPO. The resulting EPO dependent decrease of the lifespan of normal RBCs can again be expressed as a function of the location parameter s_2 depending on the dose of EPO in the model.

Other parameters of interest for incorporation of possible environmental effects are the random hazard rate c, controlling the rate of random destruction, and the mixing parameter m. Increasing the number of misshapen RBCs with reduced lifespan and increasing early death can be achieved by decreasing the mixing parameter m resulting in an increasing influence of the reduced additive Weibull distribution on the combined pdf. Furthermore, by fixing m to 1 Equation 1.38 simplifies to the flexible Weibull distribution. In this case, the lifespan distribution only accounts for random loss and death due to senescence controlled by the parameters s_1, s_2 and c. This option might be used for the sake of parsimony and can serve as a good approximation in healthy individuals, where the amount of early death and misshapen RBCs seem to be fairly small.

2.10.6. Possible application in pathological states

The applicability of the model in pathological states involving altered conditions of RBC destruction has been described using three examples: stress erythropoiesis, sickle cell anaemia and immune-mediated haemolytic anaemia.

Due to the flexibility of the model, all three states are assumed to result in certain alterations of the underlying RBC lifespan distribution which would be described by changes in the parameter values. This might provide a deeper insight in the pathophysiology of these diseases when applied to RBC survival data in patients suffering from any of these or similar conditions. However, this needs to be proven in future studies.

2.11. Conclusion

The newly developed model is able to simulate the survival time of RBCs and yields survival curves similar to those previously published for RBC survival studies using different labelling methods.

The underlying pdf of RBC survival time not only accounts for age-dependent removal from circulation due to senescence, but also for random destruction during the circulation in the body, for the early destruction of unviable RBCs, a shortened lifespan of misshapen cells, and for the mechanism of neocytolysis. Thus, the model reflects the known physiology of RBC destruction.

All parameter values controlling the survival of RBCs in this model could be estimated from a hypothetical *in vivo* RBC survival study using an ideal labelling method. However, future work is required to assess the ability of parameter estimation from *in vivo* survival studies using existing labelling methods that are inherently flawed. This will be assessed in the following chapter of this thesis.

Furthermore, the application of the model for various haemolytic disorders is expected to result in alterations of specific parameter values, which will provide deeper insight into the pathology of these diseases. Drug effects can be included into the model as dose-dependent functions of specific parameter values.

Chapter 3: Evaluation of red blood cell labelling methods

This chapter is based on the following peer-reviewed publication:

Korell J, Coulter C, Duffull S (2011) *Evaluation of red blood cell labelling methods based on a statistical model for red blood cell survival*. Journal of Theoretical Biology 291(0):88-98.

3.1. Synopsis of the Chapter

Here, the model developed in the previous chapter is enhanced to describe currently available red blood cell (RBC) labelling methods. It is furthermore assessed whether these methods are informative enough to support full parameter estimation of the model using a robust optimal design approach.

3.2. Introduction

An accurate estimation of the lifespan of RBCs is desirable. This would be particularly valuable for predicting the time to reach new steady state values of haematocrit in patients undergoing treatment with erythropoietin [127]. In addition, many diseases utilise measures that relate to RBC lifespan to provide an estimate of disease activity, for example, glycated haemoglobin (HbA_{1c}) as a biomarker for glycaemic control [128].

3.2.1. Estimation of the RBC lifespan using labelling methods

Estimation of the lifespan of RBCs is possible by using a label that stays with the cells during their circulation in the body. In general, two types of labelling methods have been developed: cohort labelling, where cells of a certain age are labelled, and random labelling, where all cells present at one moment in time are labelled irrespectively of their age (equivalent to population labelling). However, all current labelling methods contain significant flaws, resulting in potentially inaccurate estimates of the RBC lifespan.

A cohort labelling method such as heavy nitrogen (^{15}N), or radioactive iron (^{59}Fe), incorporates the label into RBC precursor cells during their production in the bone marrow, usually as a substrate used in haemoglobin (Hb) synthesis. Subsequently, these labelled precursor cells develop into reticulocytes and are released into the circulation where they mature into RBCs. However, the incorporation of the label into the precursors does not occur

instantly after administration of the label and may be prolonged over several days. Therefore, the resulting labelled RBCs are not exactly of the same age [38]. In addition, cohort labelling methods suffer from reuse of the label in the production of further RBCs [36]. An advantage of cohort labelling methods is that they do not require *ex vivo* manipulation of the cells. In contrast, the labelling process using random labelling methods is in general conducted *ex vivo* which can damage the cells and thus might affect their subsequent survival in the circulation. These methods also appear more sensitive to loss of the label via dissociation of the label from the viable RBC as well as radioactive decay. This is particularly notable for labelling with radioactive chromium (^{51}Cr) [57,129]. Both labelling methods, cohort as well as random, suffer from loss of label due to vesiculation of Hb from viable RBCs [68].

Other methods, such as carbon monoxide (CO) exhalation, have been developed to determine RBC turnover without depending on a label [62,63,65,130,131]. While CO exhalation avoids reuse and label loss, except vesiculation, it is confounded by environmental issues and only gives a measure of RBC turnover at a certain time point. It does not allow following the cells during their lifetime. Therefore, CO exhalation is not considered in this work.

Due to the flaws inherent in all present labelling techniques an ideal labelling method is currently not available and the mean lifespan and the distribution of lifespans of RBCs remain elusive. To obtain an accurate estimation of the RBCs lifespan it would either be necessary to overcome these flaws in the development of new labels, or to account for them during the calculation of the RBC lifespan from the data obtained. The former approach seems to be unlikely in the near future. However, modelling RBC survival data using methods that not only can account for the flaws inherent in various RBC labelling techniques but also differentiate between RBC destruction processes might prove fruitful to overcome the shortfalls associated with the current labels.

3.2.2. Previously developed RBC lifespan model

In the previous chapter, a statistical model for the survival time of RBCs with respect to the physiology of RBC destruction was developed. The underlying distribution of RBC lifespans accounts for known processes of RBC destruction in the body, including death due to senescence, random loss during circulation, as well as death due to early or delayed failure. The resulting survival model was used to simulate *in vivo* RBC survival studies using different ideal RBC labelling methods. Predictions from the model agreed well with models from the literature for both, cohort and random labelling methods. However, the model did not account for the above mentioned flaws associated with real labels for RBCs.

3.3. Objectives

The overall aim of this chapter was to compare the performance of commonly used labelling methods for RBCs. This was carried out by addressing the following three specific objectives: 1) To include known flaws associated with ^{51}Cr and ^{15}N into the RBC survival model. 2) To assess the theoretical identifiability of the model, i.e. the ability to estimate all parameters in the model, based on ideal cohort and random labelling experiments. 3) To evaluate whether the parameter values can be estimated from non-ideal labelling studies using ^{51}Cr or ^{15}N, and to compare the performance of both labelling methods in terms of the precision of parameter estimation.

It is important to assess the theoretical behaviour of the model so that any future experiments can distinguish issues with model performance (i.e. structural identifiability issues) from issues associated with the particular data set to be analysed (i.e. deterministic identifiability issues).

3.4. Methods

3.4.1. RBC survival model

A statistical model for the survival time of RBCs was implemented in MATLAB® (The MathWorks™ Inc., Natick USA) based on the work of Bebbington *et al*. [83] and adapted to describe RBC survival in the previous chapter. This RBC survival model was enhanced in this chapter to provide the survival function $S(t)$ (Equation 3.1) that is mathematically equivalent to the RBC lifespan distribution used earlier (see Introduction Section 1.1.3 and Appendix A.3.1).

The survival of a single cohort of RBCs born on the same day to time t is described by the following equation

$$S(t) = m \cdot \left(\exp\left(-\exp\left(s_1 \cdot t - s_2/t\right) - c \cdot t\right)\right) + (1-m) \cdot \left(\exp\left(-\left(r_1 \cdot t\right)^{r_2} - \left(r_1 \cdot t\right)^{1/r_2}\right)\right)$$

Equation 3.1: *Survival function for a cohort of RBCs born on day t = 0.*

where s_1 and s_2 control normal death of RBCs due to old age (senescence), r_1 and r_2 describe a reduced survival of faultily produced RBCs resulting in early or delayed failure, c controls random destruction irrespectively of age, and m is a mixing parameter combining both parts of the function. Figure 3.1 shows the underlying RBC lifespan distribution and indicates which parts of the distribution are controlled by these parameters.

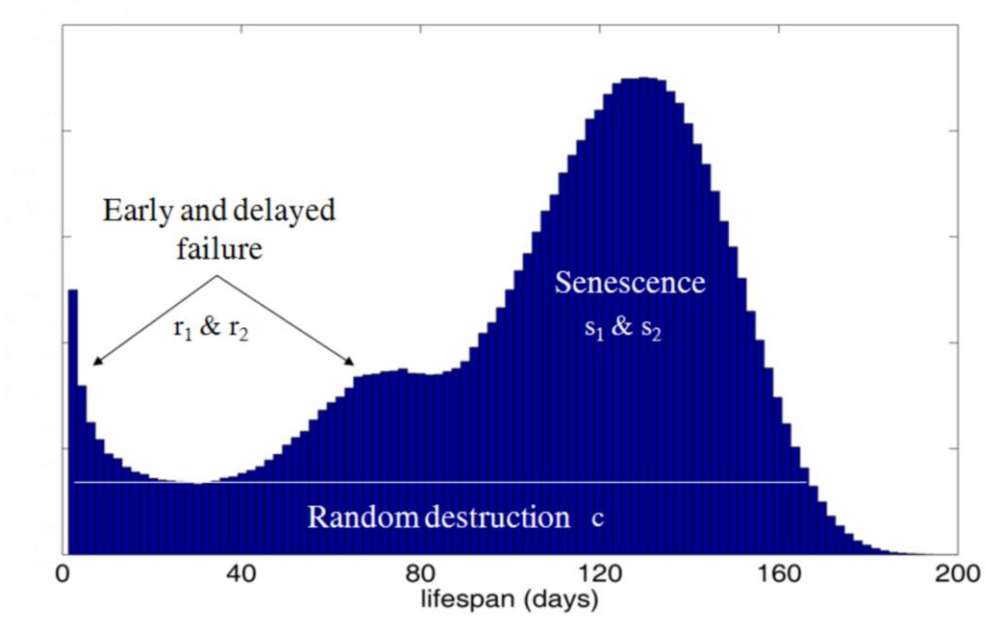


Figure 3.1: *Underlying RBC lifespan distribution that accounts for four plausible physiological mechanisms of RBC destruction. Parameters in the survival model controlling these mechanisms are indicated.*

Previously, values for these parameters were estimated by assuming a median RBC lifespan of 115 days (Table 3.1). The assumption of 115 days is for convenience and it is unnecessary to apply this assumption when modelling RBC survival data.

Correspondingly, the mean lifespan $\overline{L}$ based on these parameter values is 100 days, calculated according to [132]:

$$\overline{L} = \int_0^\infty S(t)\,dt$$

Equation 3.2: *Mean RBC lifespan.*

It is assumed that the day of birth is equal to the day of release into the circulation. This implies that the model does not require a distinction between reticulocytes and mature RBCs. Here it is assumed that reticulocytes represent the fraction of the youngest RBCs in the model and their survival along with mature RBCs can be estimated.

Table 3.1: *Survival function parameters, corresponding processes of RBC destruction, and nominal parameter values with upper and lower point estimates of the robust design.*

Parameter	Mechanism of RBC destruction	Nominal value	Lower point estimate	Upper point estimate	Units
s_1	Senescence	0.0241	0.0227	0.0256	$days^{-1}$
s_2	Senescence	440.78	414.26	468.99	days
r_1	Reduced lifespan due to delayed failure	0.0140	0.0132	0.0149	$days^{-1}$
r_2	Reduced lifespan due to initial failure	8.9681	8.4295	9.5404	-
c	Constant random destruction	0.0024	0.0023	0.0026	$days^{-1}$
m	Mixing parameter*	0.8941	0.8403	0.9512	-

*Defines the probability density of the senescence and random destruction component, while the probability density of the reduced lifespan part is scaled to (1-m).

The product of the survival of RBCs from Equation 3.1 and the number of RBCs born per day provides the model for labelling a single cohort of RBCs born on the same day. This would be equal to an ideal cohort labelling method where the label is only incorporated on the day of administration.

The combined survival of (multiple) cohorts born on different days can be described by the following integral over time *t*:

$$N_{RBCs}(t) = \int_0^t k_p(\tau) \cdot S(t-\tau)\, d\tau \quad \text{with } 0 \le \tau \le t$$

Equation 3.3: *Combined survival of multiple cohorts of RBCs.*

where k_p is the production rate of RBCs on each day τ (number of RBCs produced per cohort), $S(\cdot)$ is the survival of each cohort according to Equation 3.1, and $N_{RBCs}(t)$ is the total number of cells present at time *t* [76,132].

To simulate an ideal random labelling method, RBC production has to occur until the number of RBCs has reached steady state. Stopping the production at an arbitrary time point at steady state and subsequently observing the disappearance of the cells is equivalent to labelling all RBCs present at this time point and then observing the disappearance of the label [132]. In all simulations presented here and without loss of generality one unit of label is assumed to bind to one RBC.

3.4.2. Simulation of random labelling studies using radioactive chromium

^{51}Cr labelling is the most common method used for labelling RBCs. It is conducted *ex vivo* by using radioactive sodium chromate ($Na_2{}^{51}CrO_4$). The chromate anion containing hexavalent chromium can penetrate into RBCs, where the chromium binds to Hb and, subsequently, is reduced to the trivalent state [133]. ^{51}Cr is lost during the circulation from viable RBCs due to three different mechanisms (Figure 3.2): 1) vesiculation of Hb together with the bound label [68], 2) decay of the radioactive label, and 3) dissociation of the chromium-haemoglobin complex and subsequent diffusion out of the RBC for which the overall process has been termed "elution" [57,67].

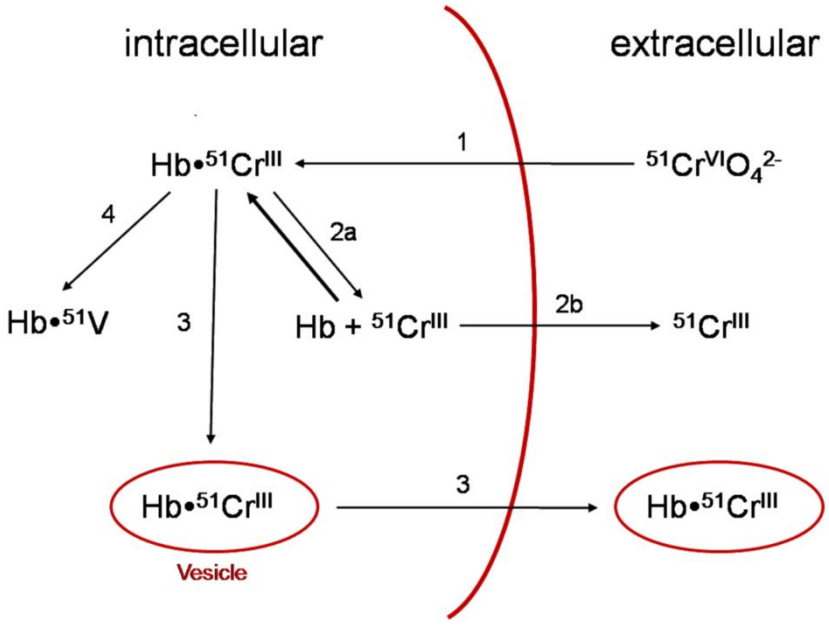


Figure 3.2: *Labelling with radioactive chromium and loss of the label from RBCs during their circulation.* **1** *Ex vivo labelling using radioactive chromate and binding to haemoglobin with reduction of the hexavalent chromium to the trivalent state.* **2** *Elution:* **a** *Dissociation of the chromium-haemoglobin complex* (*reversible process*), **b** *loss of trivalent radioactive chromium from the cell* (*irreversible process*). **3** *Vesiculation of haemoglobin together with bound label.* **4** *Radioactive decay with a half-life of 27.7025 days to stable vanadium.*

Reuse of the lost label for labelling other RBCs is not possible, as the cationic trivalent chromium cannot penetrate into RBCs again [56]. All three processes of loss were incorporated into the RBC model independently and their effect on the disappearance of the label was evaluated individually and in combination.

Vesiculation was reported to result in a loss of approximately 20% of the total Hb content over the lifetime of RBCs [68,134,135]. It was also suggested that the rate of vesiculation might be higher in older RBCs [68,136,137]. Therefore, loss of label due to vesiculation of Hb was incorporated into the survival model as a percentage loss of each RBC, with 80% of the Hb cell content still being present at the median survival of 115 days. A linear loss and an increasing loss with increasing age were considered according to the

following equation where summation was used to approximate the solution of Equation 3.3:

$$N_v(t) = \sum_{b=0}^{t} k_p(b) \cdot S(t\text{-}b) \cdot \left(1 - k_v (t-b)^v\right)$$

with $k_v = 0.2/(115 \text{ days})^n$

where $v = 1$ for linear loss

and $v > 1$ for increasing loss.

Equation 3.4: *RBC survival corrected for vesiculation.*

$N_v(t)$ is the total number of surviving RBCs at time t after loss due to vesiculation, b is the day of birth of the b^{th} cohort of RBCs, $k_p(b)$ is the production rate of RBCs on day b, $S(\cdot)$ is the survival of the b^{th} cohort according to Equation 3.1, k_v is the vesiculation rate constant that results in a 20% loss per RBC over 115 days, and v controls the rate of vesiculation. In this chapter, only integer values of v were considered with $v = (1, 2)$. Note here that for $v > 2$ the loss due to vesiculation would become too extensive after reaching 115 days, confounding the normal death of the cells governed by the underlying survival model. (Appendix A.3.2 provides additional information on the derivation of Equation 3.4 as well as the following equations.)

Elution and radioactive decay were both included in the model as random processes. The half-life of ^{51}Cr is 27.7025 days ($t_{1/2Cr}$) [138], while elution was reported to occur to approximately 1% of the remaining label per day in healthy individuals [57,139,140]. This results in a half-life of approximately 70 days for the loss due to elution ($t_{1/2el}$). The final model for simulating random labelling using ^{51}Cr is therefore given by the following equation:

$$N_{RL}(t) = N_v(t) \cdot 2^{-(t-t_L)/t_{1/2Cr} - (t-t_L)/t_{1/2el}}$$

Equation 3.5: *Random labelling corrected for vesiculation, decay and elution.*

where t_L is the time point of labelling with $t \geq t_L$, $N_{RL}(t)$ is the total number of labelled cells at time t after loss due to vesiculation, decay and elution.

3.4.3. Simulation of cohort labelling studies with reuse of the label based on the example of heavy nitrogen

Cohort labelling methods are subject to two major disadvantages in addition to the loss of label due to vesiculation of Hb: Firstly, the label is incorporated into newly produced RBCs over a prolonged time, resulting in labelled RBCs that do not have exactly the same age, and secondly, a possible reuse of the label can occur after the breakdown of initially labelled cells or from other sources where the label is incorporated due to a lack of selectivity for RBCs. ^{59}Fe for example is extensively reused after the breakdown of initially labelled cells [53]. Glycine tagged with ^{15}N is taken orally over several days and incorporated mainly into the protoporphyrin part of Hb during the production of RBCs in the bone marrow. This production of labelled RBCs occurs over an even longer period of time, strongly suggesting the presence of a glycine pool in the bone marrow [38]. Shemin and Rittenberg conducted a cohort labelling study, where the tagged glycine was given orally over three days [54]. This resulted in an increasing amount of label in the circulation for more than 20 days before an apparent steady state was reached. Additionally, there seems to be some degree of reuse of the label present as well [36]. See for instance the disappearance curve of label observed by Shemin and Rittenberg which does not return to zero even after more than 200 days and appears to asymptote to about 20% of the maximum amount of label that was present (see Figure 2.6 in Section 2.6.1 of Chapter 2) [54].

The simulation of a cohort labelling method with reuse was based on the data presented by Shemin and Rittenberg [54], assuming a prolonged but constant production of labelled RBCs over 25 days. Linear vesiculation (as per Equation 3.4) of Hb together with the incorporated label and a constant reuse of 20% of the released label from RBCs broken down on the previous day were chosen to provide the best description of their data.

The total number of labelled cells at time t ($N_{CL}(t)$) for a cohort labelling method with reuse and vesiculation is given by:

$$N_{CL}(t) = \sum_{b=0}^{t} k_p^*(b) \cdot S(t\text{-}b) \cdot \left(1 - k_v (t-b)^v\right)$$

$$\text{with} \quad k_p^*(b) = k_p(b) + reuse(b)$$

$$reuse(b) = loss_b \cdot rf$$

$$\text{and} \quad loss_b = \begin{cases} 0 & \text{for } b = 0 \\ N_v(b-1) - \left(N_v(b) - k_p(b)\right) & \text{for } b > 0 \end{cases}$$

Equation 3.6: *Cohort labelling with vesiculation and reuse of the label.*

where $loss_b$ is the loss of label between day b-1 and day b, rf is the fraction of lost label that is reused on the next day (reuse fraction), $reuse_b$ is the actual number of reused label on day b, and $k_p^*(b)$ is the apparent daily production rate including reuse. N_v, $k_p(b)$, $S(\cdot)$, k_v, and v are defined as previously.

Table 3.2 gives an overview of all fixed parameter values used in these simulations. The MATLAB® code for the proposed RBC lifespan model and its modifications for the different labelling methods is given in Appendix A.3.3.

Table 3.2: *Fixed parameter values used in simulations.*

Parameter	Values	Units	Description
Cohort labelling			
v	1	-	linear vesiculation
	2	-	increasing vesiculation
k_v	1.74×10^{-3}	$days^{-1}$	correction factor for linear vesiculation (v = 1)
	1.51×10^{-5}	$days^{-2}$	correction factor for increasing vesiculation (v = 2)
rf	0.1 0.2	-	reuse factor
Random labelling			
$t_{1/2Cr}$	27.7025	days	half-life of ^{51}Cr due to radioactive decay
$t_{1/2el}$	70	days	elution half-life of ^{51}Cr

3.4.4. Local parameter identifiability for ideal cohort and random labelling studies

To disentangle structural (or *a priori*) model identifiability from performance issues that might arise when applying a model to data (i.e. deterministic or *a posteriori* identifiability issues), structural local identifiability of the parameters in this model were determined informally based on ideal cohort and ideal random labelling experiments by applying the theory of design of experiments [29]. In this method the Fisher Information matrix was constructed according to Equation 1.32 in the Introduction to this thesis, and it was assessed whether the matrix was positive definite for a given fixed design for a single individual, indicating local identifiability (i.e. all parameters in the model are in theory identifiable).

The maximum units of label per ml of blood were set to be 400 for all simulations by choosing the daily constant production of RBCs (parameter k_p) accordingly. This corresponds to the maximum initial counts per minute (cpm) per ml blood found in a typical *in vivo* RBC survival study using ^{51}Cr. Measurement noise was determined based on *in vitro* experiments (unpublished data): The background count was 10 cpm with a variance of 1.73 cpm^2 for the additive error (σ^2_{add}), and a coefficient of variation (CV_{prop}) of 2.32% was determined for the proportional error. Both were included in the simulations as combined error model. Based on the ^{51}Cr background count a limit of detection of 5 cpm per sample analysed was calculated according to Currie [141], assuming 100% counting efficiency of the sample. Given a potential blood sample of 6 ml, the overall limit of detection would be 0.8 cpm/ml. For comparison of cohort labelling and random labelling methods, the maximum units of label/ml blood and the measurement noise were assumed to be the same for all labelling methods and were determined based on ^{51}Cr.

3.4.5. Optimised designs for parameter estimation

A hypothetical study with 100 healthy individuals was considered, where any set of blood samples was assumed to be exchangeable between the individuals. Blood sampling times were optimised based on the theory of design of experiments in order to obtain the maximum information on the parameter values in the model from these samples [29]. To account for uncertainty in the parameter values, a robust D-optimal design criterion (HClnD) was applied as introduced in the Introduction of this thesis (Section 1.1.5.1.2) [33]. The lower and upper point estimates were chosen to be symmetric around the nominal parameter values determined in the previous chapter and are given in Table 3.1. The median lifespan of RBCs based on 2^6 = 64 possible combinations of these upper and lower bounds is between 100 and 120 days (a list of the combinations is given in Appendix A.3.4). As before, the exchange algorithm (Appendix A.2.3) was used to maximise the optimal design criterion. The performance of an optimised design was then assessed based on the asymptotic lower bound of the percentage standard errors (%SE) of the parameter estimates.

100 data sets were simulated under the optimised design for the ideal labelling methods and the parameter values were re-estimated for these simulated data sets. A combined error model (additive and proportional) was considered, with three different error levels for these simulations: 1) the error used for the optimisation of the design (CV_{prop} = 2.32%, σ^2_{add} = 1.732 cmp^2), 2) CV_{prop} = 5%, σ^2_{add} = 3.5 cmp^2, and 3) CV_{prop} = 10%, σ^2_{add} = 7 cmp^2. The typical assay error levels for these type of experiments have a $\%CV_{prop}$ in the order of <5% [142-144]. A local search algorithm applying the simplex method was used for parameter estimation. The algorithm is based on the built-in MATLAB® function *fminsearch* but had been modified to allow for boundary conditions (*fminsearchbnd*, available online via file exchange at: http://www.mathworks.com/matlabcentral/fileexchange/8277, accessed on 12.10.2010).

In addition to the ideal labelling methods, two different labelling scenarios were considered for the optimisation of blood sampling times in a hypothetical *in vivo* study: 1) random labelling with ^{51}Cr and 2) cohort labelling with ^{15}N. For the random labelling study, it was furthermore investigated whether a superior (hypothetical) label that does not elute or decay appreciably during the time course of the experiment would improve the precision of the parameter estimates. The cohort labelling method was initially based on the previously described simulation scenario (constant incorporation over 25 days and a fixed reuse fraction of 20% of lost label). Furthermore, the influence of a shorter (ten days) or longer (50 days) incorporation of label, as well as a smaller fraction of reuse (10% fixed) on the precision of the parameter estimates was investigated. Finally, it was determined whether the fraction of reuse can be estimated as well, and how well a design would perform that was determined under the assumption of the wrong fraction of reuse.

3.5. Results

3.5.1. RBC survival model

Figure 3.3 shows the disappearance of an ideal cohort label and an ideal random label as predicted by the survival model.

In the case of an ideal cohort label (Figure 3.3A) all labelled RBCs are produced on the first day. Immediately afterwards, a sharp drop in the disappearance curve of the label is observed. This drop of about 5% results from the early destruction of unviable RBCs shortly after their release from the bone marrow. Subsequently, labelled RBCs are removed due to random destruction, delayed failure and senescence, resulting in an s-shaped disappearance curve. To simulate an ideal random labelling method, constant production of RBCs was assumed to occur over 500 days. This ensures steady state in the total number of RBCs as 500 days is more than double the survival time of the longest lived RBCs in the cohort labelling simulation (approximately 200 days). The disappearance of the cells after stopping the production at day 500, which is assumed to be equal to labelling all cells

present at this time point irrespectively of their age, is shown in Figure 3.3B. Instead of the s-shaped pattern observed for the simulation of a cohort labelling method, ideal random labelling results in an apparent linear disappearance with tailing at the end.

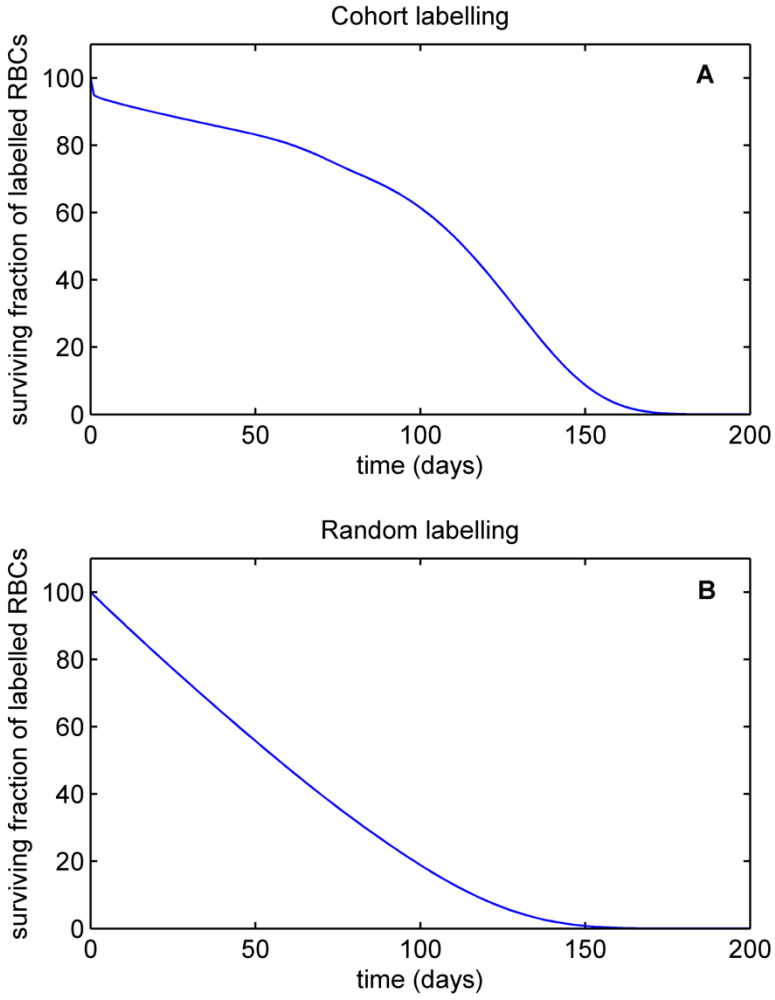


Figure 3.3: *Disappearance curves predicted for an ideal cohort label* (**A**) *and an ideal random label* (**B**).

3.5.2. Simulation of random labelling studies using radioactive chromium

The influence of all three mechanisms of loss of ^{51}Cr during the circulation was investigated individually. The resulting disappearance of label is compared with the disappearance of an ideal label in Figure 3.4.

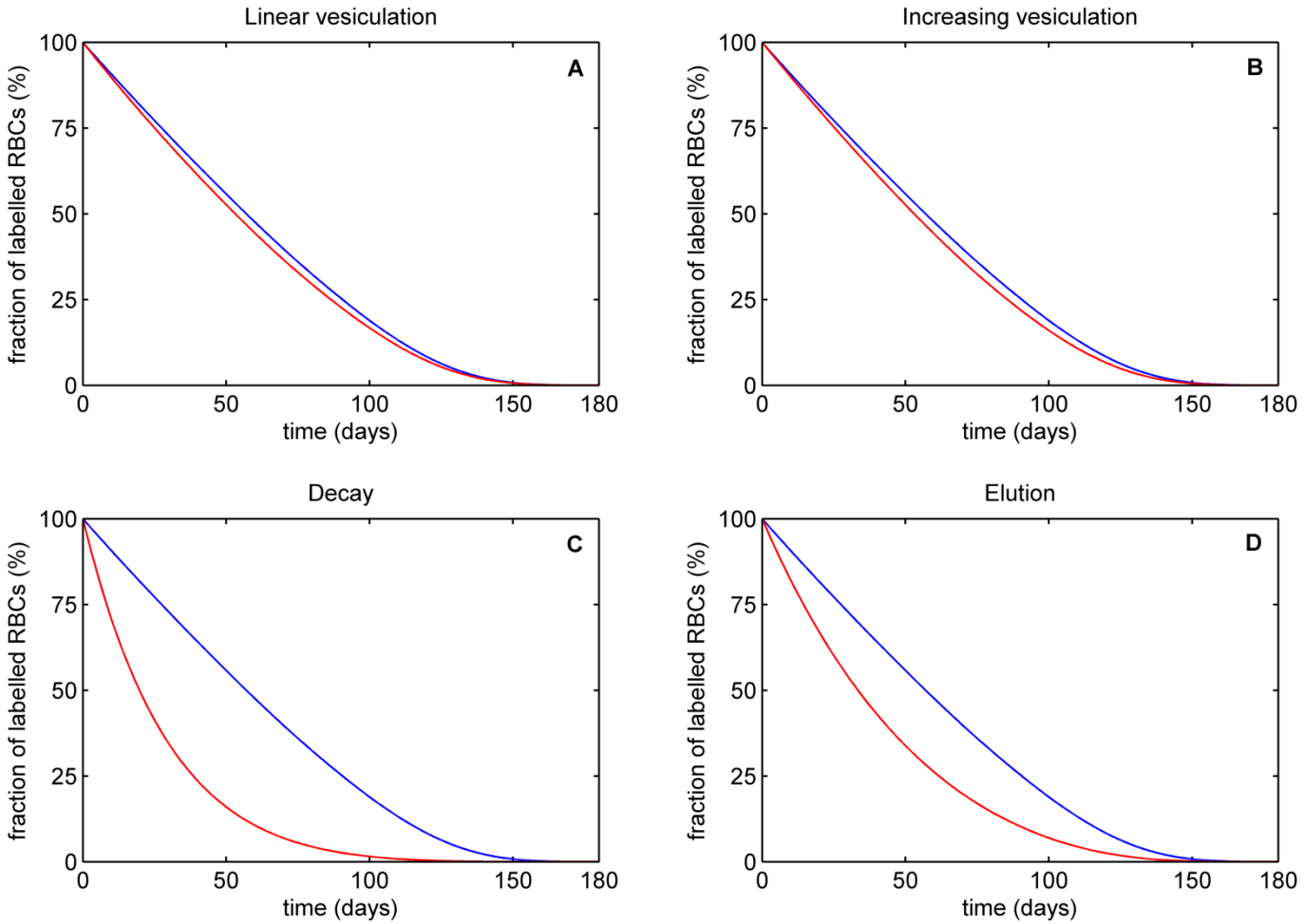
First, the loss of Hb together with bound label due to vesiculation was incorporated into the model according to Equation 3.4. Figure 3.4A shows the influence of linear vesiculation; whereas in Figure 3.4B increasing vesiculation with cell age was assumed. Slight differences in the two types of vesiculation are observed after approximately day 75, where the disappearance curve with an increasing vesiculation rate over cell age remains depressed below the disappearance curve of an ideal label for a longer period. Due to this slight deviation, further simulations were focused solely on linear vesiculation. Figure 3.4C compares the disappearance of the radioactive label with a decay half-life of 27.7025 days with the disappearance of an ideal label. Instead of the linear disappearance curve observed with an ideal label, the decay results in an exponentially decreasing disappearance curve. Finally, the influence of the dissociation of the chromium-haemoglobin complex and the subsequent loss of trivalent chromium from the cells is shown in Figure 3.4D, where a dissociation half-life of 70 days was assumed according to literature for the rate limiting step of the combined mechanism termed elution [57,140]. Elution is similar to radioactive decay, as both processes are exponential. Consequently, the influence of elution also results in an exponentially decreasing disappearance curve. However, the decline of the disappearance curve is slower due to the apparent longer half-life of elution compared to the half-life of the decay.

The combined loss of ^{51}Cr due to all three mechanisms is illustrated in Figure 3.5 in comparison to the disappearance of an ideal label.

---

following page:

Figure 3.4: *Influence of the individual mechanisms of loss on the disappearance of radioactive chromium as random label. Blue: disappearance of an ideal label; red: disappearance of the label with the corresponding mechanism of loss:* **A** *linear vesiculation of haemoglobin,* **B** *increasing vesiculation of haemoglobin,* **C** *radioactive decay* (*half-life 27.7025 days*), **D** *elution* (*half-life 70 days*).


Linear vesiculation
A
Increasing vesiculation
B
Decay
C
Elution
D
fraction of labelled RBCs (%)
time (days)
0
25
50
75
100
150
180

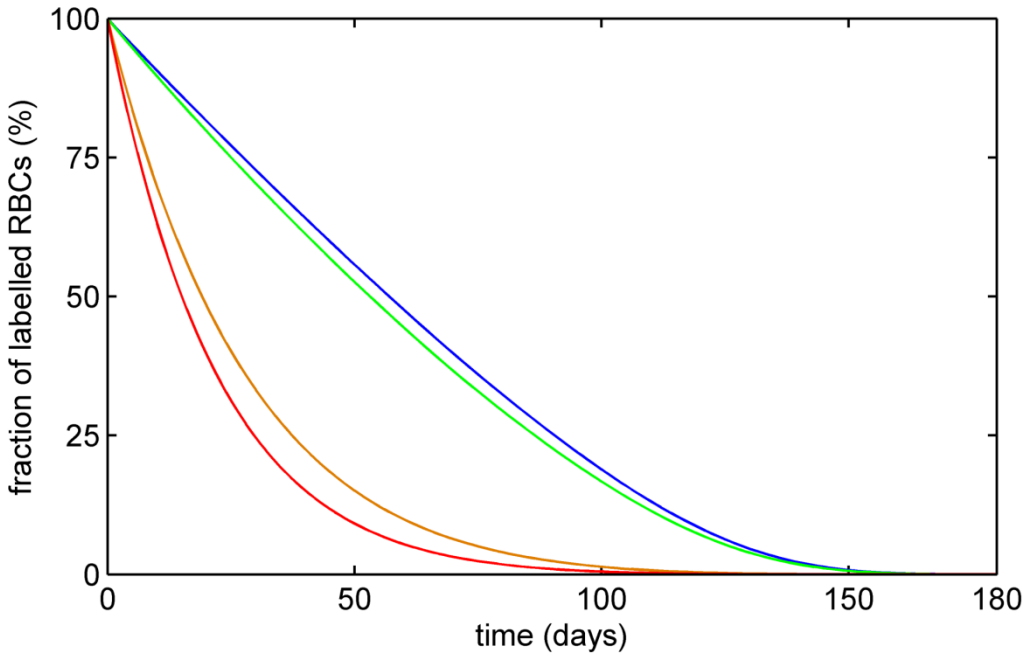


Figure 3.5: *Combined mechanisms of loss of radioactive chromium compared to an ideal random labelling method (blue). Green: linear vesiculation of haemoglobin, brown: linear vesiculation and decay, red: linear vesiculation, decay and elution.*

It is expected in clinical RBC survival studies using ^{51}Cr as the random labelling method, that one would observe the combined, very fast exponentially declining disappearance curve instead of the linear disappearance of an ideal label. Correcting this curve for the decay of the label, as it is commonly done in clinical practice, results in an apparent half-life of RBCs in the circulation of 32 days in this simulation. Note, although reported in terms of half-life, which implies a random destruction process, the principle natural mechanism of RBC destruction is in general believed to be senescence and hence not a random process. This illustrates the oversimplification caused by application of empirical corrections for decay and elution.

3.5.3. Simulation of cohort labelling studies with reuse of the label based on the example of heavy nitrogen

Cohort labelling was simulated in Figure 3.6 on the basis of survival studies using ^{15}N with a prolonged constant incorporation of the label over 25 days and a constant reuse fraction of 20% of label lost on the previous day. In addition linear loss due to vesiculation of Hb from the intact viable RBC was included.

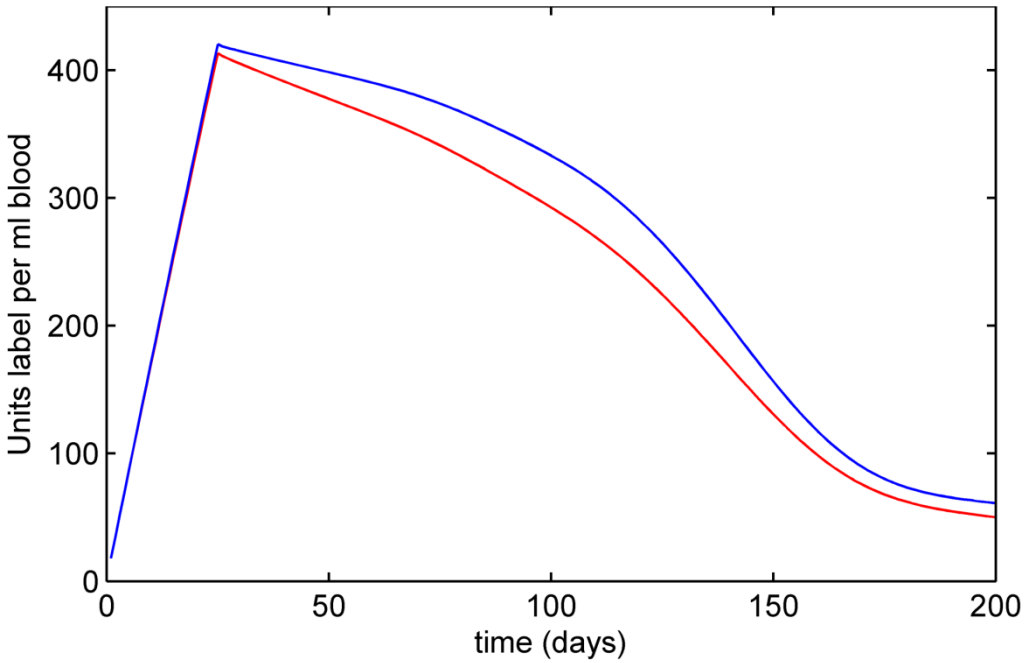


Figure 3.6: *Disappearance of a cohort label with prolonged incorporation over 25 days and constant reuse of 20% without (blue) and with (red) linear vesiculation of haemoglobin.*

The amount of label in the circulation is rising as long as production of labelled RBCs occurs, and the subsequent disappearance curve shows a similar s-shaped pattern as the disappearance of an ideal cohort label. However, in contrast to the ideal cohort labelling method, an initial drop resulting from the early death of unviable RBCs is not obvious due to the prolonged incorporation of label. Furthermore, approximately 12% of the maximum amount of label present during the entire simulation is still present in the circulation after 200 days.

3.5.4. Local parameter identifiability and optimised designs for ideal labels

The Fisher Information matrix was found to be positive definite for both the ideal cohort and ideal random labelling studies, indicating that the model is locally identifiable for a given finite design. The results for the optimal blood sampling times after labelling on day 0 and %SE of the parameter estimates calculated for the mean parameter values of the prior parameter distribution are shown in Table 3.3.

Ideal random labelling requires six blood samples between day 1 and day 140 (Table 3.3a). The %SE for five of the parameters are approximately 20% or less, while parameter r_2 controlling the initial destruction of unviable RBCs was associated with the highest %SE of about 130%. Nevertheless, all parameter values could in theory be estimated from such a study. In comparison, ideal cohort labelling is associated with an improved precision, where the blood samples are to be taken at slightly later time points (Table 3.3b). The %SE values for all but one parameter are 5% and less. Again, the initial destruction of unviable RBCs described by parameter r_2 is associated with a higher %SE of about 23%.

Table 3.3: *Sampling times and corresponding percentage standard errors (%SE) for the parameter estimates.*

Simulation	**Optimal sampling times (days)**						**%SE parameter**						
							r_1	r_2	s_1	s_2	c	m	*reuse*
Ideal labelling methods													
a) Random labelling	1	46	68	83	113	140	11.4	132.8	2.6	3.8	21.1	3.8	-
b) Cohort labelling	5	60	72	84	129	153	2.3	22.7	0.6	0.9	5.1	0.9	-
Random labelling													
c) Radioactive chromium	1	27	51	67	82	107	54.3	674.9	67.8	67.0	118.9	19.3	-
d) Fixed initial destruction	1	27	54	77	78	107	42.5	fixed	55.4	51.0	39.4	4.2	-
e) No elution of the label	1	29	54	69	85	118	31.2	362.5	24.8	27.3	74.2	11.2	-
f) No decay of the label	1	39	64	77	97	130	20.5	195.9	8.4	11.0	40.2	6.0	-

Table 3.3: *continued.*

Simulation		Optimal sampling times (days)							%SE parameter						
									r_1	r_2	s_1	s_2	c	m	*reuse*
Cohort labelling															
g)	Incorporation over 10 days and 20% reuse	11	63	76	89	129	156		3.5	34.4	1.1	1.5	7.4	1.3	fixed
h)	Incorporation over 25 days and 10% reuse	15	68	86	104	140	167		4.2	44.8	1.0	1.5	8.2	1.5	fixed
i)	Incorporation over 25 days and 20% reuse	15	69	86	104	139	166		4.9	51.9	1.3	1.8	9.5	1.7	fixed
j)	Incorporation over 50 days and 20% reuse	13	65	105	126	156	188		11.3	107.3	2.7	3.7	17.8	3.0	fixed
k)	Incorporation over 25 days and 10% reuse	15	68	86	104	140	166	201	4.2	44.8	1.1	1.5	8.2	1.5	0.7
l)	Incorporation over 25 days and 20% reuse	15	68	86	103	139	166	204	5.1	50.9	1.4	1.9	9.5	1.7	0.5

The results of the simulation-estimation analysis (Table 3.4) for ideal cohort labelling confirm the findings of the design analysis. The empirical %SE obtained at the same error level as used for the optimal design are in the same magnitude as calculated based on the Fisher Information matrix, and therefore based on simulation-estimation all parameters are estimable in this model. Note, that the %SE based on the optimal design are asymptotic lower boundaries, slightly higher values obtained based on simulation-estimation are therefore anticipated. As expected, increasing the error level results in an increase in the empirical %SE of all parameters. In contrast, simulation-estimation of ideal random labelling resulted in higher empirical %SE than calculated based on the optimal design. Nonetheless, under the experimentally derived error level, estimation of all parameters would be possible. Higher error levels are accompanied by a decrease in the success rate of the estimation procedure and a loss of information on all parameter values.

Table 3.4: *Simulation-estimation results for ideal labelling methods in comparison to the information theoretic values.*

Simulation & Error levels	%SE parameter estimates						Success [a]
	r_1	r_2	s_1	s_2	c	m	
Ideal cohort labelling							
calculated based on $\mathbf{M_F}$ [b]	2.3	22.7	0.6	0.9	5.1	0.9	-
2.32% CV_{prop} σ^2_{add} = 1.73 cpm^2	3.4	42.7	1.2	1.5	5.9	1.1	97%
5% CV_{prop} σ^2_{add} = 3.5 cpm^2	6.8	152.2	1.8	2.2	9.6	1.8	89%
10% CV_{prop} σ^2_{add} = 7 cpm^2	11.4	240.7	2.1	2.6	10.7	2.0	96%
Ideal random labelling							
calculated based on $\mathbf{M_F}$ [b]	11.4	132.8	2.6	3.8	21.1	3.8	-
2.32% CV_{prop} σ^2_{add} = 1.73 cpm^2	58.6	123.9	28.0	34.5	38.8	16.5	95%
5% CV_{prop} σ^2_{add} = 3.5 cpm^2	96.4	318.6	16.3	20.8	38.7	14.5	60%
10% CV_{prop} σ^2_{add} = 7 cpm^2	207.2	111.0	117.1	135.9	60.6	18.7	58%

[a] % runs out of 100 that resulted in parameter estimates that are within ± 100 fold of the nominal values

[b] $\mathbf{M_F}$ = Fisher Information matrix, calculations based on a combined error model with CV_{prop} = 2.32% and σ^2_{add} = 1.73 cpm^2

3.5.5. Optimised designs for non-ideal labels

The robust D-optimal design was located for random labelling with ^{51}Cr including the various loss mechanisms of the label from RBCs. The results are shown in Table 3.3c. The optimal sampling times are shifted to earlier time points compared to the ideal random labelling method. The corresponding %SE of the parameter estimates are considerably higher than for the ideal labelling method. Four of the parameters show %SE values of between 55% and 120%, while one parameter was well estimated with a %SE of approximately 20%. However, parameter r_2 controlling the initial destruction of unviable RBCs was poorly estimated with a %SE of >600%. To overcome this problem r_2 was fixed and the %SE of the estimates for the remaining five parameters were determined together with the new set of optimal blood sampling times (Table 3.3d). The optimal sampling times were minimally different. The corresponding %SE for the parameter estimates are 55% and less for all remaining parameters. However, as these parameter estimates are still associated with a low precision in comparison to an ideal random label, it was investigated whether the elution component or the decay of the radioactive label plays a greater role in this loss of precision. First, optimal blood sampling times and the corresponding %SE were determined for a radioactive random label analogue to ^{51}Cr but without loss due to elution. The results are given in Table 3.3e. The optimal blood sampling times are located on similar days as for ^{51}Cr. Precision is particularly improved for the two parameters controlling the senescence of RBCs (s_1 and s_2). The %SE for both is reduced from >60% to <30% for a label without elution. However, parameter r_2 remains poorly estimated from a study using such a label as its %SE is >300%. Second, an analogue to ^{51}Cr that had a substantially longer half-life of decay and principally only shows loss due to elution and vesiculation was considered. In this case, the optimal blood sampling times are shifted to later days (Table 3.3f). They are now located on time points up to day 130 and are closer to the optimal sampling times for an ideal random labelling method. Correspondingly, the precision of the parameter estimates has again improved. The %SE for all

parameters but r_2 are <50%, whereas the error for r_2 has decreased to approximately 200%. Furthermore, a half-life of 360 days was found to be adequate for the hypothetical radioactive label to result in %SE of a similar magnitude as would be expected for a label without decay.

The robust D-optimal design was also located for a hypothetical cohort labelling method that shows a constant zero-order rate prolonged incorporation and reuse of label from previously labelled RBCs. The results for the optimal blood sampling times and %SE are given in Table 3.3g-l. For the first setup a prolonged incorporation over 25 days and a reuse fraction of 20% was assumed on the basis of data from Shemin and Rittenberg [54]. Blood samples are taken over a longer period of time ranging up to 5 months (Table 3.3i). The corresponding %SE for all parameters but r_2 are <10%, indicating very high precision. Parameter r_2 is associated with %SE of approximately 50%. By assuming a shorter period of incorporation of just ten days, the optimal sampling times are shifted to slightly lower time points (Table 3.3g). This improves the precision of the parameter estimates even further, with r_2 having now a %SE of about 35%. In contrast, a longer period of incorporation (50 days) reduces the precision of all parameter estimates slightly (Table 3.3j). The %SE are approximately doubled in comparison to incorporation over 25 days, while the optimal blood sampling times are spread over a longer period of time where the last sample should be taken on day 188. Furthermore, decreasing the constant reuse fraction to 10% with a prolonged incorporation over 25 days has only little influence on the optimal design and the corresponding %SE (Table 3.3h) in comparison to the corresponding design for 20% reuse and 25 days of incorporation. Correspondingly, the D-efficiencies for the design evaluated at 20% reuse but applied for a reuse fraction of 10% and *vice versa* were >99.9%. Finally, an optimal design was determined where the reuse fraction was to be estimated as an additional parameter. Two scenarios with different reuse fractions (again 10 and 20%, respectively) were evaluated, while the incorporation of label was assumed to occur over 25 days (Table 3.3k and l). Estimating this additional parameter requires an additional time point. In both

cases, this time point is located at around day 200, while the remaining six optimal blood sampling times are almost unchanged when compared with the corresponding designs with a fixed reuse fraction. Correspondingly, there is only a minimal loss in the high precision for the estimates of the parameters controlling RBC survival, while the %SE for the estimation of the reuse fraction is <1%, indicating an acceptably precise design. Again, the D-efficiencies were >99.9% when the design was optimised assuming a reuse fraction of 10% but applied to a setting with 20% reuse and *vice versa*.

3.6. Discussion

3.6.1. Enhancement of the proposed RBC survival model

Previously, an individual lifespan was randomly assigned to each RBC produced during the simulation [82]. This procedure was described in the previous chapter and the Introduction to this thesis (Section 2.4.2.1 & Section 1.2.4.3.2). However, difficulties arose from this procedure when attempting to mathematically manipulate the function. The model was modified here to eliminate the need for random sampling by locating the corresponding survival function to simulate a single or multiple cohorts of RBCs.

3.6.2. Simulation of random labelling studies using radioactive chromium

The different mechanisms of loss of ^{51}Cr during the circulation of RBCs were incorporated into the final model. Loss due to vesiculation of Hb can be incorporated as a constant process or increasing in an age-dependent manner. However, loss due to an age-dependent process was minimally different compared with the linear loss and vesiculation only plays a minor role overall. As literature conflicts on vesiculation rate [68,134,145,146], a constant loss was assumed throughout this work.

Elution of ^{51}Cr is mostly reported to occur at about 1% per day in healthy individuals [57,61,140]. However, it seems to be a variable process dependent on the labelling process [147] and the disease state [67]. Furthermore, it might vary between individuals or within an individual in different RBC

subpopulations. It would therefore be highly desirable to estimate elution as an additional parameter in the model. Yet, as elution is described as a first-order loss, equivalent to a random process, it is not possible to distinguish elution from the natural random destruction provided by the estimate of the parameter *c*. Thus, elution has either to be known from *a priori* sources (e.g. the recommendations of The International Committee on Standardization in Hematology [66]), or needs to be determined independently (e.g. from *in vitro* experiments) and applied as a fixed parameter in the model to allow for unbiased estimation of *c* as the only unknown variable resulting in exponential loss. Note that the half-life of decay of ^{51}Cr to stable vanadium is known and hence can be included as an exact quantity in the equation.

In clinical practice, it is common to correct the observed disappearance of the label for the radioactive decay of the label only (ignoring any other processes of loss, e.g. elution), and simply report an apparent half-life of ^{51}Cr in the circulation instead of further calculating the mean RBC lifespan. Applying this concept to the simulation of random labelling using ^{51}Cr presented here results in an apparent half-life of 32 days after correction for radioactive decay only, which is within the normal range reported for healthy individuals (25 – 40 days) [35]. This suggests that although the model was not built from data but developed in a theoretical framework based on plausible processes of RBC destruction in the body, it is in agreement with the current survival studies of RBCs in healthy individuals.

3.6.3. Simulation of cohort labelling studies with reuse of the label based on the example of heavy nitrogen

The settings used for the simulation of cohort labelling with reuse of the label are based on data published by Shemin and Rittenberg [54]. A constant production of labelled RBCs over 25 days provided an acceptable description for their data. Although this is a simplification of the process it appears that all designs presented here are fairly robust to this choice, and it is unlikely that a

more complex description of the incorporation phase would significantly change the optimal designs and the precision of the parameter estimation.

The value of 20% reuse was chosen based on an empirical estimate from Shemin and Rittenberg's data [54], which is in agreement with findings of Amatuzio and Evans [148]. Two caveats with respect to the interpretation of reuse arise. Firstly, the protoporphyrin part of Hb together with the bound label is mainly excreted from the body in form of bilirubin. Therefore, it is unlikely that the reused label originates solely from the breakdown of previously labelled RBCs. Secondly, glycine is not only incorporated into the heme part of Hb, but is also utilized as amino acid during the formation of many other proteins, including the globin chains of Hb itself. Therefore, a reuse from these alternative sources would seem more likely [36,38]. Additionally, the model accounts for an immediate reuse of the label released from initially labelled RBCs on the very next day. Yet, Hb is produced over several days during the maturation of RBC precursors in the bone marrow resulting in a delay of the appearance of the newly labelled RBCs in the circulation. However, only the extent of reuse is relevant in the model as neither the source of reused label nor the exact time of delay is identifiable, and reuse is assumed to be continuous from the first day. Therefore, is not expected that these simplifying assumptions influence the inferences from the results presented here.

3.6.4. Local parameter identifiability & optimised design for parameter estimation

Based on the theory of optimal design, it was shown that all parameters are locally identifiable under ideal labelling conditions. The precision of the parameter estimates is slightly lower for an ideal random labelling method. Nevertheless, it would be possible to estimate the parameter values from such a study, as only the initial destruction of unviable RBCs is associated with a slightly higher error of 130%, which is amenable to a more intensive design strategy. Yet, it has to be noted, that any deviation from the optimised

sampling schedule in clinical practice would result in a design which may lead to a reduction in the precision of the parameter estimates. The optimal design results have been confirmed based on simulation-estimation analysis at the same error level as used for the information theoretic approach. Unfortunately, ideal labelling methods are not available, as all labels developed so far are associated with significant flaws. Furthermore, higher error levels, in excess of those that would be expected experimentally, result in a loss of information on the parameter estimates even under ideal labelling conditions. This was particularly seen for the random labelling method during simulation-estimation and further highlights the difficulties associated with this particular labelling method. Note, however, that the success rate and the precision of the parameter estimates could be increased by using a global search algorithm such as simulated annealing that is less likely to converge to a local minimum than the simplex method used here. This strategy was not followed in this work due to the high computational burden associated with global search algorithms.

No between subject random effects are considered in the design. This assumption implies that the variability in the lifespan between healthy individuals is less than the variability in the lifespan within an individual and hence a model that includes only the fixed effects components is of importance. This does not mean however that the variance of the between subject effects would be poorly estimated in the current design, but rather that the current design does not incorporate these elements. In the absence of empirical data that can be used to quantify the sources of variability it is not practical to relax this assumption in the current work. However, uncertainty in the parameter values used in this work was accounted for. By choosing the variance of the parameter distribution accordingly, a median survival of RBCs between 100 and 120 days was ensured [35,36]. By applying a robust optimality criterion it was ensured that the designs reported here are optimised over a range of potential parameter values and, therefore, should be applicable for healthy individuals. The study design incorporated 100 individuals with six samples per subject.

3.6.5. Comparison of the results for the different labelling methods

The results for the optimal design showed that all parameters but the initial destruction of unviable RBCs (parameter r_2) would be estimable from the hypothesised study in the case of labelling with ^{51}Cr even in the presence of the associated label flaws. In theory, this problem could be overcome by using a different random label with a longer half-life for the radioactive decay of at least 360 days. The influence of elution is relatively less than that of radioactive decay. However, this process cannot be distinguished from natural random destruction without either strong *a priori* information or data from additional studies. This problem limits the precision of estimating the lifespan of RBCs by using ^{51}Cr as a random labelling method, and provides, in agreement with Cavill [129], a clear argument against the common use of the ^{51}Cr method. The ^{51}Cr method can only provide a semi-quantitative measure of RBC survival, but cannot be used to determine the actual lifespan of RBCs due to the significant flaws inherent in this method.

It needs to be noted, that although this work was focussed on ^{15}N as an example for cohort labelling and ^{51}Cr as the most commonly used random labelling method, more recently random labelling of RBCs using biotin has been introduced [149,150]. This label is a promising alternative to ^{51}Cr, as it is not radioactive, has an apparent small loss due to elution, and is unaffected by vesiculation since the detection method using flow-cytometric analysis only counts cells as either positive or negative. Therefore, this technique comes close to being an ideal random labelling method in the sense of the work presented here. Unfortunately, the biotin label is not yet widely available in the clinical setting and requires more manipulation of the cells during the labelling process, potentially affecting the stability of the labelled cells [36]. Further work is therefore required to ultimately decide on the quality of this promising label to determine RBC survival.

Nevertheless, in contrast to random labelling, cohort labelling with reuse and vesiculation provides more information on the underlying parameter values than even an ideal random labelling method. Unfortunately, ^{15}N has not

been used and investigated as extensively as ^{51}Cr. Hence, there is limited information available on its behaviour in the body in order to provide a more detailed description of its flaws. Yet, the results presented here indicate that cohort labelling provides more information on the survival of RBCs even when associated with certain disadvantages like reuse and prolonged incorporation.

3.7. Conclusion

In summary, RBC lifespan estimates obtained by using the currently available labelling methods have to be interpreted cautiously, and the flaws associated with these methods have to be accounted for if one aims to predict for example the effect of erythropoietin treatment on haematocrit in patients with anaemia. Model based techniques that are sufficiently flexible to account for both the limitations in the methodology as well as the likely causes for RBC destruction are recommended to be used.

This work has shown that the flaws associated with the most commonly used random labelling method, ^{51}Cr, are more substantial than those associated with cohort labelling methods. Nevertheless, an adequately designed experiment will provide informative details on RBC lifespan and associated destruction mechanisms when considered in a model based analysis framework.

Chapter 4: Modelling red blood cell survival data

This chapter is based on the following peer-reviewed publication:

Korell J, Vos F, Coulter C, Schollum J, Walker R, Duffull S (2011) *Modeling red blood cell survival data*. Journal of Pharmacokinetics and Pharmacodynamics 38(6):787-801.

4.1. Synopsis of the Chapter

In this chapter, the proposed red blood cell (RBCs) survival model is applied to clinical data obtained using the most commonly used labelling technique, labelling with radioactive chromium (^{51}Cr). RBC survival is assessed in diabetic patients with chronic kidney disease (CKD) as well as age and sex matched healthy controls.

4.2. Introduction

4.2.1. Anaemia of chronic kidney disease

Anaemia of CKD is a common complication in patients with renal impairment, especially in end-stage renal failure [151,152]. It results from an insufficient production of erythropoietin (EPO) by the damaged kidneys, reduced bone marrow responsiveness to EPO and reduced capacity to produce RBCs, as well as a decreased survival of RBCs in the circulation [43,44,153,154]. Extracorpuscular factors are believed to play a major role in the decreased RBC survival [154,155], but it remains unclear whether this decrease results from increased random destruction or premature senescence of the cells. Additionally, mechanical damage during haemodialysis might affect RBC survival, while an effect of treatment with human recombinant erythropoietin (rHuEPO) on RBC lifespan is disputed [156-159].

4.2.2. Estimation of RBC survival

In general the survival of RBCs is assessed using labelling methods, and age-independent, random labelling of the cells with ^{51}Cr is most commonly used [57,66]. Beside a short half-life of ^{51}Cr ($t_{1/2Cr}$ = 27.7025 days) due to the radioactive decay [138], this method suffers from nonspecific loss of label from viable cells due to dissociation of the chromium-haemoglobin complex with subsequent loss of chromium (termed "elution") as well as vesiculation of haemoglobin (Hb) together with the bound label [67,68,160]. It is common practice to report an apparent half-life ($t_{1/2app}$) of RBCs based on this labelling

method after correction for the decay only. If we assume an age-independent random destruction mechanism of cells as the sole method of destruction, then the RBC half-life can be converted into an apparent mean RBC lifespan (LS_{app}) based on $LS_{app} = 1.44 \cdot t_{1/2app}$ [161]. It is important to note that this conversion ignores age-dependent death due to senescence and also loss of structurally deformed RBCs due to other mechanisms. Therefore this approach is likely to provide a limited understanding of the actual physiology of RBC destruction.

Similarly, alternative mathematical descriptions of RBC survival mostly assume that cell death occurs due to a single mechanism only, either random destruction or senescence [35,66,72,76,162,163]. Only a few models have been proposed that combine both mechanisms [35,66,80].

4.2.3. Newly proposed RBC survival model

In the previous two chapters, a statistical RBC survival model has been developed that, in contrast to previously proposed models, accounts for several physiologically plausible mechanisms of RBC destruction (Chapter 2) [82], and also for flaws associated with commonly used RBC labelling methods such as ^{51}Cr (Chapter 3) [164]. The destruction mechanisms in the model include senescence and random destruction, as well as cell death due to early or delayed failure, which are described by distinct parameters. These parameters give rise to a very flexible probability density function describing the change in the instantaneous probability of death over time (equivalent to a lifespan distribution). Assessing the shape of this function would allow for deeper insight into the mechanisms of destruction, for example to determine the extent of a constant risk of death due to random destruction or the presence of significant initial failure as described in Chapter 2 (Section 2.8).

4.3. Objectives

The primary aim of this chapter was to apply the previously developed model to clinical data on RBC survival. The secondary aim was to investigate whether the data supports an increased random destruction or an altered senescence as the primary mechanism underlying the decreased RBC survival in patients with CKD undergoing haemodialysis.

4.4. Materials and Methods

4.4.1. The data

RBC survival data has been described previously by Vos *et al.* [165]. Briefly, data were available for analysis from 14 patients with end-stage renal failure undergoing haemodialysis (CKD group) and 14 age and sex matched controls. These data were collected in an observational study at the Southern District Health Board Dialysis Unit, Dunedin, New Zealand (Australian New Zealand Clinical Trials Registry: ACTRN12610000145000) [165]. Table 4.1 provides an overview of the demographics in both groups. Written informed consent was obtained from all participants.

Table 4.1: *Demographics for both study groups* (*mean* ± *SD*) *[165].*

	CKD group (n = 14)	**Controls (n = 14)**
Age (years)	57.2 ± 8.6	57.3 ± 7.9
Sex (M:F)	8:6	8:6
Haemoglobin (g/L)*	122 ± 12	143 ± 10

* $p < 0.001$

All but two haemodialysis patients were anaemic with Hb concentrations between 110 and 130 g/L, and all were receiving either continuous treatment with rHuEPO or iron supplementation. Dialysis conditions were not altered during the study.

Labelling of RBCs was conducted according to the recommendations of the International Committee of Standardization in Hematology using ^{51}Cr

(Method C) [66]. This procedure involves *ex vivo* incubation of the cellular components of an autologous blood sample with radioactive chromate, which mostly binds to Hb inside the RBCs. A suspension of the labelled cells is reinjected into the patient's circulation after reduction of excess chromate to trivalent chromium using ascorbic acid. The latter form of chromium is no long available to label Hb.

In the study by Vos *et al.*, ten to 13 blood samples were taken from each individual until 52 days after labelling and analyzed for remaining radioactivity [165]. Figure 4.1 shows the disappearance of the label from the circulation in both groups. The data is expressed as %label normalised to the amount of label present at day 1 after labelling to account for the washout phase of unbound ^{51}Cr during the first day.

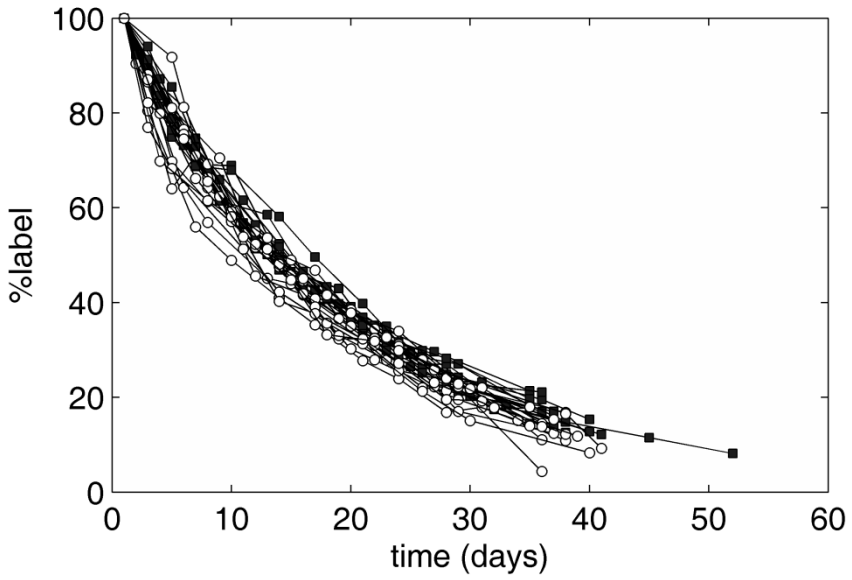


Figure 4.1: *Disappearance of ^{51}Cr from the circulation over time in CKD patients* (○) *and healthy controls* (■).

4.4.2. The model

The model used for data analysis has been described in detail in the previous chapters. To repeat briefly, the survival of a cohort of RBCs born at the same time ($t = 0$) is given by:

$$S(t) = m \cdot \left(\exp\left(-\exp\left(s_1 \cdot t - s_2/t\right) - c \cdot t\right)\right) + (1\text{-}m) \cdot \left(\exp\left(-\left(r_1 \cdot t\right)^{r_2} - \left(r_1 \cdot t\right)^{1/r_2}\right)\right)$$

***Equation 4.1**: Survival function for a cohort of RBCs born on day t = 0.*

Parameters in Equation 4.1 and the corresponding RBC destruction mechanisms controlled by these parameters are explained in Table 4.2. This table also provides their default values based on a median RBC survival of 115 days as they were derived in Chapter 2 (Section 2.4.2.2) [82].

***Table 4.2:** Mechanism of destruction and parameters controlling these.*

Parameter	Destruction mechanism	Default value [82]
s_1 (days^{-1})	Senescence (density of main maximum of lifespan distribution)	0.0241
s_2 (days)	Senescence (location of main maximum of lifespan distribution)	440.78
c (days^{-1})	Constant random destruction	0.0024
r_1 (days^{-1})	Reduced lifespan due to delayed failure	0.0140
r_2 (-)	Reduced lifespan due to initial failure	8.9681
m (-)	Mixing parameter*	0.8941

* Probability of the senescence and random destruction component given by *m*, while early and delayed failure have a probability of 1-*m*.

Based on Equation 4.1 the apparent mean lifespan $\bar{L}$ can be calculated as the integral over the survival function $S(t)$ over time:

$$\bar{L} = \int_0^\infty S(t)dt$$

***Equation 4.2:** Apparent mean RBC lifespan.*

Flaws of the ^{51}Cr random labelling method used in the study by Vos *et al.* [165] were included in the model as described in Chapter 3 (Section 3.4.2) by accounting for radioactive decay ($t_{1/2Cr}$ = 27.7025 days), elution ($t_{1/2el}$ = 70 days)

and vesiculation of Hb (20% loss over median RBC survival) according to literature values [66,68,138]. Note these values were considered to be fixed constants and not vary across cohorts or individuals. The corresponding equations are described in the Appendix A.3.2 and Chapter 3 (Section 3.4.2), while the following equation is a repetition of Equation 3.5 describing the final model for random labelling using ^{51}Cr:

$$N_{RL}(t) = N_v(t) \cdot 2^{-(t-t_L)/t_{1/2Cr} - (t-t_L)/t_{1/2el}}$$

Equation 4.3: *RBC survival model accounting for flaws associated with random labelling using ^{51}Cr.*

As it was shown in Chapter 3 (Section 3.5.5), full parameter estimation for the proposed RBC survival model would be possible in theory from a random labelling method; however the nonspecific loss of label is a significant limiting factor which compromises the power of the design to provide information on all mechanisms of RBC destruction simultaneously. In this chapter a simple overlay of the data on the model prior predictions was considered first, followed by focusing on the estimation of the parameters of highest clinical interest with respect to the available data, e.g. parameter s_2 (controlling the location of the senescence component) or random destruction (parameter c), while the remaining parameters are fixed to their default values.

This procedure was applied here based on two different approaches: (1) a two-stage approach and (2) a full population approach. For both approaches, it was assessed whether estimating senescence (Scenario A) or estimating random destruction (Scenario B) provided the better fit based on changes in objective function value (OFV). Note that this does not exclude a combination of both or even multiple contributing mechanisms as possible cause for a decreased RBC survival in patients with CKD. However, this could not be assessed in this work due to the limitations of the ^{51}Cr labelling method.

An initial exploratory two-stage approach was conducted to assess whether there was a consistent preference towards one of these scenarios across individuals within each group. Since it was not possible to estimate both

senescence and random destruction simultaneously, it was not possible to determine from a population analysis the non-dominant mechanism in the population. Furthermore, the apparent mean RBC lifespan $\bar{L}$ was calculated for both groups (CKD and controls) in each scenario based on the parameter estimates using Equation 4.2, and the differences between the groups were compared.

4.4.3. Two-stage approach

The two-stage approach was conducted in MATLAB® (The MathWorks™ Inc., Natick USA). The model was fitted to the data of each individual separately using iteratively reweighted least squares (*IRWLS*) as the objective function until convergence of the parameter estimates was obtained (as per [166]). The objective function for the i^{th} subject in the k^{th} iteration is given by:

$$OFV_{IRWLS}\left(\hat{\boldsymbol{\theta}}_i^{\{k\}}\right) = \sum_{j=1}^{n_i}\left(y_{ij} - f\left(x_{ij}, \hat{\boldsymbol{\theta}}_i^{\{k\}}\right)\right)^2 \cdot \left(w_{ij}^{\{k\}}\right)^{-1}$$

Equation 4.4: *Iteratively reweighted least squares (IRWLS) as objective function in the two-stage approach.*

where n_i is the total number of observations y_{ij} in individual i, $f(\cdot)$ is the structural model according to Equation 4.3, x_{ij} are the independent variables of individual i, $\hat{\boldsymbol{\theta}}_i^{\{k\}}$ are the individual parameter estimates of subject i during iteration k, and $w_{ij}^{\{k\}}$ are the weights for the j^{th} observation in the i^{th} individual during iteration k. Starting from ordinary least squares in the first iteration, the weights are updated during each subsequent iteration $k > 1$ based on the estimates of the previous iteration k-1:

$$w_{ij}^{\{k\}} = \begin{cases} 1 & \text{for } k = 1 \\ f\left(x_{ij}, \hat{\boldsymbol{\theta}}_i^{\{k-1\}}\right)^2 & \text{for } k > 1 \end{cases}$$

Equation 4.5*: Weights for the iteratively reweighted least squares.*

A local search algorithm based on the simplex method and which allows for boundary conditions (*fminsearchbnd*, obtained via file exchange from: http://www.mathworks.com/matlabcentral/fileexchange/8277, accessed on 12.10.2010) was used for minimization in MATLAB® to obtain the best parameter estimates in the parmeter space **Θ** during each iteration:

$$\hat{\boldsymbol{\theta}}_i^{\{k\}} = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{argmin}\left(OFV_{IRWLS}\left(\boldsymbol{\theta}_i^{\{k\}}\right)\right)$$

Equation 4.6*: Objective function minimisation.*

The estimation process was repeated until the absolute difference in the parameter estimates for two consecutive iterations was less than 10^{-6}, i.e. $\left|\hat{\theta}_{ip}^{\{k\}} - \hat{\theta}_{ip}^{\{k-1\}}\right| < 10^{-6}$ for the p^{th} parameter.

The population mean parameter values ($\bar{s}_2$ and $\bar{c}$ for Scenario A and B, respectively) were calculated for the CKD group and the controls according to Equation 4.7, while the between subject variability (BSV) was calculated based on Equation 4.8 for the p^{th} parameter [22]:

$$\bar{\theta}_p = \frac{1}{N}\sum_{i=1}^{N}\hat{\theta}_{ip}$$

Equation 4.7*: Population mean parameter value.*

$$\Omega_p = \frac{1}{N-1}\sum_{i=1}^{N}\left(\hat{\theta}_{ip} - \bar{\theta}_p\right)^2$$

Equation 4.8*: Between subject variance.*

where N is the total number of individuals in each group.

For each individual it was determined which scenario provided the better fit (lower OFV), and within each group it was assessed which scenario was preferred in the majority of individuals. The apparent mean RBC lifespans for both groups calculated according to Equation 4.2 were compared using an unpaired t-test with unequal variance for both scenarios.

4.4.4. Population approach

The population approach was conducted using the SAEM algorithm in MONOLIX® 1.1 (this version was easily amendable for the implementation of the RBC survival model) [27,167]. MONOLIX® maximizes the log likelihood (LL) denoted as OFV_{LL} in this chapter. Goodness of fit was determined based on the OFV_{LL}, where the better fit is characterized by a higher value. For nested models (when appropriate), the likelihood ratio test (LRT) is used as criterion with ΔOFV of ≥3.84 being significant for one degree of freedom.

Additive, proportional and combined error models were tested as statistical models for the residual unexplained variability (RUV).

MONOLIX® 1.1 calculates the individual estimate for the p^{th} parameter in the base model without covariates according to:

$$\hat{\theta}_{ip} = \hat{\tilde{\theta}}_p \cdot e^{\eta_{ip}}$$

Equation 4.9: *Individual parameter estimate as calculated by MONOLIX® 1.1.*

where $\hat{\theta}_{ip}$ is the individual parameter estimate of the i^{th} individual, $\hat{\tilde{\theta}}_p$ is the population mean estimate and η_{ip} is the random effect for the i^{th} individual [24].

Sex and CKD were tested as covariates on the full data set. Age was not included in the covariate analysis as there is no *a priori* evidence from the literature that it has an effect on RBC survival and also due to the homogeneous selection of both study groups with a narrow age range. In MONOLIX® 1.1, the effect of a covariate on the individual estimate of parameter *p* in the i^{th} individual is calculated according to:

$$\hat{\theta}_{ip} = \hat{\tilde{\theta}}_p \cdot e^{\beta \cdot cov_i} \cdot e^{\eta_{ip}} \quad \text{with} \begin{cases} cov_i = 0 & \text{for controls} \\ cov_i = 1 & \text{for CKD} \end{cases}$$

Equation 4.10: *Covariate model in MONOLIX® 1.1.*

where β is the estimated coefficient of the covariate and cov_i is an indicator variable taking a value of 0 for controls and 1 for CKD patients [24].

Significance is determined based on the Wald test in MONOLIX® on the level of the parameter estimates themselves, and was confirmed based on the LRT with respect to the full model. The Wald test is an approximate local test for significance of a parameter estimate based on the assumption of a symmetric confidence interval around the estimate, while the LRT is a global test assessing the significance of including the additional parameter in the model based on the overall goodness of fit without relying on the assumption of symmetry in the deviations. Therefore, the LRT is more powerful compared to the Wald test and the LRT results were considered to be decisive if both tests resulted in different outcomes with respect to the significance of a parameter. In addition, visual predictive checks (VPCs) were plotted for model evaluation.

4.5. Results

A simple overlay of the data and the model prediction without fitting is shown in Figure 4.2. Good agreement between model prior prediction and the data was observed.

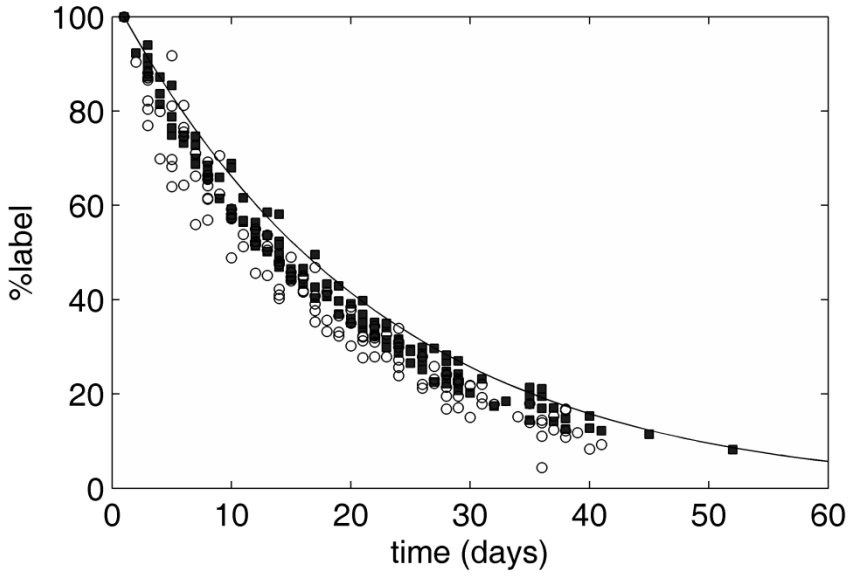


Figure 4.2: *Data (○ CKD patients , ■ controls) and model prior prediction (—) without fitting.*

4.5.1. Two-stage approach

Table 4.3 summarizes the results for the two-stage approach estimations.

Table 4.3: *Results for the two-stage approach.*

	CKD group	Controls	p value	Description
Scenario A: *Estimating senescence* $\bar{s}_2$				
$\bar{s}_2$ (days)	147.38	249.63	0.0005	population mean
Ω_{s2} (days^2)	2267.97	6527.59		BSV of population mean*
$\bar{L}_{s_2}$ (days)	59.3	75.6	0.0002	mean RBC lifespan
$\Omega_{L,s2}$ (days^2)	80.3	120.8		BSV of mean lifespan*
Scenario B: *Estimating random destruction* $\bar{c}$				
$\bar{c}$ (days^{-1})	0.0169	0.0094	0.0002	population mean
Ω_c (days^{-2})	2.62 x 10^{-5}	1.33 x 10^{-5}		BSV of population mean*
$\bar{L}_c$ (days)	49.8	69.1	0.0002	mean RBC lifespan
$\Omega_{L,c}$ (days^2)	108.1	174.8		BSV of mean lifespan*

* BSV = between subject variance

Both estimation scenarios resulted in a statistically significant reduction in the apparent mean RBC lifespan in the CKD group compared to the controls: $\Delta\bar{L}$ = 16.3 days for estimating senescence (p = 0.0002), $\Delta\bar{L}$ = 19.3 days for estimating random destruction (p = 0.0002). Estimating random destruction was preferred over estimating senescence in eleven out of 14 individuals in both groups based on the OFV. Individual fits for four representative individuals in each group are shown in Figure 4.3. (The individual fits for all individuals are shown in Appendix A.4.1)

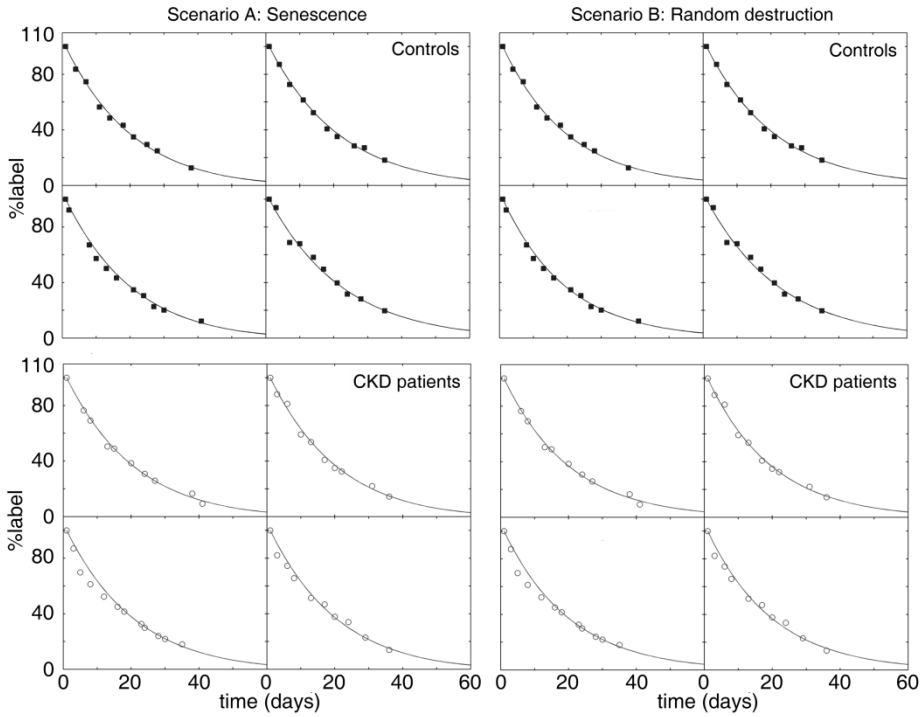


Figure 4.3: *Representative individual fits* (—) *obtained in the two-stage approach for estimating senescence* (*left column*) *and random destruction* (*right column*) *for controls* (■, *upper panels*) *and CKD patients* (○, *lower panels*).

4.5.2. Population approach

In the population approach, a combined error model was found to describe RUV best. The MONOLIX® estimates for the base model without covariates are given in Table 4.4. CKD was found to be a significant covariate based on the Wald test and LRT. Sex did not show a significant effect and was not included in the final model. Estimates for the final model using a combined error model and CKD as the only covariate are also given in Table 4.4 for both scenarios.

Table 4.4: *Fixed and random effect parameter estimates (%SE) for the population approach.*

Parameter	Estimates				Description
	Scenario A: *Estimating* $\hat{\bar{s}}_2$		Scenario B: *Estimating* $\hat{\bar{c}}$		
	Base model	Final model	Base model	Final model	
$\hat{\bar{s}}_2$ (days)	162.49 (6.9)	201.42 (7.4)	-	-	population mean estimate
$\hat{\bar{c}}$ (days^{-1})	-	-	0.0133 (7.3)	0.0106 (8.5)	population mean estimate
β_{CKD}	-	-0.4511 (23.3)	-	0.4703 (23.8)	covariate factor of CKD
Ω	0.1117 (34.1)	0.0601 (36.3)	0.1296 (33.6)	0.0721 (34.5)	between subject variance
$\bar{L}$ (days)	62.8	69.4	56.0	63.7	mean RBC lifespan
$\bar{L}_\beta$ (days)	-	56.2	-	48.1	mean RBC lifespan with covariate effect of CKD
σ_{add}^2	2.96 (12.5)	3.12 (12.4)	2.27 (12.4)	2.05 (12.1)	additive error (variance)
CV_{prop}	0.0251 (12.5)	0.0244 (12.4)	0.0234 (12.4)	0.0256 (12.1)	proportional error (coefficient of variation)
OFV_{LL}	-780.10	-773.43	-752.12	-745.51	log likelihood

Estimating random destruction was again preferred over an accelerated senescence for both models based on the OFV_{LL}, albeit these models are not nested. Figure 4.4 shows the population mean prediction for the base model as well as the full model for both scenarios, while Figure 4.5 shows the corresponding VPCs. An over-prediction during the first 15 – 20 days is apparent in these figures.

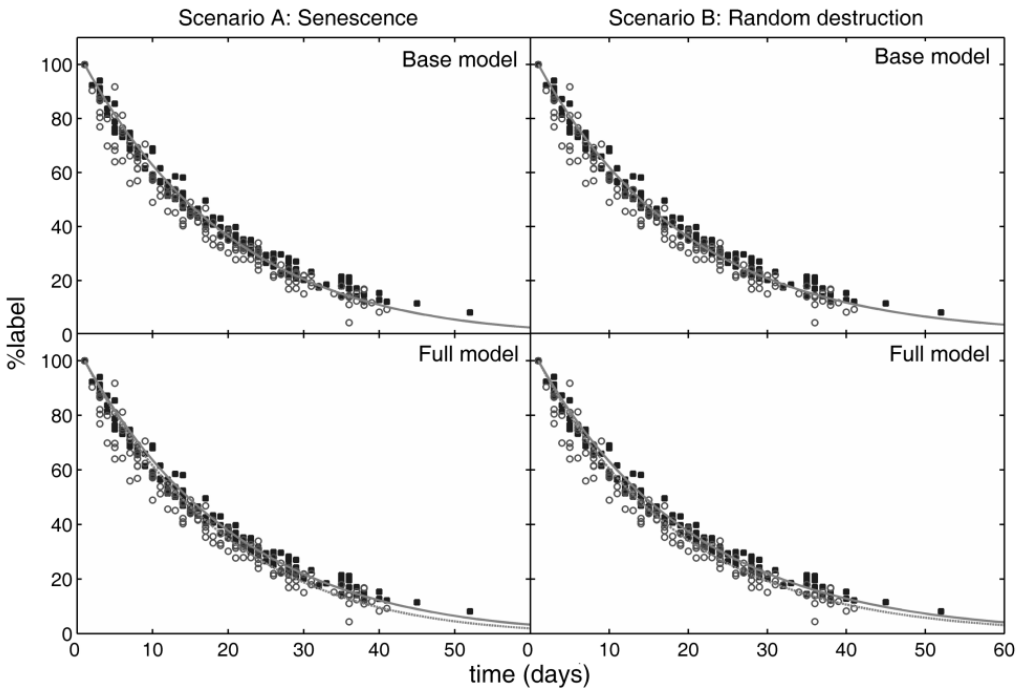


Figure 4.4: *Population fit obtained using MONOLIX®: Base model* (*upper panels*) *and full model* (*lower panels*) *for controls* (—, $cov_i = 0$) *and CKD patients* (⋯, $cov_i = 1$) *for Scenario A* (*left column*) *and Scenario B* (*right column*)*, respectively.* (■ *data of controls,* ○ *data of CKD patients*).

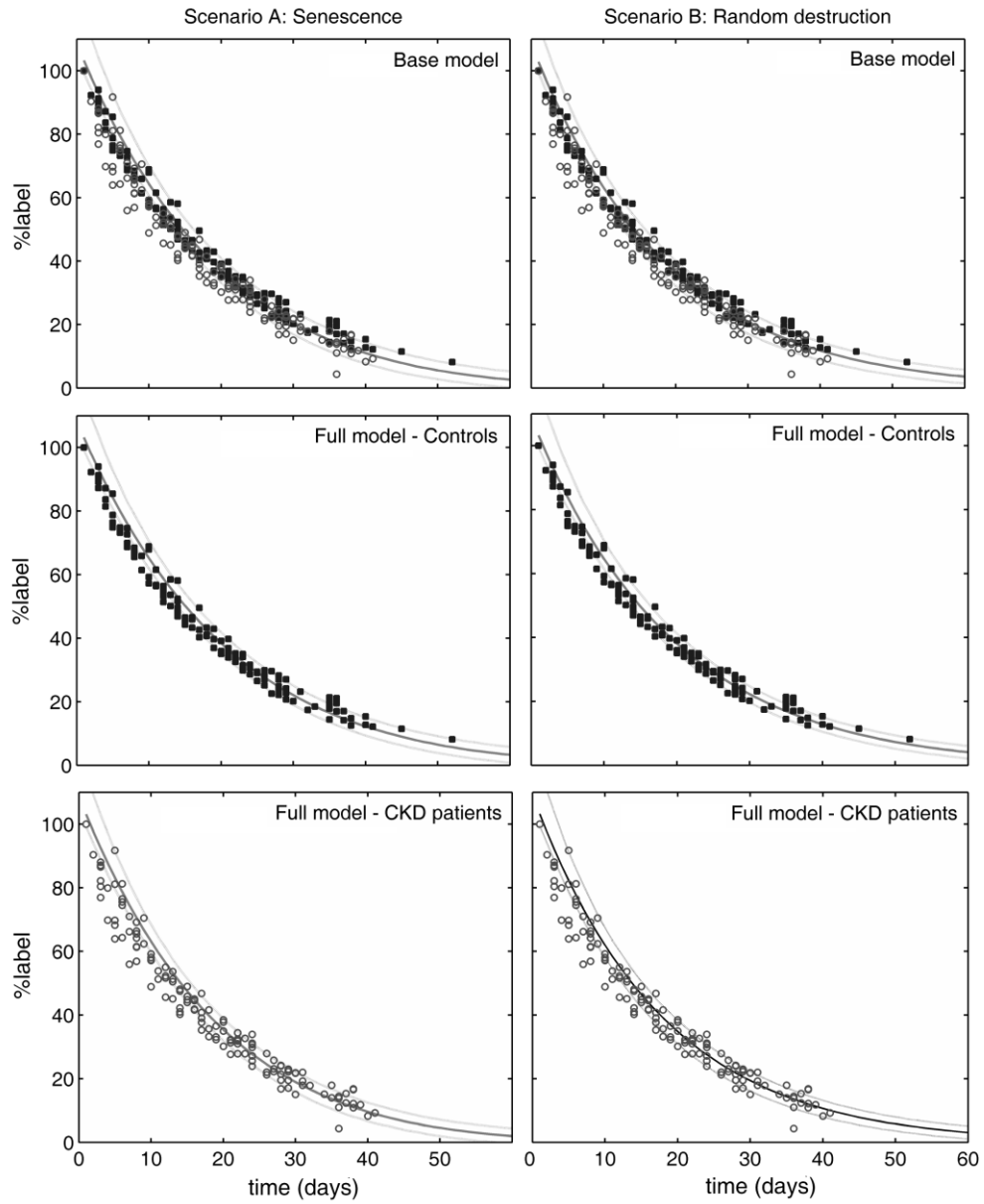


Figure 4.5: *VPCs for estimating senescence* (*left column*) *and random destruction* (*right column*) *for the base model* (*first row, no covariate*), *the full model for controls* (*second row,* $cov_i = 0$), *and the full model for CKD patients* (*third row,* $cov_i = 1$). (■ *data of controls,* ○ *data of CKD patients;* — 50^{th} *percentile,* — 10^{th} *and* 90^{th} *percentiles of model predictions*).

4.6. Discussion

4.6.1. Modelling of the clinical data

The data set does not allow full parameter estimation under the proposed model as the study was not designed for this purpose. Nevertheless, the available rich data for each individual allowed estimation of the parameters of highest interest in a classical two-stage approach and also supports a full population approach with covariate analysis. It is noted that RUV and BSV are not differentiated in the two-stage approach; therefore, BSV was naturally overestimated. A clear preference was seen for random destruction being the predominant mechanism of RBC destruction and this was consistent across individuals.

CKD was found to be a significant covariate in the population approach leading to a reduction in RBC survival and a significant reduction of approximately 45% in the unexplained BSV. In contrast, sex was found not to be a significant covariate. This was not unexpected as women of childbearing age were excluded from this study and peri-menopausal women are believed to have the same apparent RBC survival as males [35,38].

Both approaches resulted in similar estimates for the apparent mean RBC lifespan with a significant reduction in RBC survival in the CKD group compared to the healthy controls with a relative reduction of 20 - 30% depending on the analysis approach. Although this reduction is less than reported in the literature of up to 50% [44,153,168], it has to be noted that previous studies did not rely on age and sex matched controls, and had only a small number of healthy individuals in the control group, or normalise their results to an assumed survival value of 120 days in healthy individuals [169]. Recently, Vos *et al.* [27] published the clinical data that was used here in this analysis. Their own analysis used standard techniques and found a reduction in the median RBC survival of 20% between CKD patients and controls. This outcome shows good agreement with the model based approach presented

here. However, in this work it was possible to tease out a deeper insight into possible underlying mechanisms of RBC destruction.

4.6.2. Preferred mechanism of RBC destruction

Estimating the random destruction component was preferred over estimating the main parameter controlling senescence on an individual level (two-stage approach) as well as in the full population setting. This suggests the presence of a higher degree of random destruction (approximately 1%) in healthy individuals than assumed in the previous chapters (0.24%). However, random destruction and age-independent loss of the label from viable RBCs (elution) are both first-order processes that cannot be distinguished [170]. Elution was fixed to 1% per day in this work according to literature [171]; but it was also shown that the degree of elution depends on the methods used during *ex vivo* labelling of the cells (ranging from 0.95% - 1.7% per day) [147]. Since the elution process is based on physicochemical properties of ^{51}Cr binding to Hb it is plausible that pathological differences between individuals in this study are not likely to drive this process, albeit that this assertion is unprovable in this work. Therefore, the common practise to correct for a constant elution rate of 1% per day as the only mechanism of non-RBC related loss of label was retained in this analysis.

Nevertheless, CKD patients showed an additional increase in random destruction compared to healthy individuals which would support an increased random destruction as a possible underlying mechanism of anaemia in CKD, although this finding could also be explained by a higher degree of elution in CKD patients. With respect to these assumptions, the apparent mean lifespan values obtained in this analysis (as in any previous study using ^{51}Cr) have to be regarded as relative values and do not reflect the actual mean survival of RBCs in the circulation.

Moreover, nonspecific binding of excess ^{51}Cr to plasma proteins with a short half-life is also likely to be present with this labelling method. In addition, ^{51}Cr binds to platelets and leukocytes during the *ex vivo* labelling process in

whole blood [172-174]. This would be expected to result in an additional loss of label from the circulation that is not related to RBC survival or the loss of label from RBCs. At this point, the impact of these processes have not been quantified sufficiently in literature to be incorporated into the model. The model appears to over-predict the initial phase of loss of label from the circulation up to approximately day 20, as was seen in the VPCs. This discrepancy is in keeping with the nonspecific labelling of other cellular and protein elements in the autologous blood matrix. The higher initial loss has also been attributed to fragile RBCs damaged during the *ex vivo* labelling process. An exploratory compartmental analysis testing for mono- *versus* biphasic behaviour was conducted which did not reveal the presence of two random processes of loss of label in this data set as the one-compartment model was preferred in both groups (results not shown). It was also investigated whether the model supported estimation of initial or delayed destruction by assessing the parameters r_1 and/or r_2. Using a two-stage approach, the initial failure r_2 was estimated consecutively after estimating random destruction c with an adjusted mixing parameter m = 0.5 to allow for a higher contribution of the initial failure, and the estimation of c was then repeated based on the obtained value of r_2. The fit improved to a negligible extent (not statistically significant). However, due to the limited information inherent in random labelling with ^{51}Cr it is not possible to determine the contribution of initial or delayed failure simultaneously with estimating random destruction or senescence in a full population approach. Therefore, the desired deeper insight into the various RBC distruction mechanisms contributing to anaemia of CKD cannot be achieved at the moment. Nevertheless, the model captures the later phase of loss of label well, and thus provides a good description of the clinically more important long-term survival of RBCs.

4.7. Additional analyses

In addition to the analyses presented here in this chapter, two more analyses were conducted on the data of Vos *et al.* [165]:

- The administration of rHuEPO was tested as covariate on the survival on RBCs based on the population approach.
- An additional cohort of five patients receiving peritoneal dialysis (PD) in the study by Vos *et al.* was analysed using the two-stage approach.

As these analyses are not directly related to the objectives of this chapter, only a brief background and a concise discussion of the results is provided here, while the materials, methods and full results can be found in the corresponding appendices as indicated below.

4.7.1. Additional analysis testing for rHuEPO administration as covariate

4.7.1.1. Background

It is discussed in literature that rHuEPO treatment has an effect on RBC survival; however the nature of this effect is disputed. Some authors report an increased RBC lifespan under rHuEPO treatment [157,175]. Yet, others have found no evidence for an increased RBC survival in CKD patients under rHuEPO treatment [158,176]. In contrast, in haematologically normal rats, administration of rHuEPO resulted in a shortened RBC lifespan that could be attributed to an inefficient RBC production due to stress erythropoiesis [177].

4.7.1.2. Results & Discussion

An additional analysis based on the population approach was conducted to test whether the administration of rHuEPO is a significant covariate on RBC survival. The materials and methods used for this analysis together with the full results are presented in Appendix A.4.2.

In summary, rHuEPO proved to be a significant covariate. However, the inclusion of this covariate effect in the model resulted in the effect of CKD on

RBC survival to become insignificant. CKD patients receiving only iron supplementation had an RBC survival similar to the healthy controls. The same preference towards estimating random destruction was seen as with the model using CKD as covariate. It needs to be noted that using rHuEPO as a covariate instead of CKD resulted in a slightly improved fit for both scenarios according to the higher OFV:

Table 4.5: *Comparison of OFV_{LL} for different covariates in the model.*

Covariate in model	Scenario A	Scenario B
CKD	-773.43	-745.51
rHuEPO	-769.65	-742.10

Although it is tempting to conclude from these results that the administration of rHuEPO and not the underlying pathology of CKD results in the decreased RBC survival, such a conclusion cannot be drawn irrevocably due to the small numbers of patients not receiving rHuEPO treatment in this study. A type I error cannot be excluded. Importantly, it is likely that these effects are principally due to confounding, as it is likely that only CKD patients with more severely decreased RBC survival and therefore more severe anaemia received rHuEPO treatment. In fact, the Hb concentrations of two of the three patients not receiving rHuEPO were higher than the average in the CKD cohort, while the third was close to the mean (136, 146, and 126 g/L *versus* mean = 122 g/L). Note, that the definition of anaemia is a Hb concentration of <130 g/L.

4.7.2. Additional analysis for patients receiving peritoneal dialysis

4.7.2.1. Background

In patients undergoing haemodialysis (HD) RBCs are subjected to high mechanical stress during the extracorpuscular circulation and additional damage of the cells by the dialysis membrane is likely. This is supported by results of Medina *et al.* who found increased RBC turnover during HD [64].

However, the literature remains unclear as to whether RBC survival is truly decreased further by HD, e.g. in comparison to less mechanically stressful dialysis methods such as peritoneal dialysis (PD).

Vos *et al*. included five additional CKD patients into their study that received PD rather than HD [165]. Surprisingly, patients receiving PD had a mean RBC lifespan similar to, but slightly shorter than patients undergoing HD (55.3 *versus* 58.1 days) based on their empirical analysis. However, they suggested that their results could have been confounded with the older age of the patients in the PD cohort as the effect of age on RBC survival is unknown.

4.7.2.2. Results & Discussion

Here, the data of the five PD patients in the study by Vos *et al*. were analysed based on the more mechanistic RBC survival model using the two-stage approach as it was described for the HD patients and controls. The full results for the PD patients are presented in Appendix A.4.3.

Briefly, the calculated apparent mean RBC lifespan values in these patients fall between those estimated for HD patients and healthy controls for both scenarios (Table 4.6). Yet, none of these differences were found to be statistically significant based on unpaired t-tests with unequal variances. Note, that due to the small number of PD patients this cohort was not included in the population analysis presented in this chapter.

Table 4.6: *Comparison of apparent mean RBC lifespan values obtained using the two-stage approach for the different patient cohorts.*

Destruction mechanism	HD patients	PD patients	Controls
Scenario A	59.3	63.2	75.6
Scenario B	49.8	55.4	69.1

In contrast to the somewhat surprising results of the empirical analysis by Vos *et al.*, RBC survival in PD patients was found to be slightly longer than in HD patients when using the more mechanistic based RBC survival model for analysis. This new analysis strengthens the hypothesis that increased

mechanical stress during HD results in a decrease in RBC survival in addition to the shortening caused by the underlying pathology of CKD; yet the extent of this seems to be minor. Further studies in a larger patient cohort (preferably age and sex matched between the PD and HD groups) are required to ultimately determine the effect of HD on RBC survival. However, the difference between the results obtained by Vos *et al.* and the results based on the mechanistic RBC survival model shown here also indicates the importance of the choice of an adequate model for RBC survival.

4.8. Conclusion

In conclusion, individual parameters of the proposed RBC survival model can be estimated from available clinical data in a two-stage approach as well as using a full population approach. RBC survival is significantly reduced in CKD patients undergoing haemodialysis despite iron supplementation and treatment with rHuEPO. Without ruling out multiple contributing mechanisms, this analysis suggests that an increased random destruction is more likely as underlying pathological mechanism than an accelerated senescence.

In additional analyses it appeared that administration of rHuEPO or CKD is related to the decreased RBC survival and the current data are insufficient to delineate these processes.

Albeit not statistically significant, patients on PD showed a somewhat longer RBC survival than patients undergoing HD but still had a shorter apparent mean RBC lifespan than healthy controls. This suggests that the increased mechanical stress on the RBCs during HD has a small additional negative effect on their survival.

PART III

PHARMACOKINETICS OF METHOTREXATE IN RED BLOOD CELLS

Chapter 5: A population pharmacokinetic model for methotrexate measured in red blood cells

This chapter will form the basis for a publication that will be submitted as:

Korell J, Stamp L, Duffull S *et al.* (2012) *A population pharmacokinetic model for methotrexate measured in red blood cells*.

5.1. Synopsis of the Chapter

In this chapter, a population pharmacokinetic (PK) model for methotrexate (MTX) and its polyglutamated metabolites (MTXPGs) is developed using a top-down approach based on data measured in red blood cells (RBCs). Model development is broken into two parts: 1) the development of a PK model for the parent drug only; and 2) the development of a full parent-metabolite model. The final model is then assessed based on simulations and compared with previous findings. Inferences from the model will be explored in Chapter 6.

5.2. Introduction

MTX is the gold standard disease modifying anti-rheumatic drug (DMARD) in the treatment of rheumatoid arthritis (RA) [108]. MTX is a so-called folate antagonist, yet its mechanism of anti-inflammatory activity is poorly understood. Besides inhibiting key enzymes in the folate pathway, such as dihydrofolate reductase (DHFR), MTX also interferes with DNA synthesis and inhibits the enzyme ATIC (5-aminoimidazole-4-carboxamide ribonucelotide transformylase) which in turn reduces the production of proinflammatory cytokines such as interferon-γ, TNF-α and interleukin 1β due to an accumulation and secretion of adenosine [100-105].

Usually, low MTX doses (5 – 20 mg weekly) are used orally in the treatment of RA [108]. However, parenteral administration, either subcutaneous (sc) or intramuscular, can lead to better disease control in patients that show poor responsiveness to oral MTX [178-184], and parenteral administration is also associated with fewer adverse effects. Both can be attributed at least partially to the higher bioavailability via the parenteral route [87,88,91,185], as oral MTX is taken up via a saturable, active transport mechanism from the gut. When higher oral MTX doses (15 – 20 mg) are administered saturation of the active uptake transporters can occur, which results in MTX not being completely absorbed from the gastrointestinal tract.

This can increase local gastrointestinal adverse effects such as nausea and vomiting. Parenteral administration on the other hand allows higher doses to reach the systemic circulation than can be achieved orally, which leads to a better disease control than obtained with oral administration.

Fast disease control is desired in the therapy of RA to prevent joint damage, which is largely irreversible. However, MTX doses required to achieve adequate disease control are highly variable between patients and difficult to predict as no target for treatment monitoring has been established. The most common biomarker for therapeutic drug monitoring (TDM), plasma concentration, has not proved suitable for monitoring low-dose once weekly MTX treatment, as no correlation between MTX plasma concentrations and clinical outcomes of low-dose MTX has been established [115]. In addition, MTX plasma concentration rapidly fall below the limit of quantification (LOQ) and steady state plasma concentrations cannot be measured.

More recently, MTX and MTXPG concentrations measured in RBCs have been suggested as a potential biomarker for MTX monitoring [114]. However, RBCs are not involved in any of the postulated anti-inflammatory mechanisms of action of MTX. Therefore, MTX concentrations measured in RBCs are unlikely to have a causal relationship with disease outcomes or MTX treatment effects in RA. Yet, as MTX and MTXPGs accumulate inside RBCs (and other cells) [95], MTX and MTXPG concentrations measured in RBCs might be easily accessible measures for cumulative MTX exposure and therefore could have a potential as a biomarker for MTX treatment if a predictable correlation with pharmacodynamic (PD) outcomes exists.

Intracellular uptake of MTX takes place via a transporter known as reduced folate carrier (RFC). Inside the cells, the enzyme folylpolyglutamate synthetase (FPGS) adds glutamate moieties to the molecule in a stepwise manner, resulting in the sequential formation of polyglutamated MTX metabolites, MTXGluX, where X stands for the total number of glutamate moieties in the molecule. It needs to be noted that the parent MTX molecule itself already contains one moiety of glutamate and is therefore also referred to

as MTXGlu1 in this thesis. Up to four additional molecules of glutamate are added to MTXGlu1 inside RBCs, resulting in the metabolites MTXGlu2 up to MTXGlu5. Another enzyme, γ-glutamyl hydrolase (γGH), removes terminal glutamate moieties from MTXPGs. The monoglutamated form MTXGlu1 is then removed from the cells via efflux pumps such as multi drug resistance transporters (MDRT). The affinity of these efflux pumps to transport MTX decreases with increasing chain length (highest for MTXGlu1, lowest for MTXGlu5), and all MTXGluX accumulate intracellularly [99]. This accumulation seems to play a major role in the activity of MTX as MTXPGs are active metabolites [95,107]. Several studies have shown a correlation between disease control and MTXPG concentrations, mainly MTXGlu3 and the long-chain polyglutamates MTXGlu4 and MTXGlu5 [114-116]. However, these findings were not verified in a more recent cross-sectional study [117].

A better understanding of the accumulation kinetics of MTXPGs inside RBCs might help to get a better description of the relationship between MTXPG concentrations in RBCs and disease control. Yet, MTXPG kinetics in RBCs have rarely been described using mathematical models in literature so far. One approach by Dalrymple *et al.* used an empirical exponential model to describe the PK of MTX and its polyglutamates measured in RBCs [186]. In their study, each patient was modelled individually using a two-stage approach and the metabolites were treated as independent compounds. Large variability in the results between the individual patients was found.

5.3. Objectives

The objective of this work was to develop a population PK model for low-dose MTX and its polyglutamated metabolites measured in RBCs as first step towards a full PKPD model for MTX.

To achieve this objective, a top-down modelling approach is applied to clinical data resulting in an empirical model on the basis of a classical compartmental PK analysis.

5.4. Methods

5.4.1. The Data

5.4.1.1. Patients

Clinical data was available from two studies conducted in Christchurch, New Zealand, between 2005 and 2009 (ACTRN 012606000275561) [186,187]. Both studies were approved by the Upper South B Regional Ethics committee, and all patients gave written informed consent.

The first study (in the following referred to as "oral study") included patients with RA who either started ("starters") or stopped ("stoppers") weekly low-dose oral MTX treatment [186]. Patients stopping MTX had been on MTX for at least three months prior to being enrolled in the study. All patients were followed for a minimum of 24 weeks. Clinical data, MTX dosing data and MTXGluX measurements were available from 18 patients, equally divided in nine starters and stoppers, respectively. One of the stoppers restarted MTX treatment during the time course of the study. (Note that this is the same data that was analysed previously by Dalrymple *et al.* using the empirical model [186].)

The second study ("sc study") was conducted by the same group of researchers and involved patients with RA on a stable weekly dose of oral MTX treatment that were switched to sc administration at the beginning of the study due to insufficient disease control and/or adverse effects under oral treatment [187]. In this study one patient started immediately on sc MTX without prior oral MTX treatment. Again, all patients were followed for at least 24 weeks. Data from a total of 30 patients were available for the analysis presented here (one of which ceased MTX treatment during the time course of the study).

In both studies, clinical assessment of the patients was undertaken at study begin (baseline at week 0) and weeks 8, 16 and 24. Standard patient demographics and laboratory data (including mean corpuscular volume (MCV) and haemoglobin (Hb) concentrations) were recorded at these visits together with the assessment of disease activity and adverse effects associated with

MTX treatment. Table 5.1 shows the demographics of patients included in this analysis. All patients received concomitant folic acid supplementation of 5 mg per week, administered on a weekday different from the MTX dosing.

Table 5.1: *MTX treatment schedule and patient characteristics* (*mean ± SD*) *in the oral [186] and sc study [187], together with the polled data used in the analysis here.*

Patient characteristics	oral study	sc study	pooled
Number of individuals	18	30	48
Starters	9	1	10
Stoppers	9 [a]	1	10
Continuous treatment	-	28	28
Weekly MTX dose range [mg]	5 – 20	10 – 20	5 – 20
M:F	7:11	7:23	14:34
Age [years]	58.1 ± 8.8	53.1 ± 9.7	55.0 ± 9.6
Weight [kg]	75.2 ± 13.5	78.4 ± 16.6	77.2 ± 15.4
Height [cm]	172.2 ± 6.6 [b]	162.2 ± 31.1	164.7 ± 27.3
Hb [mg/L]	128.3 ± 17.4	129.7 ± 11.4	129.2 ± 13.8
MCV [fL]	90.7 ± 7.0	93.1 ± 4.5	92.2 ± 5.6

[a] one patient restarted MTX treatment later on during the study
[b] values for height missing for eight individuals in the oral study

5.4.1.2. Blood sampling and analysis

In patients starting MTX treatment, trough samples (within 36 hours prior to MTX dosing) were collected weekly for eight weeks, followed by fortnightly sampling until a stable maintenance dose of MTX was achieved. Thereafter samples were collected every four weeks until at least week 24, or until the patient withdrew from the study. For stoppers in the oral study and all patients in the sc study weekly (trough) samples were obtained until week 8, followed by fortnightly sampling until week 16, and then every four weeks until at least week 24, or until withdrawal from the study.

MTXGluX were measured in duplicate samples using high-performance liquid chromatography as described previously [186], separately measuring the

concentration of each MTXGluX in packed RBCs. The average of the raw duplicate MTXGluX measurements, not normalised for RBC count, were used for the PK modelling analysis. Initial runs were conducted using the non-averaged observations with an adjusted error model accounting for correlation between the replicates but no advantage was seen in parameter estimation and due to excessive run times this was abandoned.

The LOQ of the analytical assay was 5 nmol/L_{RBCs} for all MTXGluX. Data below LOQ (BLQ data) was censored and reported as 0 in the data sets, without distinction from the limit of detection, which in turn was not reported. For a limited number of samples actual measurements of the BLQ data could be retrieved. The remaining BLQ data remained reported as BLQ. The percentage of BLQ data was 16%, 13%, 11%, 23% and 45% for measurements of MTXGlu1 to MTXGlu5, respectively.

5.4.2. Structural model development

5.4.2.1. Plasma PK model for MTX

No plasma MTX concentrations were measured in either of the studies available for analysis. Therefore, MTX plasma PK was assumed to follow the population PK model previously published by Hoekstra *et al.* [91]. These authors investigated the differences in bioavailability (F) of MTX after oral *versus* sc administration in patients with RA using higher weekly doses of MTX (25 – 40 mg) than in these studies here (5 – 20 mg). They found a two-compartment model with absorption lag times for both routes of administration best described their data. F was assumed to be 100% for sc administration (F_{sc} = 1), while the median F_{oral} was estimated in relation to F_{sc} as 64%, yet with a wide range (21 – 96%). The authors did not report whether this was dose-dependent. The mean population parameter estimates for the remaining parameters are given in Table 5.2. These values were used to predict the population mean plasma concentration-time profiles of MTX in the patients analysed here, not accounting for any between subject variability (BSV) or residual unexplained variability (RUV) in the plasma PK. To account for the

known dose-dependency of oral MTX bioavailability, F_{oral} was set to 0.7 as reported by Herman *et al.* [88]. In contrast to Hoekstra *et al.,* this study used a lower oral MTX dose (10 mg/m^2 ≈ 17.5 mg/week) which is similar to the dose used in the oral study of this analysis (median oral dose 12.5 mg/week).

Table 5.2: *Population mean parameter estimates for the plasma MTX model published by Hoekstra et al. [91], oral bioavailability* F_{oral} *as published by Herman et al. [88].*

Parameter	Population mean	Description
ka_{oral} [hr^{-1}]	0.87	oral absorption rate constant
$tlag_{oral}$ [hrs]	0.36	lag time for oral absorption
F_{oral}	0.70	oral bioavailability (based on [88]**)**
ka_{sc} [hr^{-1}]	0.36	sc absorption rate constant
$tlag_{sc}$ [hrs]	0.06	lag time for sc absorption
F_{sc}	1.0	sc bioavailability
V_1 [L]	9.6	central volume of distribution
CL_1 [L/hr]	8.4	clearance of MTX from plasma
k_{12} [hr^{-1}]	0.81	intercompartmental transfer rate constant from central to peripheral compartment
k_{21} [hr^{-1}]	0.55	intercompartmental transfer rate constant from peripheral to central compartment

5.4.2.2. Parent model for MTXGlu1 in RBCs

Intracellular RBC MTXGlu1 kinetics were modelled by adding a third compartment to the plasma PK model published by Hoekstra *et al*. [91], which is not in mass balance with the plasma PK and therefore similar to a hypothetical effect compartment in delayed effect PKPD models [188]. It needs to be noted however, that in this work all observations were in the effect compartment as only measurements of RBC MTXGluX concentrations were available in both studies. In contrast, classical effect compartment models do not usually have any observations in the effect compartment.

First-order kinetics as well as active transport were initially tested for the uptake of MTXGlu1 into RBCs from the central compartment. However, only

the first-order process gave stable estimates and was therefore retained in this analysis. Simple first-order elimination of MTXGlu1 from RBCs was assumed in the parent model. A schematic of the model is shown in Figure 5.1. Structural parameters estimated in the parent model were k_{in}, the rate constant of MTXGlu1 uptake into RBCs, the apparent volume of distribution of MTXGlu1 inside RBCs (V_{Glu1}) and CL_{Glu1}, the clearance of MTXGlu1 from RBCs, while the elimination rate constant from RBCs is given by CL_{Glu1}/V_{Glu1}.

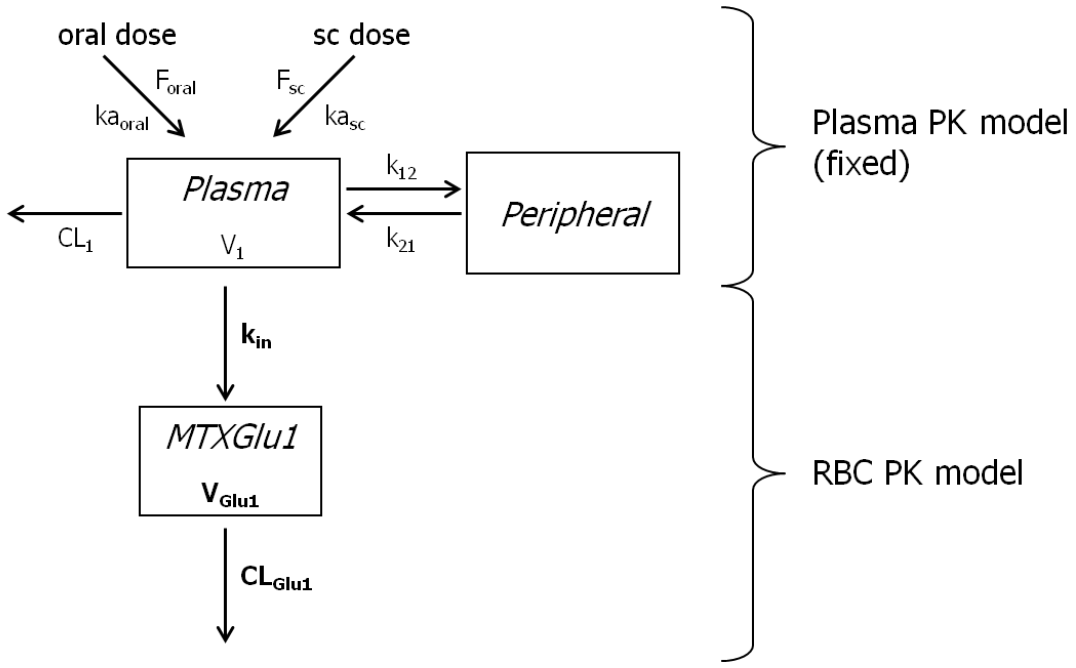


Figure 5.1: *Parent model for MTXGlu1 based on the two-compartment plasma PK model of Hoekstra et al. [91] with an additional compartment for the RBC PK model. Parameters of the plasma PK model are explained in Table 5.2 and were fixed at the population mean values. Parameters for the RBC PK model* k_{in}, V_{Glu1} *and* CL_{Glu1} *were estimated in this analysis. Note that the RBC compartment is not in mass balance with the plasma PK model, similar to an effect compartment in delayed PKPD models.*

5.4.2.3. Parent-metabolite model for MTXGluX in RBCs

The structure of the full parent-metabolite model for MTXGluX was derived by adding an additional catenary RBC compartment for each MTXPG to the parent model, resulting in a total of five catenary RBC compartments (Figure 5.2). The transfer between the compartments was described as first-order reaction with individual rate constants for each step; although these steps are catalysed by FPGS and γGH for the polyglutamation and deglutamation steps, respectively. The volume of distribution parameters (V_{GluX}) were assumed to be correlated for all MTXGluX.

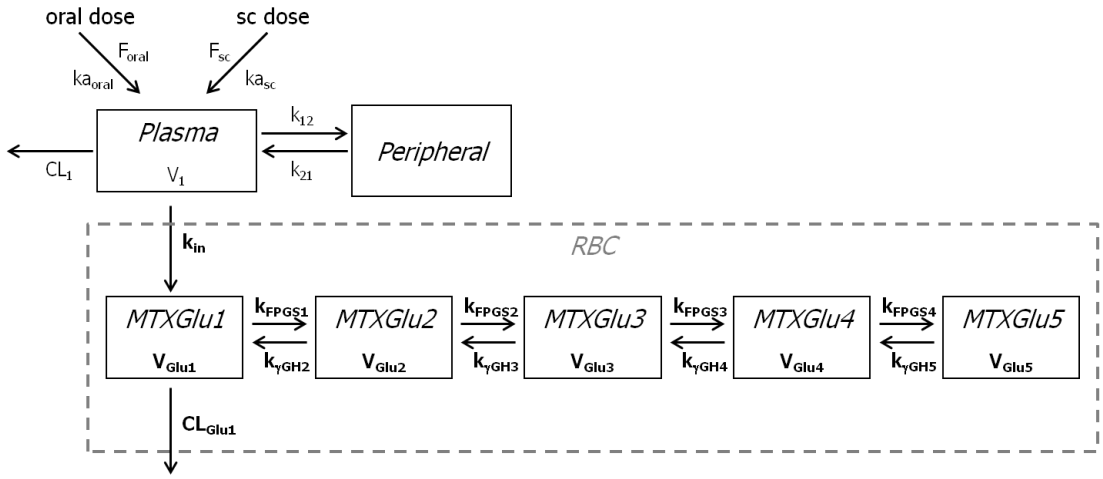


Figure 5.2: *Parent-metabolite model for MTXGluX with a catenary RBC compartment for each MTXGluX. Structural parameters estimated for the full model:* k_{in}, CL_{Glu1}, V_{Glu1}, V_{Glu2}, V_{Glu3}, V_{Glu4}, V_{Glu5}, k_{FPGS1}, k_{FPGS2}, k_{FPGS3}, k_{FPGS4}, $k_{\gamma GH2}$, $k_{\gamma GH3}$, $k_{\gamma GH4}$, *and* $k_{\gamma GH5}$.

5.4.3. Data analysis

NONMEM® version 7.2 (ICON Development Solutions, Ellicott City, MD) was used to fit the data [23]. Due to long run times, Stuart Beal's method M6 [189] was used for handling BLQ data during model development to allow for parameter estimation using the first-order conditional estimation method with interaction (FOCE INT). Using this method the first or last BLQ measurement in the data set in a sequence of decreasing or increasing concentrations is

replaced with a value of LOQ/2, which was equal to 2.5 nmol/L in this analysis, and more extreme values (either before or after this value) were censored from the data set. All parameters in the final parent model and the final parent-metabolite model were re-estimated applying the M3 method for BLQ data which uses the Laplacian method with interaction in NONMEM®, where BLQ data is treated as censored and the joint likelihood of the true and censored observations is computed.

ADVAN5 was used as subroutine in NONMEM® as all processes in the model are described by first-order rate constants. ADVAN5 uses matrix exponentials to analytically solve linear ordinary differential equations and therefore reduces the computational effort compared to time-stepping numerical integration methods [23].

5.4.3.1. Statistical models for between subject and residual variability

BSV on structural parameters was modelled according to Equation 5.1 for the p^{th} parameter.

$$\hat{\theta}_{ip} = \hat{\bar{\theta}}_p \cdot e^{\eta_{ip}}$$

Equation 5.1: *Between subject variability on structural parameter estimates.*

where $\hat{\theta}_{ip}$ is the individual parameter estimate for the i^{th} individual, $\hat{\bar{\theta}}_p$ is the population mean parameter estimate and η_{ip} is the random effect for the i^{th} individual, following a normal distribution with mean = 0 and variance = ϖ_{pp}^2.

RUV was modelled independently for each MTXGluX as a combined error model with:

$$y_{ij}^{\{X\}} = f\left(\hat{\boldsymbol{\theta}}_i, t_{ij}\right) \cdot \left(1 + \varepsilon_{1j}^{\{X\}}\right) + \varepsilon_{2j}^{\{X\}}$$

Equation 5.2: *Statistical model for residual unexplained variability for each MTXGluX.*

where $y_{ij}^{\{X\}}$ is the j^{th} observation for MTXGluX in the i^{th} individual observed at time t_{ij}, $f(\cdot)$ is the structural model, $\hat{\boldsymbol{\theta}}_i$ are the individual parameter estimates of the i^{th} individual, $\varepsilon_{1j}^{\{X\}}$ is proportional error component having the variance σ^2_{prop}GluX, and $\varepsilon_{2j}^{\{X\}}$ is the additive error component with variance σ^2_{add}GluX. The RUV of the individual MTXGluX were assumed to be independent, despite the fact that all metabolite concentrations were measured from a single blood sample at each observation, i.e. the L2 data item function in NONMEM® was not applied here. The L2 function would require the RUV parameters σ^2_{prop}GluX and σ^2_{add}GluX to be coded as random effect parameters that are correlated between the individual MTXGluX, i.e. coded as EPSILONs in a $SIGMA BLOCK with off-diagonal elements for the covariances in the NONMEM® control stream. However, in this analysis the RUV parameters were coded equivalent to fixed effect parameters without correlation, i.e. as THETAs in $PK, to allow for the implementation of the M3 method for handling BLQ data.

5.4.3.2. Stability & Sensitivity analysis for the parent-metabolite model

During the structural model development process the full parent-metabolite model was found not to be globally identifiable (i.e. multiple sets of parameter estimates provided an equally good fit to the data) due to its high number of parameters and catenary structure. Subsets of the structural parameters were subsequently fixed until the resulting reduced model was found to be stable. Stability was assessed as successful convergence to the same parameter estimates from at least two sets of different initial estimates.

In addition, a sensitivity analysis on the parameter estimates was performed to assess local identifiability of the parameter estimates for the reduced model. Absolute and normalised sensitivity were calculated according to Equation 5.3 for each parameter estimate $\hat{\phi}_p \in [\hat{\boldsymbol{\theta}}, \boldsymbol{\Omega}, \sigma^2]$ in the reduced model.

$$S = \frac{\Delta OFV}{\Delta \phi_p} \quad \text{where} \quad \begin{cases} \Delta \phi_p = \delta \cdot \hat{\phi}_p - \hat{\phi}_p \\ \Delta OFV = OFV\left(\delta \cdot \hat{\phi}_p\right) - OFV\left(\hat{\phi}_p\right) \end{cases}$$

$$NSI = \frac{\Delta OFV}{\Delta \phi_p} \cdot \frac{\hat{\phi}_p}{OFV\left(\hat{\phi}_p\right)}$$

Equation 5.3: *Sensitivity (S) and Normalised Sensitivity Index (NSI), with* $\hat{\phi}_p$ *= estimate of parameter tested,* δ *= level of change in the parameter estimate.*

An initial change of +10% in each parameter estimate (δ = 1.1) was considered in this analysis, and extended to -10%, +25% and -25% (δ = 0.9, 1.25 and 0.75, respectively) if the model was found to be insensitive to a certain parameter at the +10% change level.

5.4.3.3. Covariate analysis

5.4.3.3.1. Covariate analysis for the parent model

Total body weight (WT), lean body weight (LBW), MCV and Hb concentration ([Hb]) were tested during model building of the parent model as covariates on V_{Glu1} based on biological plausibility or prior evidence in the literature [190]. All covariates were centred based on their median value in the study population according to:

$$\hat{\theta}_{ip} = \hat{\tilde{\theta}}_p \cdot \left(\frac{COV_i}{COV_{median}}\right)^{\beta} \cdot e^{\eta_{ip}}$$

Equation 5.4: *Covariate model.*

where COV_i is the individual covariate value in the i^{th} individual and using the following median values COV_{median}: 70 kg for WT, 45 kg for LBW, 130 g/L for [Hb] and 90 fL for MCV. Here, β denotes the estimated coefficient of the covariate, which is estimated as a fixed effect parameter in NONMEM®.

LBW was calculated as proposed by Janmahasatian *et al.* [191] based on WT and body mass index (BMI) according to:

$$LBW = \frac{9270 \cdot WT}{6680 + 216 \cdot BMI} \quad \text{for males}$$

$$LBW = \frac{9270 \cdot WT}{8780 + 244 \cdot BMI} \quad \text{for females}$$

Equation 5.5: *Formulas to calculate lean body weight according to [191].*

Height (HT) was missing for eight individuals in the oral study for the calculation of BMI in Equation 5.5. Therefore, a multiple linear regression of height against sex and weight at baseline in the remaining 40 patients was used for a single value imputation of height in these eight individuals. The linear regression was derived as (see Appendix A.5.1):

$$HT = 0.21 \cdot WT + 19.69 \cdot SEX - 0.14 \cdot WT \cdot SEX + 150.25$$

Equation 5.6: *Imputation of height based on weight and sex, where SEX = 0 for females and SEX = 1 for males.*

In both studies, covariates were only recorded during the clinical visits and not at each time point of MTX measurements. Therefore, the last covariate observation was carried forward for MTX observations with missing covariate values occurring after the first clinical visit, while the first covariate observation was carried backward for MTX observations with missing covariates before the first clinical visit.

5.4.3.3.2. Covariate analysis for the parent-metabolite model

Covariates that were found to be significant on V_{Glu1} in the parent model were retested for their significance during model development of the parent-metabolite model. Only covariates that were found to reduce the OFV when applied on all V_{GluX} were included in the final parent-metabolite model.

5.4.3.4. Model selection and evaluation

Model selection was based on the objective function value (OFV) applying the likelihood ratio test (LRT) for nested models and Akaike's Information Criterion (AIC) for non-nested models, as well as reduction of random BSV and RUV. Graphical diagnostics included goodness of fit plots, individual predictions overlaid with the data and individual weighted residual plots.

Non-parametric bootstrap was used to evaluate the parent model. 2000 bootstrap runs were conducted in NONMEM®. The resampling procedure (with replacement) was stratified based on the number of patients that started or stopped MTX or were on a continuous MTX schedule in the original study. Runs that minimized successfully were used to calculate the empirical 95% confidence interval from the 2.5th and 97.5th percentiles. These also included successful minimisations were the covariance step was not successful [192].

The fully reduced parent-metabolite model including the final covariates was also evaluated using non-parametric bootstrap. However, due to the long run times for the parent-metabolite model, only 400 stratified bootstrap runs were conducted in NONMEM®. The mean and standard deviation for the runs with successful minimisation were obtained and used to calculate the asymptotic empirical 95% confidence interval.

5.4.4. Model assessment

5.4.4.1. Predictive performance of the final parent-metabolite model

The predictive performance of the final parent-metabolite model was assessed based on simulations. The time to reach steady state in all MTXGluX was computed for an average individual based on the population mean parameter estimates and compared to the findings of Dalrymple *et al.* [186], as well as the postulated normal survival of RBCs. For the simulation, continuous treatment with an unchanged dose of 10 mg MTX orally once a week over two years was assumed. The MTXGluX profiles over time were also assessed from this simulation based on the ratio of MTXGlu1 to MTXGlu2, 3, 4 and 5.

5.4.4.2. Comparison of MTXGluX kinetics in RBCs with the kinetics observed in other cell lines

The MTXGluX kinetics in RBC described by the final parent-metabolite model and the *in vitro* kinetics observed in human breast cancer cells were directly compared based on the results published by Morrison and Allegra [8]. These authors used a similar catenary model structure and the final parent-metabolite model is based on fixed values for the polyglutamation rate constants (referred to as a group as $k_{FPGS1-4}$ in the following) obtained from this publication.

In addition, an indirect comparison with *in vivo* MTXGluX kinetics observed by Panetta *et al.* [90] in acute lymphoblastic leukaemia is provided.

5.5. Results

5.5.1. Parent model for MTXGlu1 in RBCs

The parameter estimates for base and final model for the parent MTXGlu1 are shown in Table 5.3. These estimates were obtained using the M3 method to handle BLQ data. The corresponding NONMEM® code is given in Appendix A.5.2.

Table 5.3: *Parameter estimates for the base and final parent model for MTXGlu1 and bootstrap (BS) results for the final parent model.*

Parameter	Estimate	Shrinkage [%] [a]	Median (BS) [b]	Empirical 95% CI (BS) [b,c]
Base model				
k_{in} [hr^{-1}]	1.06×10^{-4}	-		
CL_{Glu1} [L/hr]	5.79×10^{-4}	-		
V_{Glu1} [L]	0.278	-		
BSV k_{in} [%]	70.5	12.2		
BSV CL_{Glu1} [%]	8.1	88.8		
BSV V_{Glu1} [%]	120.3	4.7		
CV_{prop} [%]	21.3	-		
σ_{add} [nmol/L]	3.42	-		
Epsilon shrinkage [%]	-	8.3		
Final model with covariates on V_{Glu1}				
k_{in} [hr^{-1}]	1.35×10^{-4}	2.5	1.40×10^{-4}	9.09×10^{-5} – 2.00×10^{-4}
CL_{Glu1} [L/hr]	7.05×10^{-4}	-	7.00×10^{-4}	4.90×10^{-4} – 1.03×10^{-3}
V_{Glu1} [L]	0.287	8.9	0.286	0.184 – 0.434
BSV k_{in} [%]	63.7	-	63.6	47.8 – 87.2
BSV V_{Glu1} [%]	110.1	-	110.1	71.2 – 139.2
CV_{prop} [%]	20.3	8.9	20.3	16.3 – 24.8
σ_{add} [nmol/L]	3.58		2.43	1.32 – 5.21

[a] η-shrinkage given for BSV parameters, ε-shrinkage given together with the RUV parameters CV_{prop} and σ_{add}.
[b] Non-parametric bootstrap statistics based on 1126 runs with successful minimisation out of 2000 runs.
[c] Empirical 95% confidence interval (CI) constructed from the 2.5th and 97.5th percentiles.

The combination of LBW, [Hb] and MCV was found to have a significant covariate effect on V_{Glu1} in the final parent model resulting in a reduction in the OFV of -16.93. However, estimation of the covariate coefficient β was not supported and this parameter was fixed to 1. This covariate is biologically plausible as the product of LBW, [Hb] and MCV can be regarded as an approximation of the true total volume of RBCs (true V_{RBCs}) circulating in the body as shown in Equation 5.7 where A_{Hb} is the total amount of Hb in the body.

$$\text{with} \quad MCV = \frac{\text{true}\, V_{RBCs}}{RBC_{count}} \quad \text{and} \quad [Hb] = \frac{A_{Hb}}{V_{blood}}$$

$$\text{if} \quad V_{blood} \propto LBW \quad \text{then} \quad [Hb] \propto \frac{A_{Hb}}{LBW} \quad \text{and} \quad A_{Hb} \propto [Hb] \cdot LBW$$

$$\text{if} \quad RBC_{count} \propto A_{Hb} \quad \text{then} \quad MCV \propto \frac{\text{true}\, V_{RBCs}}{A_{Hb}} \propto \frac{\text{true}\, V_{RBCs}}{[Hb] \cdot LBW}$$

$$\text{therefore} \quad \text{true}\, V_{RBCs} \propto MCV \cdot [Hb] \cdot LBW$$

Equation 5.7: *Biological plausibility of covariates on* V_{Glu1}.

By dividing all covariates by their median value, the total volume of RBCs is normalised for an average individual, and the estimated parameter V_{Glu1} describes the fractional difference between the normalised total volume of RBCs and the apparent volume of distribution of MTXGlu1.

CL_{Glu1} was associated with a small value for BSV (8.1%) and large shrinkage (88.8%) in the base model. Removing BSV for this parameter from the model decreased the OFV by an additional –1.81. Therefore, BSV on CL_{Glu1} was not included in the final model for MTXGlu1.

Figure 5.3 shows the goodness of fit plot of the observations *versus* the individual predictions. ε-shrinkage was found to be 8.9% (see Table 5.3) [193]. Good agreement between observations and predictions is achieved for the lower concentration range (<80 nmol/L), while higher concentrations show a larger deviation from the line of identity. However, no systematic trend of model misspecification is observed in the goodness of fit plots and the individual weighted residual plots (not shown).

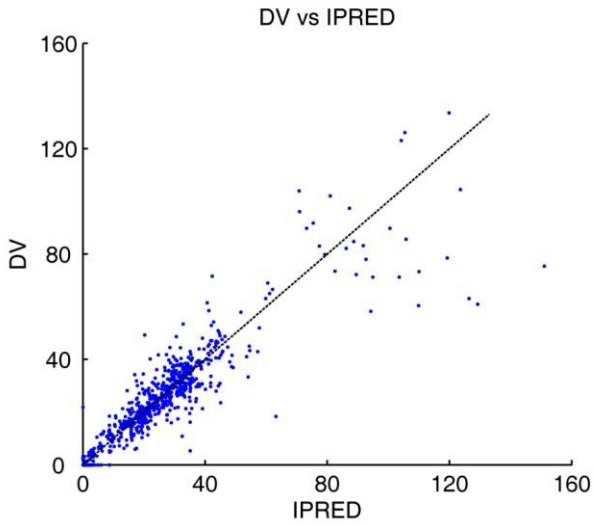


Figure 5.3: *Goodness of fit plot for the final parent model for MTXGlu1: Observations* (*dependent variable DV*) *versus individual predictions* (*IPRED*).

Individual fits obtained from the final model for six patients representing the typical profiles seen in the data set are given in Figure 5.4, while the fit for all individuals is shown in Appendix A.5.3. The model is able to capture the trend in the data for all individuals.

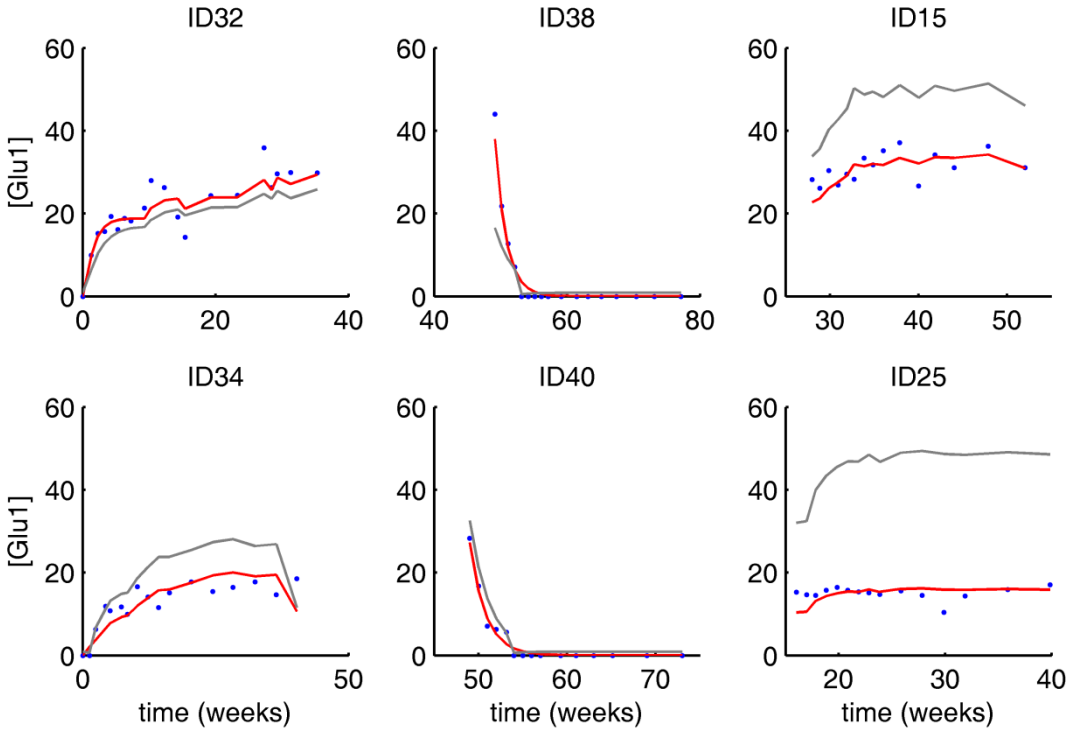


Figure 5.4: *Individual fits for six patients (ID) representing the typical profiles seen in the data set obtained with the final parent model for MTXGlu1. Left panels starters (ID32 & ID34), middle panels stoppers (ID38 & ID40), right panels patients on continuous treatment (ID15 & ID25). Blue dots: observations, red line: individual prediction, grey line: population mean prediction.*

The non-parametric bootstrap results for the final parent model are also presented in Table 5.3. The median and the 95% confidence interval for all parameters were calculated based on 1126 runs that minimised successfully out of a total of 2000 runs (success rate 56.3%). The median values are in close agreement with the population mean estimates for all parameters.

5.5.2. Parent-metabolite model for MTXGluX in RBCs

As previously mentioned, convergence issues became apparent during the structural development of the parent-metabolite model, and the full model was found not to be globally identifiable. It was therefore necessary to simplify the model by fixing structural parameters to achieve successful convergence and stability. In the following section the results for the final reduced model will be presented.

5.5.2.1.1. Structural model development, Stability & Sensitivity analysis

Stability of the parent-metabolite model was achieved by fixing all volume of distribution parameters (V_{GluX}) and polyglutamation rate constants $k_{FPGS1-4}$. Any arbitrary value could be chosen for these fixed values; however meaningful values are desirable that are either based on prior knowledge from literature or previous analyses. The volume of distribution parameters were therefore fixed to 0.3 L, the rounded value estimated for V_{Glu1} in the parent model, while the polyglutamation rate constants were fixed to values reported by Morrison and Allegra in 1987 [8] based on *in vitro* experiments in human breast cancer cells, where a similar catenary compartment model was used for data analysis (Table 5.4).

Table 5.4: *Fixed parameter values in the reduced parent-metabolite model.*

Parameter	V_{GluX} [a]	k_{FPGS1} [b]	k_{FPGS2} [b]	k_{FPGS3} [b]	k_{FPGS4} [b]
Value	0.3 L	0.171 hr^{-1}	0.344 hr^{-1}	0.097 hr^{-1}	0.141 hr^{-1}

[a] Fixed based on the mean population estimate for the parent drug MTXGlu1 only.
[b] Fixed based on Model I published by Morrison and Allegra, 1987 [8].

BSV was supported in the reduced model on the structural parameters k_{in}, CL_{Glu1}, V_{Glu1}, V_{Glu2}, V_{Glu5}, and $k_{\gamma GH3-5}$, but not on $V_{Glu3\&4}$, $k_{FPGS1-4}$ or $k_{\gamma GH2}$. The combined error model was retained for all MTXGluX to describe RUV.

The parameter estimates for the basic reduced parent-metabolite model using the M6 method to handle BLQ data and the results for the sensitivity analysis are provided in Table 5.6. The model was found to be sensitive to all

estimated parameters. Sensitivity was the lowest for BSV on $k_{\gamma GH5}$ using a ±10% in the estimated parameter value, but increased when a ±25% change was considered.

Repeating the estimation for the best structural parent-metabolite model using the M3 method did not change the parameter estimates significantly, and the M6 method was used again for the subsequent analyses due to its considerably shorter run times.

5.5.2.1.2. Covariate analysis

Based on the reduced parent-metabolite model, the significance of the covariates MCV, [Hb] and LBW on the volume of distribution parameters V_{GluX} was reassessed. Table 5.5 provides the corresponding OFV all covariate combinations tested.

Table 5.5: *Reassessment of covariates on* V_{GluX}.

Covariates on V_{GluX}	**OFV**	**change in OFV**
none	15438.69	
MCV, LBW & [Hb]	15476.49	37.8
MCV & LBW	15433.09	-5.6
MCV & [Hb]	15473.82	35.13
MCV	15430.96	-7.73
LBW & [Hb]	15467.05	28.36
LBW	15441.30	2.61
[Hb]	15463.74	25.05

Based on these results, only MCV was included as covariate on V_{GluX} in the final reduced parent-metabolite model. Table 5.7 shows the parameter estimates for the corresponding final reduced parent-metabolite model. The NONMEM® code for the final reduced parent-metabolite model is given in Appendix A.5.4.

Table 5.6: *Parameter estimates and sensitivity analysis for the basic reduced parent-metabolite model without covariates on* V_{GluX}.

Parameter	Estimate	ΔOFV	S	NSI
OFV	15438.69		δ = 1.1	
k_{in} [hr^{-1}]	2.21×10^{-4}	0.73	32895.33	4.70×10^{-4}
CL_{Glu1} [L/hr]	1.45×10^{-3}	0.79	5481.99	5.14×10^{-4}
$k_{\gamma GH2}$ [hr^{-1}]	0.1736	2.31	133.18	1.50×10^{-3}
$k_{\gamma GH3}$ [hr^{-1}]	0.1915	1.28	66.73	8.28×10^{-4}
$k_{\gamma GH4}$ [hr^{-1}]	0.2426	2.25	92.72	1.46×10^{-3}
$k_{\gamma GH5}$ [hr^{-1}]	0.2984	1.08	36.33	7.02×10^{-4}
BSV k_{in}	0.5122	0.16	3.14	1.04×10^{-4}
BSV CL_{Glu1}	0.4111	0.13	3.19	8.49×10^{-5}
BSV V_{Glu1}	0.1058	0.06	5.86	4.02×10^{-5}
BSV V_{Glu2}	0.1005	0.06	5.67	3.69×10^{-5}
BSV V_{Glu5}	0.2296	0.09	3.79	5.64×10^{-5}
BSV $k_{\gamma GH3}$	0.2832	0.14	5.01	9.20×10^{-5}
BSV $k_{\gamma GH4}$	0.1737	0.18	10.31	1.16×10^{-4}
BSV $k_{\gamma GH5}$	0.0625	7.00×10^{-3}	1.12	4.53×10^{-6}
CV^2_{prop} Glu1	0.0346	0.57	165.66	3.71×10^{-4}
σ^2_{add} Glu1	35.34	1.10	0.31	7.14×10^{-4}
CV^2_{prop} Glu2	0.0460	1.44	313.83	9.35×10^{-4}
σ^2_{add} Glu2	7.32	0.32	0.44	2.09×10^{-4}
CV^2_{prop} Glu3	0.0142	0.71	500.20	4.60×10^{-4}
σ^2_{add} Glu3	39.15	1.10	0.28	7.11×10^{-4}
CV^2_{prop} Glu4	0.0641	1.51	234.95	9.75×10^{-4}
σ^2_{add} Glu4	3.78	0.29	0.76	1.87×10^{-4}
CV^2_{prop} Glu5	0.0771	0.97	125.35	6.26×10^{-4}
σ^2_{add} Glu5	2.28	0.14	0.63	9.39×10^{-5}
			δ = 0.9	
BSV $k_{\gamma GH5}$	0.0625	8.00×10^{-3}	-1.28	-5.18×10^{-6}
			δ = 1.25	
BSV $k_{\gamma GH5}$	0.0625	0.04	2.82	1.14×10^{-5}
			δ = 0.75	
BSV $k_{\gamma GH5}$	0.0625	0.05	-3.14	-1.27×10^{-5}

Table 5.7: *Parameter estimates for the final reduced parent-metabolite model and bootstrap (BS) results with MCV as covariate on* V_{GluX}.

Parameter	Estimate	Shrinkage [%] [a]	Mean (BS) [b]	Empirical standard deviation (BS) [b]	Asymptotic empirical 95% CI (BS) [b,c]
k_{in} [hr^{-1}]	2.27×10^{-4}	-	2.31×10^{-4}	2.43×10^{-5}	1.83×10^{-4} – 2.79×10^{-4}
CL_{Glu1} [L/hr]	1.48×10^{-3}	-	1.54×10^{-3}	2.03×10^{-4}	1.14×10^{-3} – 1.93×10^{-3}
V_{Glu1} [L]	0.3 fixed	-	-	-	-
V_{Glu2} [L]	0.3 fixed	-	-	-	-
V_{Glu3} [L]	0.3 fixed	-	-	-	-
V_{Glu4} [L]	0.3 fixed	-	-	-	-
V_{Glu5} [L]	0.3 fixed	-	-	-	-
k_{FPGS1} [hr^{-1}]	0.171 fixed	-	-	-	-
k_{FPGS2} [hr^{-1}]	0.344 fixed	-	-	-	-
k_{FPGS3} [hr^{-1}]	0.097 fixed	-	-	-	-
k_{FPGS4} [hr^{-1}]	0.141 fixed	-	-	-	-
$k_{\gamma GH2}$ [hr^{-1}]	0.174	-	0.173	9.33×10^{-3}	0.155 – 0.191
$k_{\gamma GH3}$ [hr^{-1}]	0.192	-	0.190	0.0154	0.159 – 0.220
$k_{\gamma GH4}$ [hr^{-1}]	0.243	-	0.243	0.0151	0.213 – 0.272
$k_{\gamma GH5}$ [hr^{-1}]	0.299	-	0.301	0.0202	0.262 – 0.341
BSV k_{in} [%]	69.4	10.0	68.1	9.1	51.8 – 87.9
BSV CL_{Glu1} [%]	63.4	14.1	61.8	10.5	43.2 – 85.9

Table 5.7: *Continued.*

Parameter	Estimate	Shrinkage [%] [a]	Mean (BS) [b]	Empirical standard deviation (BS) [b]	Asymptotic empirical 95% CI (BS) [b,c]
BSV V_{Glu1} [%]	31.8	27.5	31.0	6.5	19.5 – 47.0
BSV V_{Glu2} [%]	32.3	25.1	30.5	13.5	5.65 – 115.4
BSV V_{Glu5} [%]	48.2	24.1	41.4	12.8	8.79 – 151.8
BSV $k_{\gamma GH3}$ [%]	53.3	8.8	51.2	5.2	41.4 – 62.6
BSV $k_{\gamma GH4}$ [%]	41.7	4.5	40.5	4.6	32.2 – 50.3
BSV $k_{\gamma GH5}$ [%]	24.3	61.6	28.9	16.7	0.49 – 484.9
CV_{prop} Glu1 [%]	18.3	3.7	16.6	4.10	13.7 – 22.2
σ_{add} Glu1[nmol/L]	5.96		6.17	0.60	4.82 – 7.44
CV_{prop} Glu2 [%]	21.5	4.0	21.1	2.42	16.1 – 25.9
σ_{add} Glu2[nmol/L]	2.69		2.75	0.43	1.93 – 3.67
CV_{prop} Glu3 [%]	11.8	4.0	11.6	1.42	9.59 – 14.6
σ_{add} Glu3[nmol/L]	6.29		6.43	0.65	4.82 – 7.59
CV_{prop} Glu4 [%]	25.4	3.7	25.9	3.09	18.8 – 31.9
σ_{add} Glu4[nmol/L]	1.95		2.05	0.29	1.44 – 2.54
CV_{prop} Glu5 [%]	27.9	3.9	27.6	3.47	20.3 – 33.7
σ_{add} Glu5[nmol/L]	1.50		1.31	0.64	0.23 – 2.51

[a] η-shrinkage given for BSV parameters, ε-shrinkage for each MTXGluX given together with the corresponding RUV parameters CV_{prop} and σ_{add} GluX.
[b] Non-parametric bootstrap statistics based on 100 runs with successful minimisation out of 400 runs.
[c] Asymptotic empirical 95% confidence interval (CI) constructed based on mean and empirical standard deviation of the 100 successful BS runs.

5.5.2.1.3. Model evaluation

Goodness of fit plots for all five MTXGluX are shown in Figure 5.5. As was seen for the parent drug, good agreement between observations and individual predictions is achieved for the lower concentrations, while higher concentrations show a somewhat larger deviation from the line of identity. However, no systematic model misspecification is seen in these plots. For all five MTXGluX, ε-shrinkage was ≤4% (see Table 5.7) [193].

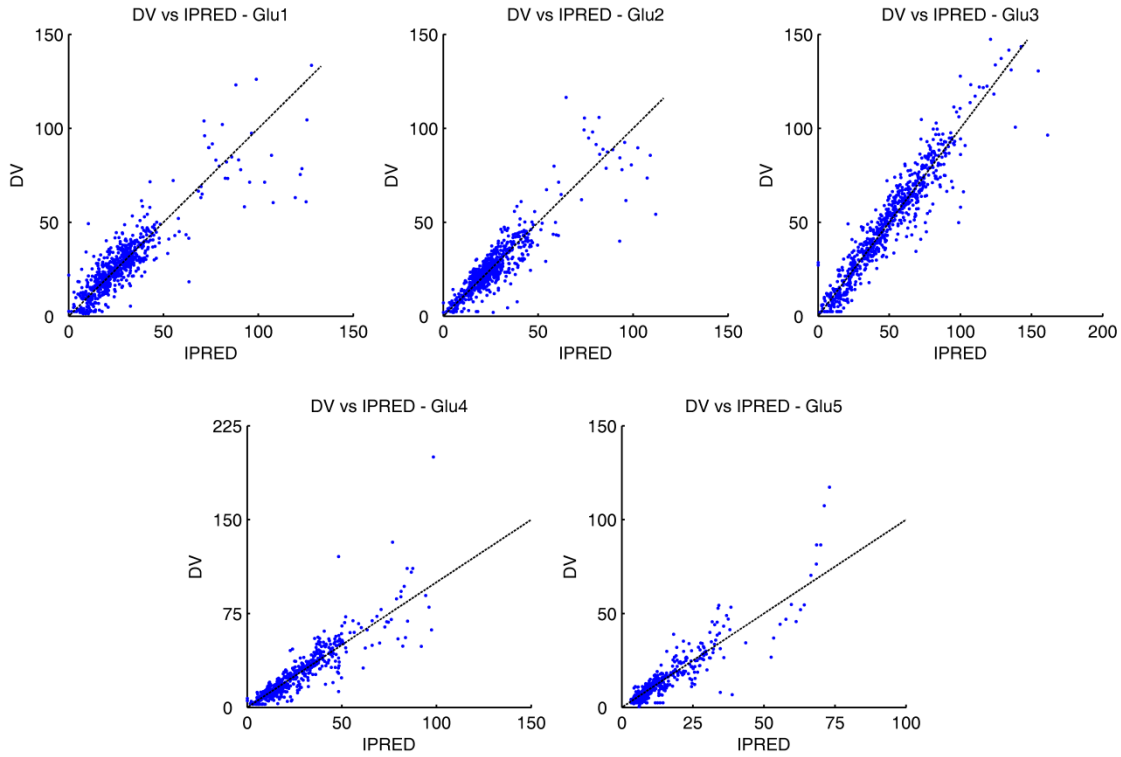


Figure 5.5: *Goodness of fit plots for all five MTXGluX obtained with the final reduced parent-metabolite model. Observations (DV) versus individual predictions (IPRED).*

following page:

Figure 5.6: *Individual fits for all five MTXGluX obtained with the final reduced parent-metabolite model for six patients: ID32 & ID34 starters, ID38 & ID40 stoppers, ID15 & ID25 patients on continuous treatment. Blue dots: observations, red line: individual prediction, grey line: population mean prediction.*

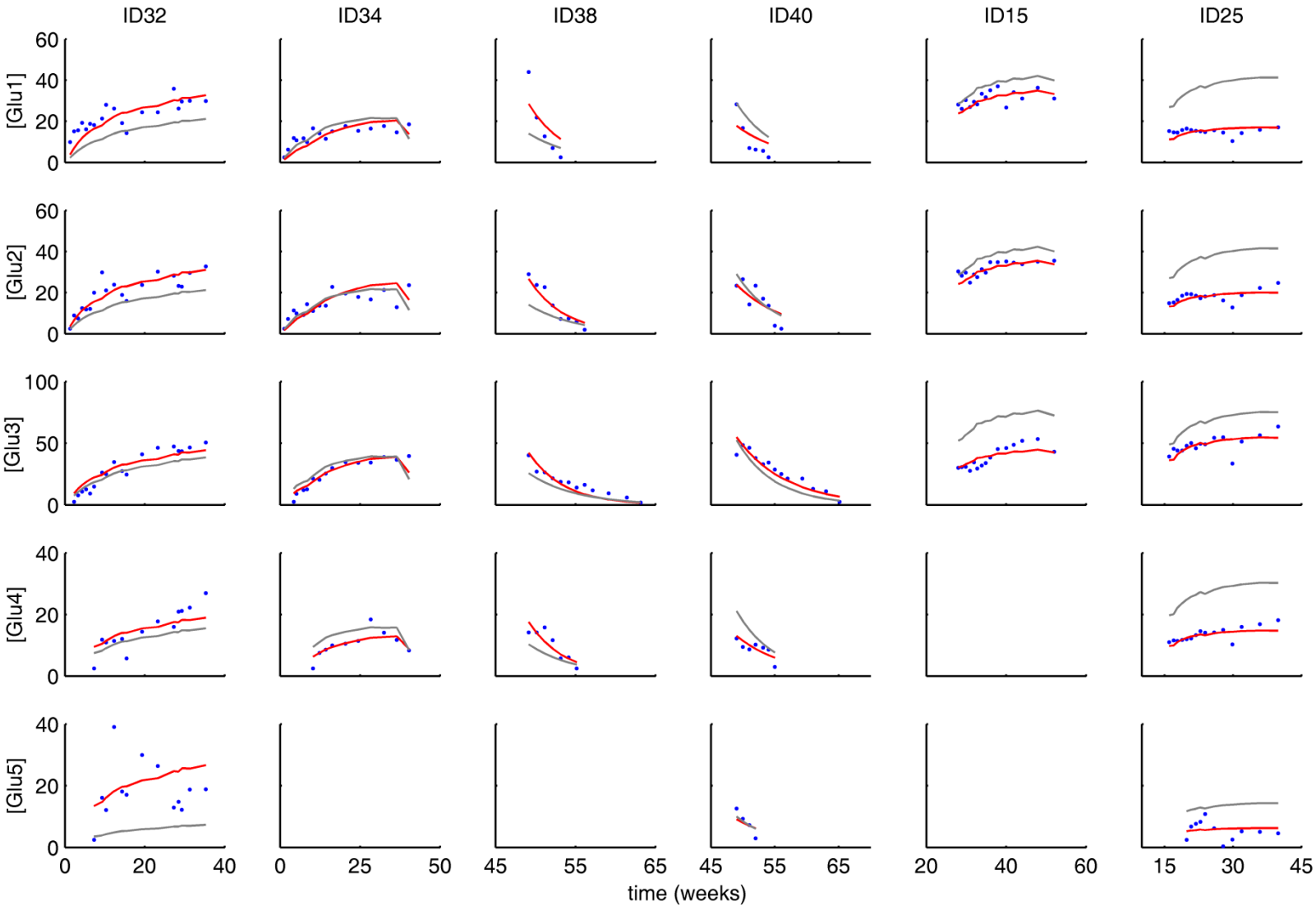

ID32
ID34
ID38
ID40
ID15
ID25
[Glu1]
[Glu2]
[Glu3]
[Glu4]
[Glu5]
time (weeks)

Figure 5.6 shows the individual fits obtained from the final reduced parent-metabolite model in the same six patients as were shown for the parent model. It can be seen in these plots that long chain MTXPGs (MTXGlu4 and MTXGlu5) are not measurable in some individuals, e.g. ID15, ID34 and ID38 in this figure. The fits for all patients are shown in Appendix A.5.5. The final parent-metabolite model shows reasonably good agreement with the data for all MTXGluX and in all individuals. However, there remains unexplained variability in some individuals (note for example ID25 where it appears that there was some period of non-compliance around week 30); yet it was not possible to elucidate either a mechanism or clinical reason for this.

The bootstrap results for the final reduced parent-metabolite model are also provided in Table 5.7. Out of the 400 conducted runs 100 minimized successfully (success rate 25%). The mean parameter values $\bar{\phi}_p$ and empirical standard deviations *s* of these 100 runs were calculated and used to construct the empirical asymptotic empirical 95% confidence interval (CI) according to:

$$95\%\,\mathrm{CI} = \bar{\phi}_p \pm 1.96 \cdot s$$

Equation 5.8: *Formula to calculate the upper and lower bound of the asymptotic empirical 95% confidence interval (CI) of the bootstrap parameter estimates, with an α-error level of 0.05.*

Good agreement between the estimates for the final model and the bootstrap results is seen, although several of the random effect parameters in the final model fall slightly outside of the 95% confidence interval, e.g. BSV on $V_{Glu4\&5}$ and the deglutamation rate constant $k_{\gamma GH5}$.

5.5.3. Model assessment

5.5.3.1. Predictive performance of the final parent-metabolite model

In the original analysis of the data in the oral study by Dalrymple *et al.* [186] the authors calculated the time to reach steady state in all five MTXGluX (Table 5.8). Their calculations were based on the estimated half-life of accumulation that was determined using a monoexponential accumulation model for each patient starting MTX treatment in the study. They defined the time to reach steady state as the time that was required to achieve 90% of the maximum concentration observed in each individual.

Table 5.8: *Calculated time to reach 90% of the steady state concentration for MTXGluX in the study by Dalrymple et al. [186].*

Time to reach steady state	Median [weeks]	Range [weeks]
MTXGlu1	6.2	0.0 – 13.9
MTXGlu2	10.6	7.0 – 77.2
MTXGlu3	41.2	19.8 – 66.7
MTXGlu4	149	16.2 – 831.6
MTXGlu5	139.8	15.5 – 264.0

Using the population estimates for the final parent-metabolite model, the RBC MTXGluX kinetics for a typical patient on a stable oral dose of 10 mg MTX per week were simulated over a period of two years. The time to reach 90% of the steady state concentration and true steady state in all MTXGluX was calculated from the simulated data for comparison. In this analysis the time to reach 90% of the true steady state concentration did not differ between the MTXGluX. It was found to be 14 weeks for all, while true steady state concentrations were observed after 45, 60, 63, 57 and 70 weeks for MTXGlu1 to MTXGlu5, respectively. These values notably exceed the postulated average lifespan of RBCs of 120 days or approximately 17 weeks.

The ratios between the measured concentrations of the individual MTXPGs and MTXGlu1 were also computed at each sampling time point in

this simulation and are shown together with the true steady state concentrations for each MTXGluX in Table 5.9. The ratios were found to be stable throughout the entire simulation time course with only a marginal fluctuation of ±0.001 nmol/L.

Table 5.9: *Ratios of individual MTXGluX concentrations and true steady state concentration ([nmol/L]) for an average individual.*

	MTXGlu1	**MTXGlu2**	**MTXGlu3**	**MTXGlu4**	**MTXGlu5**
ratio	1	1.007	1.824	0.734	0.348
[nmol/L]	14.38	14.49	26.23	10.56	5.00

As can be seen from Table 5.9 the intracellular RBC concentrations of MTXGlu1 and MTXGlu2 are fairly similar, while MTXGlu3 is the most abundant metabolite in the RBCs. The long chain polyglutamates MTXGlu4 and MTXGlu5 have a lower relative concentration compared to the parent drug, with MTXGlu5 being the least abundant metabolite. Furthermore, the true steady state concentration of MTXGlu5 in an average individual on a stable dose of 10 mg MTX once weekly is just 5.0 nmol/L, which is equivalent to the LOQ of the analytical MTXGluX assay. This explains why MTXGlu5 was not detectable in a number of patients in this study.

5.5.3.2. Comparison of MTX kinetics in RBCs with the kinetics observed in other cell lines

5.5.3.2.1. RBC kinetics versus in vitro kinetics in human breast cancer cells

Morrison and Allegra [8] used a similar catenary model to describe the *in vitro* kinetics of MTX in human breast cancer cells. They estimated naïve pooled parameter values for the polyglutamation and deglutamation steps. Here, the rate constants determined for the polyglutamation steps were used as fixed values for $k_{FPGS1\text{-}4}$. It is therefore possible to directly compare the deglutamation rate constants estimated for the RBC kinetics ($k_{\gamma GH2\text{-}5}$) with the corresponding values published by these authors (Table 5.10):

Table 5.10: *Comparison of deglutamation rate constants [hr^{-1}] observed in vitro by Morrison and Allegra [8] for human breast cancer cells and estimated for RBCs in this work based on the final parent-metabolite (PM) model.*

Source	$k_{\gamma GH2}$	$k_{\gamma GH3}$	$k_{\gamma GH4}$	$k_{\gamma GH5}$
Morrison & Allegra	0.637	0.114	0.065	0.122
Final PM model	0.174	0.192	0.243	0.299

As can be seen from this table there is a relative difference in the rate constants between cancer cells and RBCs. The observed deglutamation rates are two to four times faster in RBCs than in the breast cancer cells for all reaction steps other than the last step from MTXGlu2 to MTXGlu1, which Morrison and Allegra found to be much faster than was estimated here for RBCs. In addition, using both the published rate constant values for poly- as well as deglutamation and estimating the volume of distribution terms in the parent-metabolite model was examined, but did not provide a satisfactory fit to the RBC data. Thus, MTXGluX kinetics in RBCs seem to differ considerably from those observed in the cancer cell line.

However, direct comparison of these results is difficult. Morrison and Allegra also accounted for binding of MTXGluX to DHFR, which is not incorporated here in the RBC parent-metabolite model, and they used Michaelis-Menten kinetics to describe the active uptake of MTXGlu1 into the cells. In addition, the MTXGluX concentrations measured in the cancer cell experiments are much higher than in RBCs (µmol/L rather than nmol/L).

5.5.3.2.2. RBC kinetics versus in vivo kinetics in lymphocytes

Panetta *et al.* [90] developed a model for the *in vivo* accumulation of MTXPGs in acute lymphoblastic leukaemia, were patients receive a high dose regimen of MTX (0.25 – 8 g/m^2 body surface area as intravenous infusion over 24 hours). MTXGluX concentrations in their study were measured in T-lymphocytes as well as B-lymphocytes, and were found to considerably exceed RBC concentrations as well (again µmol/L rather than nmol/L). Their final

model included only two intracellular compartments, one for MTXGlu1 and the other comprising all metabolites, i.e. the sum of the measured MTXGlu2 to MTXGlu5 concentrations. Active and passive uptake into the cells was incorporated in the model, where the rate constant for the passive uptake was fixed and only the active component was estimated based on Michaelis-Menten kinetics. Efflux was considered to occur only for MTXGlu1 and described by first-order kinetics. The polyglutamation step was estimated as active process, while the deglutamation reaction was assumed to follow first-order kinetics.

Due to these major differences between the Panetta model and the RBC model developed here, a direct comparison is not possible. However, during the initial model development of the RBC model, the model structure used by Panetta *et al.* was also considered. For this, the MTXGlu2 – MTXGlu5 concentrations in RBCs were summed to describe the cumulative concentration of MTXPGs in the second intracellular compartment. Yet, it was impossible to apply this model structure successfully to the RBC data in NONMEM®. The program terminated due to integration errors when the full data set was used, while successful minimisation could only be achieved for a data set containing only the patients on continuous MTX treatment, e.g. the steady state data in the RBC study. This model structure seems incapable of adequately describing the accumulation and disappearance kinetics of MTXGlu1 and MTXPGs inside and from RBCs, which leads to the conclusion that the kinetic profiles observed in RBCs must differ considerably from those observed by Panetta *et al.* in lymphocytes.

5.6. Discussion

In this chapter, previously published data of MTXGluX concentrations measured in RBCs were used to develop an empirical population parent-metabolite model for low-dose MTX treatment in RA using a top-down modelling approach. The model was developed in the style of a classical compartmental PK analysis with only loose guidance based on mechanistic principles.

5.6.1. Parent model for MTXGlu1 in RBCs

Initially, a model for the parent drug MTXGlu1 only was developed. The model described the data very well, and good agreement between the final parameter estimates and the bootstrap results were found.

Plasma concentrations for MTXGlu1 were predicted based on a previously published model. Only population mean parameter values could be used for this prediction which does not allow for variability in the plasma PK. The PK inside RBCs was then described by adding an additional compartment to the plasma PK model. The RBC compartment is not assumed to be in mass balance with the plasma PK model, similar to an effect compartment in a delayed effect PKPD model. This assumption is justified, as the rate constant of uptake into RBCs k_{in} is only 1.35×10^{-4} hr^{-1} compared to the rate of elimination from the plasma, $k_{el} = V_1/CL_1 = 9.6\ L / 8.4\ L \cdot hr^{-1} = 1.14\ hr^{-1}$. Due to this almost 10,000 fold difference, the uptake of MTXGlu1 into RBCs has only a marginal influence on the plasma concentration and can be neglected in mass balance. However, the BSV parameter on k_{in} is likely inflated as it also takes into account the lack of variability in the predicted plasma concentrations, and it is possible that k_{in} is biased itself. In future studies, measuring MTX plasma concentrations in addition to RBC concentrations would be valuable to obtain a better estimate of k_{in} and its BSV.

Although the uptake of MTX into cells is facilitated by an active transport via the reduced folate carrier, it was assumed to follow first-order kinetics in this analysis. This approximation is justified as the peak MTX plasma

concentrations (C_{max}) under low-dose treatment are well below the reported K_m value for the transporter (C_{max} = 700 nmol/L for a 20 mg sc dose of MTX which is the highest dose in this study, compared to K_m values between 3,900 and 8,200 nmol/L reported for various tumour cell lines [93]). In addition, active transport into RBCs was considered during initial model development but did not provide a better fit to the data than the simpler model assuming first-order kinetics.

LBW, MCV and [Hb] were found in combination to yield a significant covariate on V_{Glu1} and it was shown that these covariates are also physiologically plausible. As LBW already takes into account structural differences in body composition between males and females no additional gender difference in the median [Hb] value used for normalisation was supported in this analysis.

CL_{Glu1} in the parent model accounts for elimination of MTXGlu1 from the cells as well as metabolism to MTXGlu2. BSV on this composite parameter was not supported in the final model. In addition, the input of MTXGlu1 via the deglutamation of MTXGlu2 into the RBC compartment cannot be quantified in the parent model, which potentially leads to bias in the structural parameter estimates.

5.6.2. Parent-metabolite model for MTXGluX in RBCs

The final reduced parent-metabolite model for MTXGluX kinetics in RBCs is able to describe the observed data; however the model development process was hampered by identifiability issues with the full parent-metabolite model. The catenary structure of the model and the high number of estimated parameters results in a highly flexible model with multiple solutions and hence lack of consistent convergence. Fixing of several structural parameters was required to ensure local identifiability of the model. The apparent volume of distribution of all MTXGluX (V_{GluX}) was set to the rounded value that was estimated in the parent model, while the polyglutamation rate constants ($k_{FPGS1-4}$) were fixed to literature values. Any arbitrary values could have been

chosen and the estimated parameters for the deglutamation steps would have changed accordingly; however the chosen values allow for a comparison of the results with the parent model as well as the previous model published by Morrison and Allegra [8]. Nevertheless, the estimated values in the final reduced parent-metabolite model do not have any direct physiological meaning as the true values for the polyglutamation rate constants ($k_{FPGS1-4}$) and V_{GluX} in RBCs are unknown.

All processes in this model were assumed to follow first-order kinetics, although the poly- as well as deglutamation reactions are catalysed by FPGS and γGH, respectively, and the uptake and loss of MTXGlu1 are facilitated by active transport. Yet, the MTXGluX concentrations measured inside RBCs are well below the K_m values for all these enzymes and the transporter and the first-order approximation is therefore assumed to be valid. This approximation also allows using the ADVAN5 subroutine in NONMEM®, which reduced the run times considerably.

Model development was conducted using the M6 method to handle BLQ data and the final reduced parent-metabolite model was then re-analysed with the M3 method. The use of M3 did not result in a significant change in the final estimates for the parent-metabolite model, despite the large percentage of data reported as BLQ especially for the long chain MTXPGs. However, as BLQ data was not distinguished from data below the limit of detection in the data set, the actual percentage of BLQ data could be much less. In addition, the performance of the different methods to handle BLQ data has been assessed for classic mammillary compartmental models [189,194-196], yet not for a catenary model structure, and especially not in a situation where reversible formation of metabolites occurs and where a large number of parameters are fixed *a priori*. It is therefore unknown what differences in performance are to be expected.

At each time of observation all MTXGluX concentrations were measured in a single blood sample. This means that a correlation in the residual error between these measurements is expected, for example as the process noise introduced during sample taking and handling will be the same. In

NONMEM® such a correlation is usually accounted for by using the L2 data item function. However, the L2 function requires RUV parameters to be coded as random effects (i.e. as EPSILONs), while the M3 method to handle BLQ data requires these parameters to be coded equivalent to fixed effect parameters (i.e. as THETAs). To account for error correlation between the individual MTXGluX measurements within a blood sample while using the M3 method, a covariance between the fixed effect parameters that describe RUV would need to be included in the code. Yet, this is very difficult to implement in NONMEM® for fixed effect parameters, especially if they are not associated with random effects (i.e. when they do not have a corresponding $OMEGA BLOCK in NONMEM®). As this analysis did not aim to quantify the magnitude of the error correlation and the ability of using the M3 method to handle BLQ data was regarded as having a higher priority, such a complex coding procedure was not explored here. For simplification it was instead assumed that the RUV of the different MTXGluX are independent. However, if future studies confirm that using the M3 method for handling BLQ data has truly no benefit compared to simpler methods under the current model structure, the L2 data item function should be used in future work.

The final parent-metabolite model was evaluated using non-parametric bootstrap. Due to the long run times only 400 bootstrap runs were conducted and the success rate was only 25%. The main reason for unsuccessful minimisation in NONMEM® was rounding errors in the parameter estimates. As only 100 successful bootstrap runs were available for analysis, the means instead of the medians of the parameter estimates were calculated and the standard deviations were used to construct the asymptotic empirical 95% confidence intervals around the mean estimates. These results showed reasonable good agreement with the final estimates for the original data set.

Comparing the parameter estimates for MTXGlu1 with those obtained with the parent model, it can be seen that the structural parameters CL_{Glu1} and k_{in} both have increased. BSV on CL_{Glu1} was supported in the final parent-metabolite model, yet not in the parent model. BSV on V_{Glu1} dropped

considerably from >100% in the parent model to 32% in the parent-metabolite model, while BSV on k_{in} increased only slightly. The proportional error component was approximately the same (18% *versus* 20%), yet the additive error increased from 3.58 nmol/L to 5.96 nmol/L in the parent-metabolite model. This is also reflected by a less good fit observed for MTXGlu1 to the individual data as was obtained in the parent model only, especially when looking at the elimination of MTX in the stoppers, e.g. ID38 and ID40 in Figure 5.4 and Figure 5.6.

In the final parent-metabolite model only MCV retained a significant effect as covariate on all V_{GluX}. This is in contrast to the parent model were MCV, LBW and [Hb] had a significant effect on the apparent volume of distribution. It is possible that fixing the population mean value for V_{GluX} impedes the covariate analysis in the parent-metabolite model, especially as BSV was also only found to be significant on three of the five volume of distribution parameters (V_{Glu1}, V_{Glu2} and V_{Glu5}).

5.6.3. Model assessment

The final parent-metabolite model was assessed based on simulations. The time required to reach 90% of the true steady state concentration for an average individual was compared to the analysis conducted by Dalrymple *et al.* [186] on a subset of the same data set. These authors found the required time to be highly variable between different individuals and also between the individual MTXGluX, with median values well above 100 weeks for MTXGlu4 and MTXGlu5. In contrast, the population analysis conducted here led to an equal time to reach 90% steady state of all MTXGluX which was found to be just 14 weeks due to interdependence between the metabolites. The time to reach true steady state on the other hand varied between the different MTXGluX: 45 – 70 weeks. These values correspond to 315 to 490 days and are much longer than any of the values reported for the average RBCs lifespan in humans. It is therefore impossible for this accumulation to occur within a single RBC during its circulation in the body. It is more likely that the actual

accumulation takes part in the stem cells and/or RBC precursor cells in the bone marrow and that over time reticulocytes with increasing MTX content are released into the circulation until steady state is achieved in the bone marrow cell populations. This theory is in accordance with suggestions by Schalhorn *et al.* [197]. In their study, MTX concentrations in RBCs declined for approximately 24 - 48 hours after administration of a 6 hour infusion of high-dose MTX, but the concentrations started to rise again after day 4 - 7. They attributed this second increase in RBC MTX concentrations to the release of young RBCs that were preloaded with MTX and its metabolites during their production in the bone marrow.

It should be noted that if MTX accumulation truly takes place on the level of pluripotent stem cells in the bone marrow, then long-lasting post MTX treatment consequences on fertility might be expected if a similar long-term accumulation also takes place in gametocytes. In this case, a MTX free interval prior to conception that is longer than the three to six months that are currently recommended by manufacturers would be appropriate. This was also suggested in a review of clinical studies on the effect of MTX on pregnancy and fertility by Lloyd *et al.* [198]. Nevertheless, further clinical studies are required to determine whether MTX accumulates in bone morrow precursor cells after their commitment to form a specific blood cell line or on the stage of pluripotent stem cells, and also in gametocytes.

Nevertheless, the kinetics observed in RBCs were found to differ from those observed in other cell lines. The rate constants for deglutamation in the final parent-metabolite models do not match with those obtained by Morrison and Allegra [8] for human breast cancer cells *in vitro*, and the structural model developed for leukocytes by Panetta *et al.* [90] is not able to adequately describe the concentration-time course in RBCs observed during the elimination of MTX in the starters and stoppers. These differences can be explained by the limited enzyme capacity in RBCs compared to nucleated cells as well as the inability of RBCs to perform mitosis. These physiological disparities are likely to result in altered intracellular MTX kinetics in RBCs compared to other cell lines.

Overall, these findings lead to the conclusion that the MTX kinetics observed in RBCs differ considerably from the kinetics observed in other cell lines, and do not to reflect the concentration-time profile of MTX in a single RBC, but rather a cross-section of the concentration in RBCs of different ages in a random blood sample.

5.6.4. Further analyses and hypotheses testing

The developed parent-metabolite model for MTX in RBCs can provide a basis to test further hypotheses. For example, in the following chapter of this thesis, alternative structural models will be evaluated which if supported would provide insight into alternative mechanisms of MTX accumulation in RBCs.

Literature suggests that γGH is able to cleave two moieties of glutamate at a time [97,199]. This alternate cleaving mechanism can be incorporated into the structural model in a future analysis despite the need to fix $k_{FPGS1\text{-}4}$ and V_{GluX}. Furthermore, loss of MTXPGs can also be incorporated into the structural model as was done by Morrison and Allegra [8]. If supported, such a loss could either be attributed to the destruction of RBCs or active transport of MTXPGs out of the cells.

In addition, genotypic covariates can be incorporated in to the model and tested for significance. PK relevant polymorphisms have been reported for γGH as well as the influx and efflux transporters of MTX [200,201]. However, the clinical relevance of pharmacogenetic covariates for MTX has not yet been established conclusively [106]. The developed parent-metabolite model can be used to test whether covariates that possibly affect the intracellular PK of MTX show a significant effect.

Lastly, this PK model can form the basis for a full PKPD model in future work which can then be used to test the suitability of MTXGluX concentrations measured in RBCs as biomarker for monitoring of low-dose MTX treatment in RA. Yet, any possible association between PD outcomes and RBC concentrations has to be seen in the light of the unique PK profile observed in

RBCs and the lack of a causal relationship as RBCs are not on the postulated pathway of action of MTX.

5.7. Conclusion

This is the first time that the PK of MTX and its polyglutamated metabolites in RBCs has been described in a population modelling analysis. Although identifiability issues occurred during the model development process, a stable parent-metabolite model could be developed by fixing selected parameter values. The resulting reduced model is able to describe the observed data and can be used for further hypotheses testing before being applied in a full PKPD model of low-dose MTX in the treatment of RA in the future.

It is notable that the kinetics of MTX measured in RBC differ considerably from the kinetics in other cell lines and likely represent a cross-section of RBC with different MTX concentrations over time rather than the intracellular kinetics within an individual RBC.

Future work is required to assess in more detail the encountered identifiability issues and the mathematical features of catenary models in general.

Chapter 6: Hypotheses testing for methotrexate pharmacokinetics in red blood cells

This chapter will be part of a publication that will be submitted as:

Korell J, Stamp L, Duffull S *et al.* (2012) *A population pharmacokinetic model for methotrexate measured in red blood cells*.

6.1. Synopsis of the Chapter

In the previous chapter, a population pharmacokinetic (PK) model for methotrexate (MTX) and its polyglutamated metabolites (MTXPGs) measured in red blood cells (RBCs) was developed. In this chapter several hypotheses related to the intracellular PK of MTX will be tested based on this model, and the model will be updated as required based on the results.

6.2. Introduction

In the following section four different hypothesis tests are described that are related to the intracellular PK of MTX in RBCs. The first two hypothesis tests are with respect to the structure of the model. They concern the observed deglutamation mechanism and a potential loss of MTXPGs from RBCs. The third and fourth hypothesis tests consider the influence of covariates, genotypic as well as phenotypic, on the PK of MTX in RBCs and in the plasma.

6.2.1. Hypothesis test 1: Cleaving mechanism of γ-glutamyl hydrolase

The intracellular metabolism of MTX involves two enzymes: folylpolyglutamate synthetase (FPGS) and γ-glutamyl hydrolase (γGH). FPGS adds additional glutamate moieties to MTX in a stepwise manner resulting in MTXPGs, i.e. MTXGlu2 to MTXGlu5. Here, MTXGluX stands for a specific polyglutamated metabolite containing a total of X glutamate moieties in the molecule. γGH removes terminal glutamate moieties from these metabolites. For the development of the original population PK model for MTXGluX in the previous chapter it was assumed that γGH cleaves one moiety of glutamate at a time. Yet, *in vitro* experiments by Rhee *et al.* [97] suggest that the enzyme is able to cleave two terminal glutamate moieties simultaneously and that this cleaving mechanism might in fact be preferred.

This leads to four alternative hypotheses on the structural model for the intracellular PK of MTX:

1) γGH cleaves only one terminal glutamate moiety as assumed in the original model (Model A).
2) γGH cleaves only two terminal glutamate moieties, without conversion of MTXGlu2 to MTXGlu1 (Model B).
3) γGH preferentially cleaves two terminal glutamate moieties, but allows for the conversion of MTXGlu2 to MTXGlu1 by cleaving one glutamate for this step (Model C).
4) γGH cleaves one and two terminal glutamate moieties (Model D).

For a graphical representation of these hypotheses based on the proposed model please refer to Figure 6.1 in the corresponding Methods section (Section 6.4.1).

6.2.2. Hypothesis test 2: Loss of MTXPGs from RBCs

The structure of the original model developed in the previous chapter only allows the parent drug MTXGlu1 to leave the cells. This assumption is supported by the findings by Zeng *et al.* [99] who showed that the transporters that facilitate the efflux of MTXGlu1 have very poor transport capacity for MTXPGs.

Yet, Morrison and Allegra found in their modelling analysis of *in vitro* MTX PK data measured in human breast cancer cells that a model that accounts for loss of MTXGlu3, MTXGlu4 and MTXGlu5 (in the following referred to as MTXGlu3-5) from the cells provided a better fit to the data than a model that only allowed MTXGlu1 to leave the cells [8]. Loss of MTXGlu2 was assumed to occur as well in these experiments, but the corresponding rate constant of loss from the cell was not found to be identifiable in their model. The authors simplified their model further by using a single rate constant to describe the loss of MTXGlu3-5. These results suggest that MTXPGs are able to leave cancer

cells and it can be hypothesised that a similar efflux might also be possible in other cell lines such as RBCs.

In addition, RBCs have a finite survival time. The removal of individual RBCs from the circulation at the end of their lifespan means that their MTX content is lost from the total RBC MTX pool in the circulation. If the same rate constant is able to describe the loss of all MTXPGs in the system, this loss might be attributed to RBC death rather than efflux of MTXPGs.

These findings lead to three alternative hypotheses, each with two sub-hypotheses (a & b), that will be tested against the original model in this chapter:

1) All MTXGluX are lost from the cells, where the loss of all individual MTXGluX is described by:
    a. a different rate constant $k_{out,GluX}$.
    b. the same rate constant $k_{out,Glu1}$.
2) Besides MTXGlu1, MTXGlu3-5 are lost from the cells (but not MTXGlu2), where the loss of MTXGlu3-5 is described by:
    a. a different rate constant $k_{out,GluX}$. (This is a similar to Hypothesis 1a.)
    b. the same rate constant $k_{out,Glu3-5}$, but which is different from $k_{out,Glu1}$.

    This hypothesis follows the structure of the model developed my Morrison and Allegra [8].
3) Besides MTXGlu1, only the long chain MTXPGs (MTXGlu4&5) are lost from the cells, where the loss of MTXGlu4&5 is described by:
    a. a different rate constant $k_{out,GluX}$. (This again is similar to 2a and 1a).
    b. the same rate constant $k_{out,Glu4\&5}$, but which is different from $k_{out,Glu1}$.

A graphical representation of the proposed model structure including loss of all MTXGluX is provided in the corresponding Methods section (Figure 6.2, Section 6.4.2).

6.2.3. Hypothesis test 3: Significance of genotypic covariates on MTXGluX PK in RBCs

Recently, there has been an increasing interest in genotypic covariates that might be suitable to explain some of the large variability seen in MTXGluX concentrations and clinical outcomes of MTX treatment in rheumatoid arthritis (RA) [108]. A variety of candidate genes and polymorphisms have been identified within the MTX and folate pathway that are relevant to the PK and/or effect of MTX treatment. A number of studies report associations between individual polymorphisms and MTX efficacy, adverse effects and/or RBC MTXGluX concentrations. These were recently reviewed by Stamp and Roberts [106]. However, none of the investigated polymorphisms have been proven to be a suitable biomarker for MTX treatment so far. This is partially due to a lack of conclusive evidence from the association studies as replicate studies often fail to reproduce initially observed correlations between the polymorphisms and clinical outcomes of MTX treatment and/or RBC MTXGluX concentrations. For example, Dervieux *et al.* found a significant association between polymorphisms affecting the reduced folate carrier (RFC) and γGH [114,200]. Yet, these findings could not be verified in a study by Stamp *et al.* [202].

Polymorphisms relevant to the intracellular PK of MTX have been reported for γGH as well as the influx and efflux transporters including RFC and multi drug resistance transporters (MDRTs) of the ATP-binding cassette (ABC) family, respectively. These include single nucleotide polymorphisms (SNPs) as well as copy number variations (CNVs).

A SNP occurs when a single nucleotide is exchanged in the DNA coding sequence of the corresponding gene. For example the SNP denoted rs1051266 occurs in the gene SLC19A1, which codes RFC. Here, the guanine (G) in

position 80 is exchanged for an adenine (A), and the variant is therefore denoted as 80G>A. For each SNP two forms, commonly referred to as alleles, of the gene exist (e.g. for rs1051266 having either G or A in position 80), and these two different alleles give rise to three different genotypes: homozygous for allele 1 (e.g. GG = wildtype), homozygous for allele 2 (e.g. AA) or heterozygous (e.g. GA).

A CNV occurs when the number of copies of an allele is altered, e.g. instead of having two alleles of a particular gene (one on each of the corresponding chromosomes), repetitions of the allele in one or both of the chromosomes occur. For example, the gene coding the enzyme γGH has two alleles (denoted as allele 1 and allele 2) resulting from the SNP rs11545078 and multiple copies of both alleles can be present in the genome of an individual.

Table 6.1 provides an overview of the genotypic information that was available for all but two patients (46 out of 48 patients) in the data set used for model building in Chapter 5.

It can be hypothesised that these polymorphisms have an effect on RBC MTXGluX concentrations. This hypothesis will be tested for each polymorphism independently based on the developed population PK model for MTX in RBCs by including a covariate effect on the related PK parameter and testing for its significance as described in the corresponding Methods section (Section 6.4.3). For example, for rs1051266 the covariate effect will be included in the model on the rate of uptake into RBCs (k_{in}) as this SNP affects the gene coding the influx transporter RFC.

Table 6.1: *Polymorphisms relevant to intracellular MTX PK and their frequency in the study population (n = 46). A = adenine, C = cytosine, G = guanine, T = thymine.*

Gene	SNP	Variant	Genotype		Frequency (%) [a]
γ-glutamyl hydrolase					
γGH	rs11545078	CNV	allele 1	1 copy	100
			allele 1	2 copies	0
			allele 2	1 copy	91
			allele 2	2 copies	9
Multi drug resistance transporters					
ABCC1	rs35592	CNV	allele 1	1 copy	98
			allele 1	2 copies	2
			allele 2	1 copy	74
			allele 2	2 copies	26
	rs3784862	CNV	allele 1	1 copy	90
			allele 1	2 copies	10
			allele 2	1 copy	62
			allele 2	2 copies	38
ABCG2	rs17731538	CNV	allele 1	1 copy	98
			allele 1	2 copies	2
			allele 2	1 copy	74
			allele 2	2 copies	26
ABCC2	rs4148396	CNV	allele 1	1 copy	93
			allele 1	2 copies	7
			allele 2	1 copy	48
			allele 2	2 copies	52
	rs2273697	1249G>A	GG		68 [55]
			GA		28 [40]
			AA		4 [5]
ABCB1	rs1045642	3435C>T	CC		24 [20]
			CT		43 [50]
			TT		33 [30]
Reduced folate carrier					
SLC19A1	rs1051266	80G>A	GG		37 [25]
			GA		43 [55]
			AA		20 [20]

[a] Frequency in the European population [in square brackets if available] obtained from: *SNPedia: a wiki supporting personal genome annotation, interpretation and analysis* (http://www.snpedia.com/index.php/SNPedia) [203], accessed on 17.04.2012.

6.2.4. Hypothesis test 4: Significance of phenotypic covariates to predict between subject variability in the plasma PK of MTX

Lastly, the original RBC PK model developed in Chapter 5 is based on a fixed plasma PK model for MTX published by Hoekstra *et al.* [91]. The plasma PK parameters were fixed to the population mean estimates of Hoekstra *et al.* and no random effects on these parameters could be estimated during the model development process. This was due to a lack of plasma samples in the data set used for this analysis as well as the temporal delay between the sampling of RBC PK data and the plasma PK which did not allow recreating the plasma PK from the RBC data. Therefore, no within or between subject variability is accounted for in the plasma PK, although both types of variability are expected to be present.

However, MTX is largely eliminated via the kidneys involving passive glomerular filtration (81%) as well as active secretion [86,112]. Reduced kidney function, i.e. decreased glomerular filtration rate (GFR), is therefore expected to result in a reduced renal clearance of MTX which in turn will result in higher MTX plasma concentrations and hence higher intracellular uptake of MTX. Differences in kidney function between individuals will therefore contribute to the between subject variability (BSV) observed in the MTX RBC concentrations, and this contribution can be predicted by including a suitable biomarker for kidney function as covariate on the plasma clearance parameter (CL_1) in the developed model.

In the data set used for model building the estimated glomerular filtration rate (eGFR) was available as a biomarker for kidney function as calculated according to the Modification of Diet in Renal Disease (MDRD) formula [204] This biomarker will be tested for its significance as a covariate on CL_1 in the developed population PK model as described in the corresponding Methods section (Section 6.4.4).

6.3. Objectives

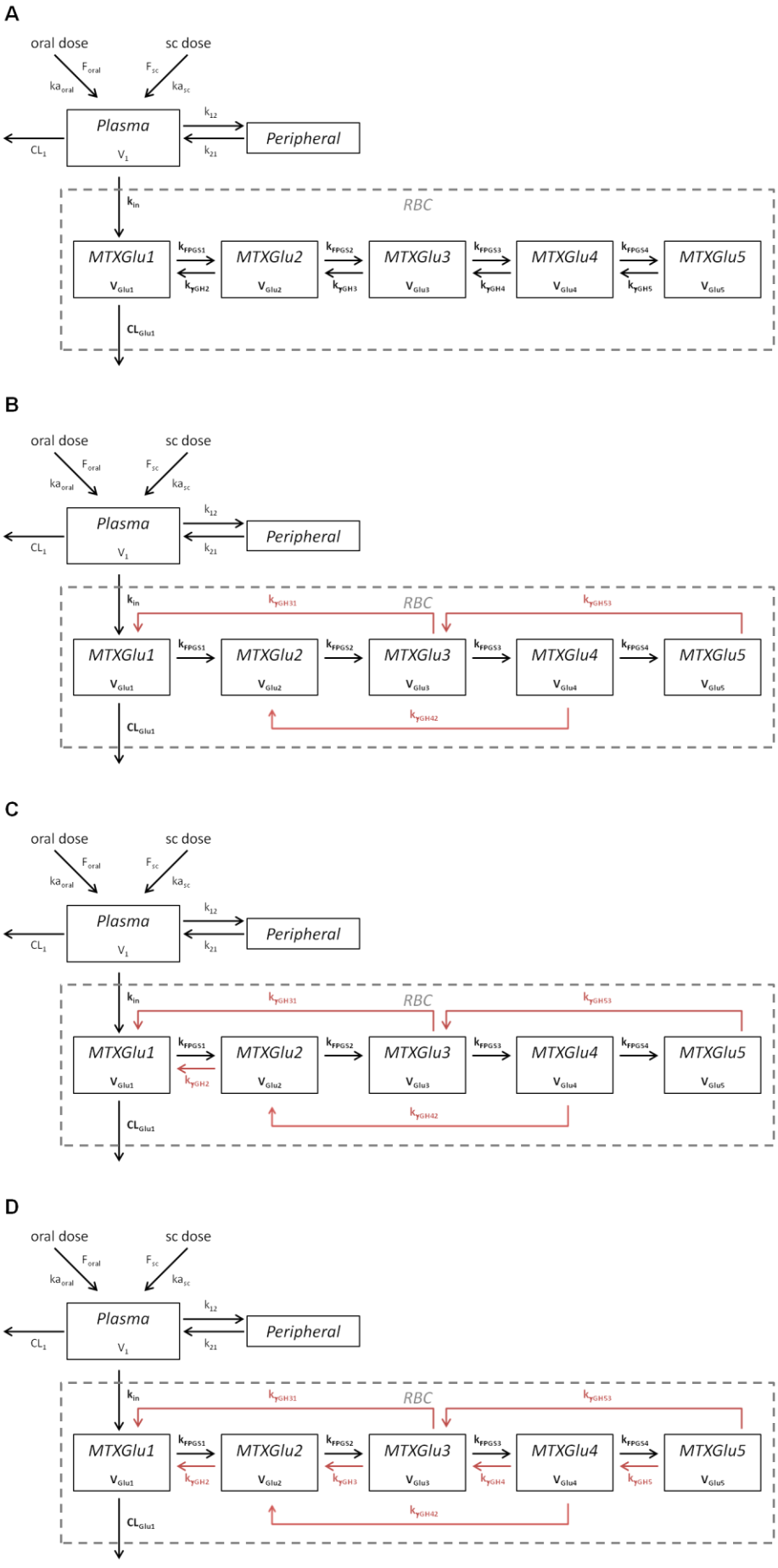
The general objectives of this chapter were to assess the validity of these hypotheses based on the original population PK model for MTX in RBCs that was developed in the previous chapter, and to update the model as required based on the results.

The following questions will be addressed specifically:

1) What is the preferred cleaving mechanism of γGH in RBCs?
2) Are MTXPGs lost from RBCs, either via efflux or due to death of individual RBCs?
3) Do genotypic covariates have significant effects on the intracellular PK of MTXGluX in RBCs?
4) Can phenotypic covariates such as eGFR be used to include predictable BSV in the otherwise fixed plasma PK of MTX in the developed model?

following page:

Figure 6.1: *Structural models tested to assess the preferred cleaving mechanism of γGH.* **A** *Original model with cleaving of one glutamate moiety only,* **B** *simultaneous cleaving of two glutamate moieties without conversion of MTXGlu2 to MTXGlu1,* **C** *simultaneous cleaving of two glutamate moieties with conversion of MTXGlu2 to MTXGlu1, and* **D** *cleaving of one and two glutamate moieties possible.*


A
oral dose
sc dose
Plasma
Peripheral
RBC
MTXGlu1
MTXGlu2
MTXGlu3
MTXGlu4
MTXGlu5
B
oral dose
sc dose
Plasma
Peripheral
RBC
MTXGlu1
MTXGlu2
MTXGlu3
MTXGlu4
MTXGlu5
C
oral dose
sc dose
Plasma
Peripheral
RBC
MTXGlu1
MTXGlu2
MTXGlu3
MTXGlu4
MTXGlu5
D
oral dose
sc dose
Plasma
Peripheral
RBC
MTXGlu1
MTXGlu2
MTXGlu3
MTXGlu4
MTXGlu5

6.4. Methods

6.4.1. Hypothesis test 1: Cleaving mechanism of γ-glutamyl hydrolase

To determine the preferred cleaving mechanism of γGH, three alternative structural models were fitted to the data used for model development in Chapter 5. The alternative model structures together with the structure of the original model are shown in Figure 6.1.

The same estimation procedure as described in the Methods section of Chapter 5 (Section 5.4.3) was applied here, using the FOCE method with interaction in NONMEM® 7.2 and the M6 method to handle data below the limit of quantification.

It was assessed whether any of these models provide a better fit than the original model based on Akaike's Information Criterion (AIC) for non-nested models (the lower the AIC value the better) and the likelihood ratio test (LRT) for nested models. For the addition of one parameter in the model (one additional degree of freedom) the LRT requires a decrease in the OFV of at least -3.84 to be statistically significant at an α-error level of 5%.

Due to the identifiability issues observed during initial model development, the parameters describing the polyglutamation steps by FPGS ($k_{FPGS1-4}$) and the volume of distribution parameters (V_{GluX}) were fixed in the original model as described in Chapter 5. Changing the structural model with respect to the cleaving mechanism of γGH does not resolve the identifiability problems observed for the original model and therefore the same parameters were also fixed in the three alternative models.

6.4.2. Hypothesis test 2: Loss of MTXPGs from RBCs

Testing the potential loss of MTXPGs from RBCs was performed in a similar fashion than testing the preferred cleaving mechanism of γGH. The general structure of the alternative models is shown in Figure 6.2.

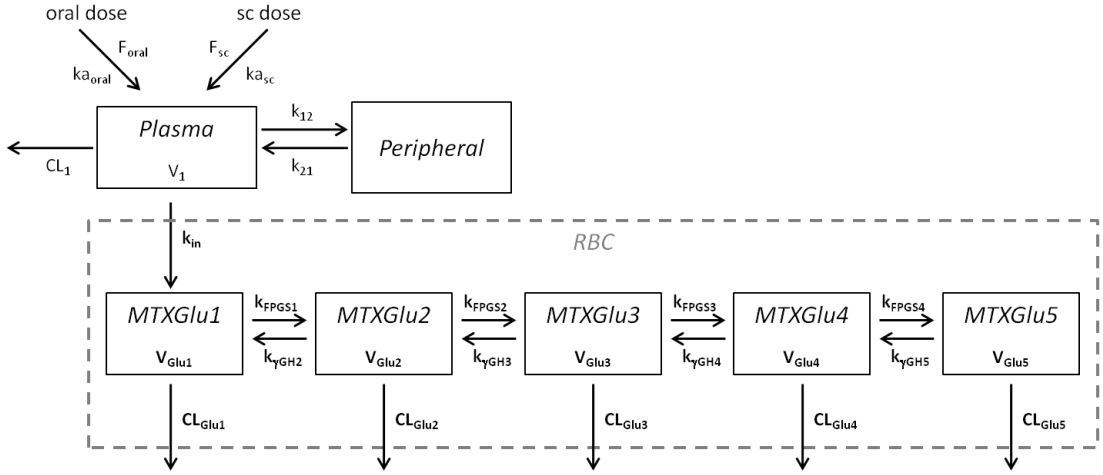


Figure 6.2: *General model structure used for assessing a potential loss of MTXPGs from RBCs in addition to the loss of MTXGlu1, where* $k_{out,GluX} = CL_{GluX}/V_{GluX}$.

This model structure was fitted to the clinical data using NONMEM® and the individual $k_{out,GluX}$ were estimated as CL_{GluX}/V_{GluX}. The polyglutamation rate constants $k_{FPGS1-4}$ and individual V_{GluX} were kept fixed as described in Chapter 5 due to the observed identifiability issues.

For each fit of the six different hypotheses described in the Introduction of this chapter (Section 6.2.2) only the relevant parameters of the general model structure were included in the tested model. This means that for Hypothesis 1a all individual rate constants of loss from RBCs ($k_{out,GluX}$) were estimated, while for Hypothesis 1b only $k_{out,Glu1}$ was estimated and $k_{out,Glu2-5}$ were set to the same value as $k_{out,Glu1}$, i.e. $k_{out,Glu1} = k_{out,Glu2} = k_{out,Glu3} = k_{out,Glu4} = k_{out,Glu5}$. For Hypothesis 2, $k_{out,Glu2}$ was fixed to 0. The remaining four $k_{out,GluX}$ parameters were estimated individually for Hypothesis 2a, while in Hypothesis 2b it was assumed that $k_{out,Glu3} = k_{out,Glu4} = k_{out,Glu5}$ and $k_{out,Glu3-5}$ was allowed to differ from $k_{out,Glu1}$, resulting in two estimated clearance parameters: $k_{out,Glu1}$ and $k_{out,Glu3-5}$. For Hypothesis 3 the same approach as for Hypothesis 2 was applied, yet here

$k_{out,Glu2}$ and $k_{out,Glu3}$ were fixed to 0, resulting in the estimation of $k_{out,Glu1}$ and $k_{out,Glu4\&5}$ for Hypothesis 3b, while in Hypothesis 3a $k_{out,Glu1} = k_{out,Glu4\&5}$.

Note, that Figure 6.2 is shown using the structure of the original PK model where γGH cleaves only one glutamate moiety; however, the best model found from Hypothesis test 1 will be used.

To determine the model that provides the best fit AIC and LRT were used as model discrimination criteria as described before.

6.4.3. Hypothesis test 3: Significance of genotypic covariates on MTXGluX PK in RBCs

As can be seen in Table 6.1, data from eight different polymorphisms relevant to the intracellular PK of MTX were available from the data set used for model building. These eight polymorphisms included three SNPs; two affecting genes coding for MDRT efflux transporters of the ABC family while the third affects the gene coding the influx transporter RFC. The other five polymorphisms are CNVs; four occurring in genes coding MDRT efflux transporters, one coding the deglutamation enzyme γGH.

These polymorphisms were tested individually for their significance as covariates on the corresponding parameters in the population PK model. The six polymorphisms (rs35592, rs3784862, rs17731538, rs4148396, rs2273697, and rs1045642) affecting MDRT efflux transporters were included on all CL_{GluX} in the model (depending on whether these parameters were supported based on the results of the previous hypothesis tests). The SNP rs1051266 affects RFC and was therefore included on k_{in}, while the CNV of rs11545078 affecting γGH was included on all four deglutamation rate constants $k_{\gamma GH2\text{-}5}$ and it was assumed that the covariate affects all four rate constants equally, i.e. the covariate coefficients added to these parameters were assumed to be the same.

For SNPs, the influence of the three different genotypes (homozygous wildtype, heterozygous, and homozygous for minor allele) on the corresponding parameter were coded according to Equation 6.1:

$$\hat{\theta}_{ip} = \left(\hat{\bar{\theta}}_p + a \cdot \beta_1 + b \cdot \beta_2 \right) \cdot e^{\eta_{ip}}$$

where a = 0 and b = 0 for a wildtype individual

a = 1 and b = 0 for a heterozygous individual

a = 0 and b = 1 for an individual homozygous for the minor allele

Equation 6.1: *Covariate model for a SNP.*

Here, $\hat{\theta}_{ip}$ is the individual estimate of the p^{th} parameter in the i^{th} individual, $\hat{\bar{\theta}}_p$ is the population mean parameter estimate, η_{ip} is the random effect for the i^{th} individual, a and b are control parameters equal to 0 or 1 dependent on the genotype of the i^{th} individual, while β_1 and β_2 are the estimated covariate coefficients for an individual heterozygous or homozygous for minor allele, respectively.

Similarly, the influence of a CNV as genotypic covariate on the corresponding parameter was coded according to Equation 6.2:

$$\hat{\theta}_{ip} = \hat{\bar{\theta}}_p \cdot (\beta_1)^a \cdot (\beta_2)^b \cdot e^{\eta_{ip}}$$

where a = 0 and b = 0 for a individual with one copy of both alleles

a = 1 and b = 0 for a individual with two copies of allele 1 and one copy of allele 2

a = 0 and b = 1 for a individual with one copy of allele 1 and two copies of allele 2

a = 1 and b = 1 for a individual with two copies of both alleles

Equation 6.2: *Covariate model for a CNV.*

Genotypic information was lacking for two individuals in the data set. These individuals were assumed to be homozygous for the wildtype of all SNPs and to have one copy of both alleles for the CNVs, i.e. a and b were set to zero in Equation 6.1 and Equation 6.2 for all tested genotypes for these individuals.

The LRT was used to assess the significance of the inclusion of the two estimated coefficients β_1 and β_2 into the model in comparison to the nested model without covariate effects (base model). Two additional parameters results in two additional degrees of freedom and the corresponding critical value for the LRT is 5.99 at an α-error level of 5%. This means that the inclusion of the two additional parameters must result in a decrease in the OFV of at least -5.99 points in comparison to the base model to be statistically significant.

6.4.4. Hypothesis test 4: Significance of phenotypic covariates to predict between subject variability in the plasma PK of MTX

Table 6.2 lists the demographics available in the data set used for model building in Chapter 5 that were tested here as potential covariates to predict BSV in the clearance of MTX from plasma.

Table 6.2: *Covariates* (*mean ± SD*) *with potential influence on the plasma clearance of MTX available in the data set* (*oral study [186], sc study [187]*).

Patient characteristics	oral study	sc study	pooled
Number of individuals	18	30	48
eGFR [ml/min/1.73m^2]	76.7 ± 14.3	81.9 ± 15.1	80.0 ± 14.8
Weight [kg]	75.2 ± 13.5	78.4 ± 16.6	77.2 ± 15.4
Height [cm]	172.2 ± 6.6 [a]	162.2 ± 31.1	164.7 ± 27.3

[a] values for height missing for eight individuals in the oral study

Based on these covariates two alternative covariate models were hypothesised. First, eGFR was tested for its significance as covariate on total MTX plasma clearance CL_1. Yet, the eGFR values available in the data set were calculated according to the MDRD formula that automatically normalises the calculated GFR values to a body surface area (BSA) of 1.73 m^2. To account for this normalisation, the covariate model was corrected for BSA, and both covariates were centred on typical values: 100 ml/min/1.73 m^2 for eGFR and 1.73 m^2 for BSA (Equation 6.3).

$$CL_{1,i} = CL_{1,pop} \cdot \left(\frac{eGFR_i}{100} \cdot \frac{BSA_i}{1.73} \right)^{\beta}$$

Equation 6.3: *Covariate model for BSA adjusted eGFR on total plasma clearance* CL_1.

Here, as well as in the following equations, $CL_{1,i}$ denotes the individual total plasma clearance of the i^{th} individual, $CL_{1,pop}$ is the population mean value of total MTX plasma clearance (fixed to 8.4 L/hr according to Hoekstra *et al.* [91]), and β is the estimated covariate coefficient.

BSA was calculated form total body weight (WT) and height (HT) according to the Mosteller formula [205]:

$$BSA\left[m^2\right] = \sqrt{\frac{WT\,[kg] \cdot HT\,[cm]}{3600}}$$

Equation 6.4: *Mosteller formula to calculate body surface area* (*BSA*) *from total body weight* (*WT*) *and height* (*HT*).

Note that for eight individuals from the oral study height had already been imputed based on weight and sex for the analysis presented in Chapter 5 (see Appendix A.5.1). Here, the imputed values were again used to calculate BSA of these individuals.

It needs to be noted that GFR is a biomarker for renal clearance, yet only 81% of an administered MTX dose is renally cleared [86]. Total MTX clearance also involves hepatic metabolism whose capacity is determined by body size. Therefore, the second covariate model tested the effect of BSA adjusted eGFR as covariate on the renally cleared fraction (rCL_1) only. In addition, allometrically scaled total body weight with a fixed exponent of 0.75 [206] and centred on a typical value of 75 kg was included as additional covariate describing body size on the non-renal fraction ($nrCL_1$). Renal and non-renal clearance of MTX were assumed to be linear independent processes in this model, and it was furthermore assumed that renal clearance of MTX is either affected linearly (A) or non-linearly (B) by the influence of the covariates (Equation 6.5).

$$(\mathbf{A})\ \mathrm{rCL}_{1,i} = f \cdot \mathrm{CL}_{1,pop} \cdot \left(\frac{\mathrm{eGFR}_i}{100} \cdot \frac{\mathrm{BSA}_i}{1.73} \right) \cdot \beta$$

$$(\mathbf{B})\ \mathrm{rCL}_{1,i} = f \cdot \mathrm{CL}_{1,pop} \cdot \left(\frac{\mathrm{eGFR}_i}{100} \cdot \frac{\mathrm{BSA}_i}{1.73} \right)^{\beta}$$

$$\mathrm{nrCL}_{1,i} = (1-f) \cdot \mathrm{CL}_{1,pop} \cdot \left(\frac{\mathrm{WT}}{75} \right)^{0.75}$$

$$\mathrm{CL}_{1,i} = \mathrm{rCL}_{1,i} + \mathrm{nrCL}_{1,i}$$

Equation 6.5: *Covariate model for BSA adjusted eGFR on the renal fraction of* CL_1 *using either a linear model* (**A**) *or a non-linear model* (**B**)*, and allometrically scaled total body weight* (*WT*) *on the non-renal fraction, with* $f = 0.81$.

Again, the LRT was used to assess the significance of the addition of the covariate effects described by parameter β for nested models, while the AIC was used to discriminate between non-nested models.

6.4.5. Refinement of the PK model for MTX in RBCs

Throughout the hypothesis testing process, the proposed population PK model for MTX in RBCs was updated based on the results. This means that if an alternative model tested during one of the hypothesis tests was found to be superior to the corresponding base model, the alternative model was then carried forward for subsequent testing (as an updated base model). Therefore, the resulting final model constitutes the best structural model and also includes all covariates that were found to be significant throughout this analysis.

Non-parametric bootstrap was used to evaluate the updated model. 400 bootstrap runs were conducted in NONMEM®, where resampling was stratified based on the number of patients that started, ceased or received continuous MTX treatment. The asymptotic empirical 95% confidence intervals of the parameter estimates were constructed based on the means and standard deviations for the runs that minimised successfully.

The updated model was furthermore assessed based on the individual fits to the data for its suitability as a basis for a full PKPD model, while prediction

and variability corrected visual predictive checks (pvcVPCs) [207] were used to assess the predictive performance of the model.

6.5. Results

6.5.1. Hypothesis test 1: Cleaving mechanism of γ-glutamyl hydrolase

The OFV and AIC values for the four structural models in Figure 6.1 assessing the preferred cleaving mechanism γGH are given in Table 6.3

Table 6.3: *OFV and AIC for structural models assessing the preferred cleaving mechanism of γGH.*

Model	Cleaving mechanism of γGH	OFV	AIC
Model A	1 glutamate moiety (original)	15430.96	15478.96
Model B	2 glutamate moieties, without conversion of MTXGlu2 to MTXGlu1	15873.14	15921.14
Model C	2 glutamate moieties, with conversion of MTXGlu2 to MTXGlu1	15898.96	15944.96
Model D	1 and 2 glutamate moieties	15528.52	15588.52

It needs to be noted that NONMEM® had difficulties with the minimisation for Models C and D, as the algorithm was trapped in local minima for the corresponding runs presented in Table 6.3. Model D did not default to either Model A, B or C although these models are all nested within the structure of Model D, while Model C equally failed to default to the corresponding nested model B.

These runs were all started from the same initial values, with 0.1 hr^{-1} for $k_{\gamma GH31}$, $k_{\gamma GH42}$ and $k_{\gamma GH53}$ for Models B, C and D. For Model D, these initial values resulted in final estimates of 0.0608 hr^{-1}, 0.0484 hr^{-1}, and 0.0897 hr^{-1}, respectively, for these parameters. Yet, when the initial estimates for $k_{\gamma GH31}$, $k_{\gamma GH42}$ and $k_{\gamma GH53}$ were set to 1 x 10^{-4} hr^{-1}, Model D almost defaulted to Model A

with final estimates of 0.0042 hr^{-1}, 1 x 10^{-6} hr^{-1}, and 3 x 10^{-6} hr^{-1} for $k_{\gamma GH31}$, $k_{\gamma GH42}$ and $k_{\gamma GH53}$, respectively, and an OFV of 15430.83 (AIC 15490.83).

Based on these results, cleaving of only one glutamate moiety by γGH is preferred, as the original model (Model A) has the lowest AIC. This model structure was therefore retained for the following hypothesis test.

6.5.2. Hypothesis test 2: Loss of MTXPGs from RBCs

The results for the six hypotheses tested to assess a potential loss of MTXPGs from RBCs are shown in Table 6.4.

Table 6.4: *OFV and AIC for the six hypotheses (H) assessing a potential loss of MTXPGs from RBCs.*

H	Coding of loss of MTXGluX ($k_{out,GluX} = CL_{GluX}/V_{GluX}$) [a]	OFV	AIC
1a	$k_{out,Glu1} \neq k_{out,Glu2} \neq k_{out,Glu3} \neq k_{out,Glu4} \neq k_{out,Glu5}$	25627.48	25691.48
1b	$k_{out,Glu1} = k_{out,Glu2} = k_{out,Glu3} = k_{out,Glu4} = k_{out,Glu5}$	15413.88	15461.88
2a	$k_{out,Glu1} \neq k_{out,Glu3} \neq k_{out,Glu4} \neq k_{out,Glu5}$ & $k_{out,Glu2} = 0$	15422.88	15480.88
2b	$k_{out,Glu1} \neq k_{out,Glu3} = k_{out,Glu4} = k_{out,Glu5}$ & $k_{out,Glu2} = 0$	15422.03	15474.03
3a	$k_{out,Glu1} \neq k_{out,Glu4} \neq k_{out,Glu5}$ & $k_{out,\ Glu2} = k_{out,Glu3} = 0$	15421.19	15475.19
3b	$k_{out,Glu1} \neq k_{out,Glu4} = k_{out,Glu5}$ & $k_{out,Glu2} = k_{out,Glu3} = 0$	15420.05	15472.05

[a] Note, that here ≠ means that the parameters were allowed to differ in principle, but could still be estimated as having the same value.

Again, identifiability issues became apparent during this analysis with all tested models other than the model for Hypothesis 1b. None of the additionally parameters $k_{out,GluX}$ describing loss of MTXPGs that were added in these models could be estimated as changes in these parameters did not result in changes in the OFV (observed as zero gradients throughout the minimisation procedure in NONMEM® for these parameters).

However, Hypothesis 1b, where the loss of all MTXGluX from the cells is assumed to be equal to the rate constant of the parent drug ($k_{out,Glu1}$), was stable and provided a better fit to the data than the original model as the AIC value

for this hypothesis is lower than the AIC value of the non-nested original model (15461.88 *versus* 15478.96). Therefore, the population PK model for MTX in RBCs was updated according to Hypothesis 1b and this updated model was used for the following hypothesis tests.

6.5.3. Hypothesis test 3: Significance of genotypic covariates on MTXGluX PK in RBCs

The eight PK relevant polymorphisms available in the data set were individually included in the updated population model as covariate effect on the corresponding parameters. However, none of these covariates showed a significant effect. Based on the LRT a drop in the OFV of at least -5.99 is required to render the inclusion of two additional parameters significant, yet inclusion of the two additional covariate coefficients β_1 and β_2 did not result in a drop of more than -5.99 for any of the tested covariates (Table 6.5).

The biggest, yet still non-significant reduction in OFV (-5.07) is seen for the inclusion of the CNV of rs35592 in ABCC1, followed by rs4148396 in ABCC2 (-3.93). For these models, the only notable change in the population parameter estimates compared to the base model was seen for CL_{Glu1}, the parameter on which the covariate effect was included: 2.93×10^{-4} hr^{-1} in the base model *versus* 2.71×10^{-4} hr^{-1} for the model including rs35592 and 2.53×10^{-4} hr^{-1} for the model including rs4148396.

Table 6.5: *Comparison of hypothesised models without and with inclusion of genotypic covariates relevant to the intracellular PK of MTX.*

Gene	SNP	Variant	Parameter with covariate effect	Population mean estimate	Estimated coefficients β_1	Estimated coefficients β_2	OFV
Base model without genotypic covariates			-	-	-	-	15413.88
Genotypic covariate that affects γ-glutamyl hydrolase							
γGH	rs11545078	CNV	$k_{\gamma GH2}$	0.1678 hr^{-1}	1.0711	-	15413.61
			$k_{\gamma GH3}$	0.1879 hr^{-1}	same		
			$k_{\gamma GH4}$	0.2397 hr^{-1}	same		
			$k_{\gamma GH5}$	0.2966 hr^{-1}	same		
Genotypic covariates that affect the multi drug resistance transporters							
ABCC1	rs35592	CNV	CL_{Glu1}	2.71×10^{-4} hr^{-1}	2.2702	1.3986	15408.81
	rs3784862	CNV	CL_{Glu1}	2.93×10^{-4} hr^{-1}	1.0311	2.93×10^{-4}	15413.67
ABCG2	rs17731538	CNV	CL_{Glu1}	2.91×10^{-4} hr^{-1}	0.9393	1.0347	15413.84
ABCC2	rs4148396	CNV	CL_{Glu1}	2.53×10^{-4} hr^{-1}	0.9420	1.3862	15409.95
	rs2273697	1249G>A	CL_{Glu1}	2.93×10^{-4} hr^{-1}	1.00×10^{-6} [a]	1.00×10^{-6} [a]	15413.93
ABCB1	rs1045642	3435C>T	CL_{Glu1}	2.81×10^{-4} hr^{-1}	2.67×10^{-5}	1.07×10^{-5}	15413.71
Genotypic covariate that affects the reduced folate carrier							
SLC19A1	rs1051266	80G>A	k_{in}	2.00×10^{-4} hr^{-1}	2.32×10^{-5}	3.25×10^{-5}	15412.88

[a] Minimisation terminated, estimates of β_1 and β_2 near boundary

6.5.4. Hypothesis test 4: Significance of phenotypic covariates to predict between subject variability in the plasma PK of MTX

Table 6.6 gives an overview of the results obtained with the three different covariate models using eGFR as a biomarker for predictable BSV on plasma clearance of MTX (CL_1) in comparison to the base model without covariates on CL_1. Note that here the base model is the updated structural model where the loss of all MTXGluX is described by the rate constant $k_{out,Glu1}$ (Hypothesis 1b in Section 6.5.2).

Table 6.6: *Comparison of the hypothesised covariate models testing the significance of phenotypic covariates on total versus renal plasma clearance of MTX.*

Model	Covariate model	β	OFV	AIC
Base	none	-	15413.88	15461.88
Hypothesis 1	BSA adjusted eGFR on total CL_1	0.284	15407.51	15457.51
Hypothesis 2A	linear model for BSA adjusted eGFR on rCL_1, allometric WT on $nrCL_1$	0.106	15402.96	15452.96
Hypothesis 2B	non-linear model for BSA adjusted eGFR on rCL_1, allometric WT on $nrCL_1$	0.358	15406.85	15456.85

As can be seen from these results, the inclusion of the covariates on CL_1 is significant based on the LRT for all hypothesised covariate models as all three models result in a decrease in the OFV of more than -3.84 in comparison to the base model that does not include the additional parameter β.

Out of the three covariate models, Hypothesis 2A performs best as it has the lowest AIC value. Therefore, the effect of BSA adjusted eGFR on MTX plasma clearance is best described by a linear covariate model on the renal fraction (rCL_1) only, while allometrically scaled total body weight (WT) is used to predict BSV in the non-renal fraction ($nrCL_1$).

6.5.5. Refinement of the PK model for MTX and MTXPGs in RBCs

Based on the results of the four different hypothesis tests, the original population PK model for MTX in RBCs developed in Chapter 5 was refined here with respect to its structure and by adding further covariates. The structural update in the final model was based on Hypothesis 1b in Section 6.5.2 which assumed that all MTXGluX are lost from RBCs described by the same rate constant $k_{out,GluX} = k_{out,Glu1} = CL_{Glu1}/V_{Glu1}$, while BSA adjusted eGFR was included as covariate on renal plasma clearance and allometrically scaled total body weight on non-renal plasma clearance according to Hypothesis 2A in Section 6.5.4. These covariates account for predictable BSV in the otherwise fixed plasma PK of MTX. The updated NONMEM® code is shown in Appendix A.6.1.

The parameter estimates for the final updated model are shown in Table 6.7, together with their means $\bar{\phi}_p$ and standard deviations *s* obtained from the non-parametric bootstrap analysis based on 101 successful runs out of 400 runs in total (success rate 25.3%). The corresponding asymptotic empirical 95% confidence interval (CI) was calculated according to:

$$95\%\,\mathrm{CI} = \bar{\phi}_p \pm 1.96 \cdot s$$

Equation 6.6: *Formula to calculate the upper and lower bound of the asymptotic empirical 95% confidence interval (CI) of the bootstrap parameter estimates, with an α-error level of 0.05.*

As for the previous parent-metabolite model in Chapter 5, the mean parameter estimates of the bootstrap results agree reasonably well with the estimates for original data set under the updated model structure. The main reason for unsuccessful minimisations was again rounding errors.

Table 6.7: *Parameter estimates for the final updated parent-metabolite model.*

Parameter	Estimate	Shrinkage [%] [c]	Mean (BS) [a]	Empirical standard deviation (BS) [a]	Asymptotic empirical 95% CI (BS) [a,b]
k_{in} [hr^{-1}]	5.91×10^{-5}	-	5.92×10^{-5}	1.56×10^{-5}	2.87×10^{-5} – 8.97×10^{-5}
CL_{Glu1} [L/hr]	2.94×10^{-4}	-	2.92×10^{-4}	2.29×10^{-5}	2.47×10^{-4} – 3.37×10^{-4}
V_{Glu1} [L]	0.3 fixed	-	-	-	-
V_{Glu2} [L]	0.3 fixed	-	-	-	-
V_{Glu3} [L]	0.3 fixed	-	-	-	-
V_{Glu4} [L]	0.3 fixed	-	-	-	-
V_{Glu5} [L]	0.3 fixed	-	-	-	-
k_{FPGS1} [hr^{-1}]	0.171 fixed	-	-	-	-
k_{FPGS2} [hr^{-1}]	0.344 fixed	-	-	-	-
k_{FPGS3} [hr^{-1}]	0.097 fixed	-	-	-	-
k_{FPGS4} [hr^{-1}]	0.141 fixed	-	-	-	-
$k_{\gamma GH2}$ [hr^{-1}]	0.169	-	0.171	0.0100	0.151 – 0.190
$k_{\gamma GH3}$ [hr^{-1}]	0.189	-	0.187	0.0153	0.157 – 0.217
$k_{\gamma GH4}$ [hr^{-1}]	0.241	-	0.239	0.0149	0.210 – 0.268
$k_{\gamma GH5}$ [hr^{-1}]	0.296	-	0.303	0.0217	0.260 – 0.345
Covariate effect β of eGFR and BSA on CL_1	0.106	-	0.105	0.0933	-0.078 – 0.288

[a] Non-parametric bootstrap statistics based on 101 runs with successful minimisation out of 400 runs.
[b] Asymptotic empirical 95% confidence interval (CI) constructed based on mean and empirical standard deviation of the 101 successful BS runs.

Table 6.7: *continued*

Parameter	Estimate	Shrinkage [%] [c]	Mean (BS) [a]	Empirical standard deviation (BS) [a]	Asymptotic empirical 95% CI (BS) [a,b]
BSV k_{in} [%]	56.5	9.0	58.8	9.0	43.1 – 78.5
BSV CL_{Glu1} [%]	41.1	25.5	44.3	10.1	27.0 – 68.8
BSV V_{Glu1} [%]	32.2	23.6	30.8	8.7	15.7 – 55.3
BSV V_{Glu2} [%]	34.5	21.5	31.5	13.2	10.9 – 74.0
BSV V_{Glu5} [%]	51.8	18.5	45.9	10.7	19.3 – 97.8
BSV $k_{\gamma GH3}$ [%]	54.3	8.5	53.4	5.6	43.2 – 65.3
BSV $k_{\gamma GH4}$ [%]	42.0	4.5	41.0	4.7	32.4 – 51.2
BSV $k_{\gamma GH5}$ [%]	15.4	75.7	19.0	17.9	0.03 – 526.4
CV_{prop} Glu1 [%]	17.9	3.8	17.9	2.2	13.7 – 22.2
σ_{add} Glu1[nmol/L]	6.06		6.13	0.67	4.82 – 7.44
CV_{prop} Glu2 [%]	21.5	3.9	21.0	2.5	16.1 – 25.9
σ_{add} Glu2[nmol/L]	2.74		2.80	0.44	1.93 – 3.67
CV_{prop} Glu3 [%]	11.9	3.9	12.1	1.3	9.59 – 14.6
σ_{add} Glu3[nmol/L]	6.23		6.21	0.71	4.82 – 7.59
CV_{prop} Glu4 [%]	25.4	3.6	25.4	3.3	18.8 – 31.9
σ_{add} Glu4[nmol/L]	1.93		1.99	0.28	1.44 – 2.54
CV_{prop} Glu5 [%]	28.1	3.9	27.0	3.4	20.3 – 33.7
σ_{add} Glu5[nmol/L]	1.47		1.37	0.58	0.23 – 2.51

[c] η-shrinkage given for BSV parameters, ε-shrinkage for each MTXGluX given together with the corresponding RUV parameters CV_{prop} and σ_{add} GluX.

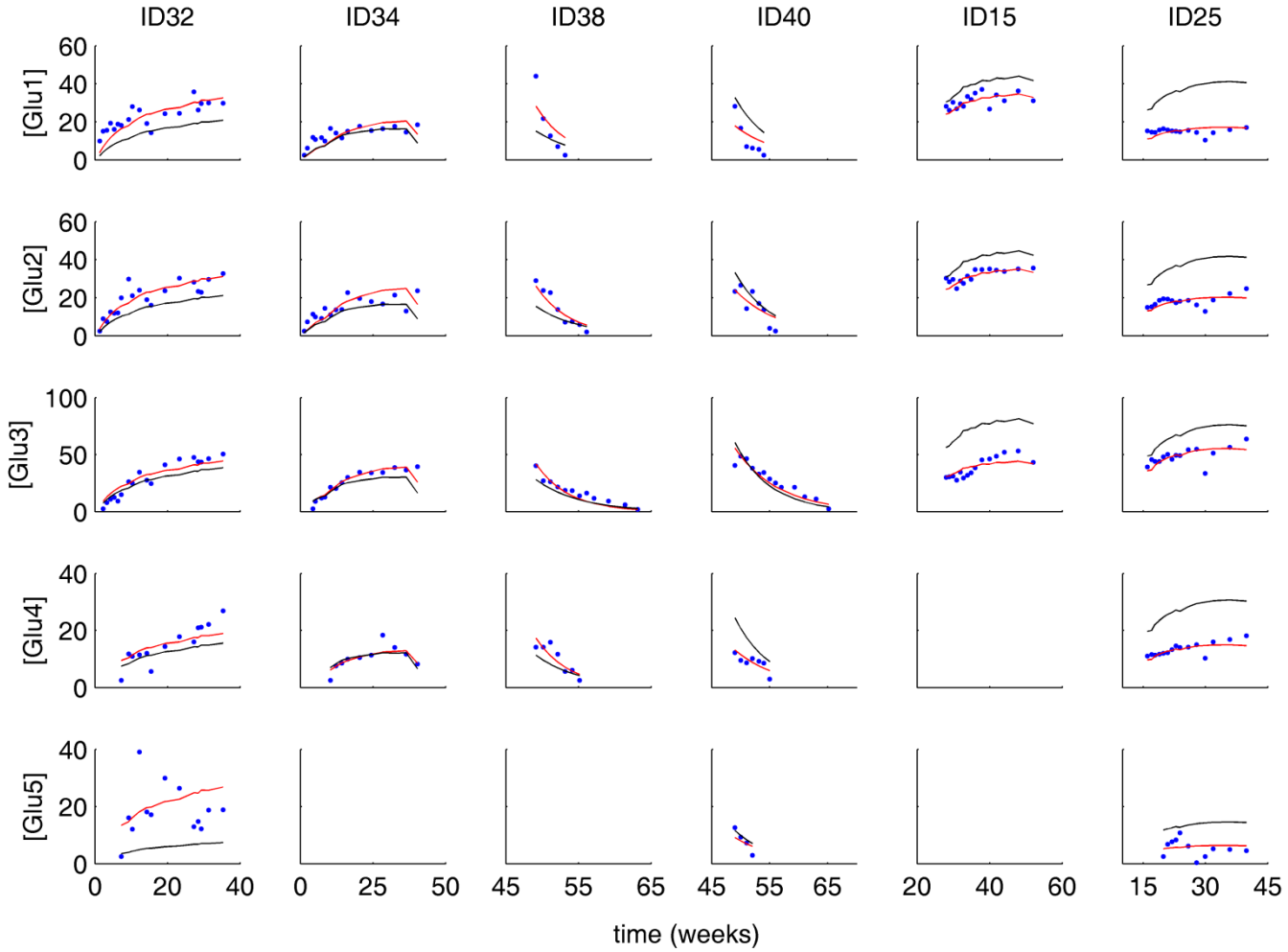
Figure 6.3 shows the individual fit of the updated model to the data of six representative individuals. These are the same individuals that were also plotted in Chapter 5. The individual fits of all 48 patients are provided in Appendix A.6.2.

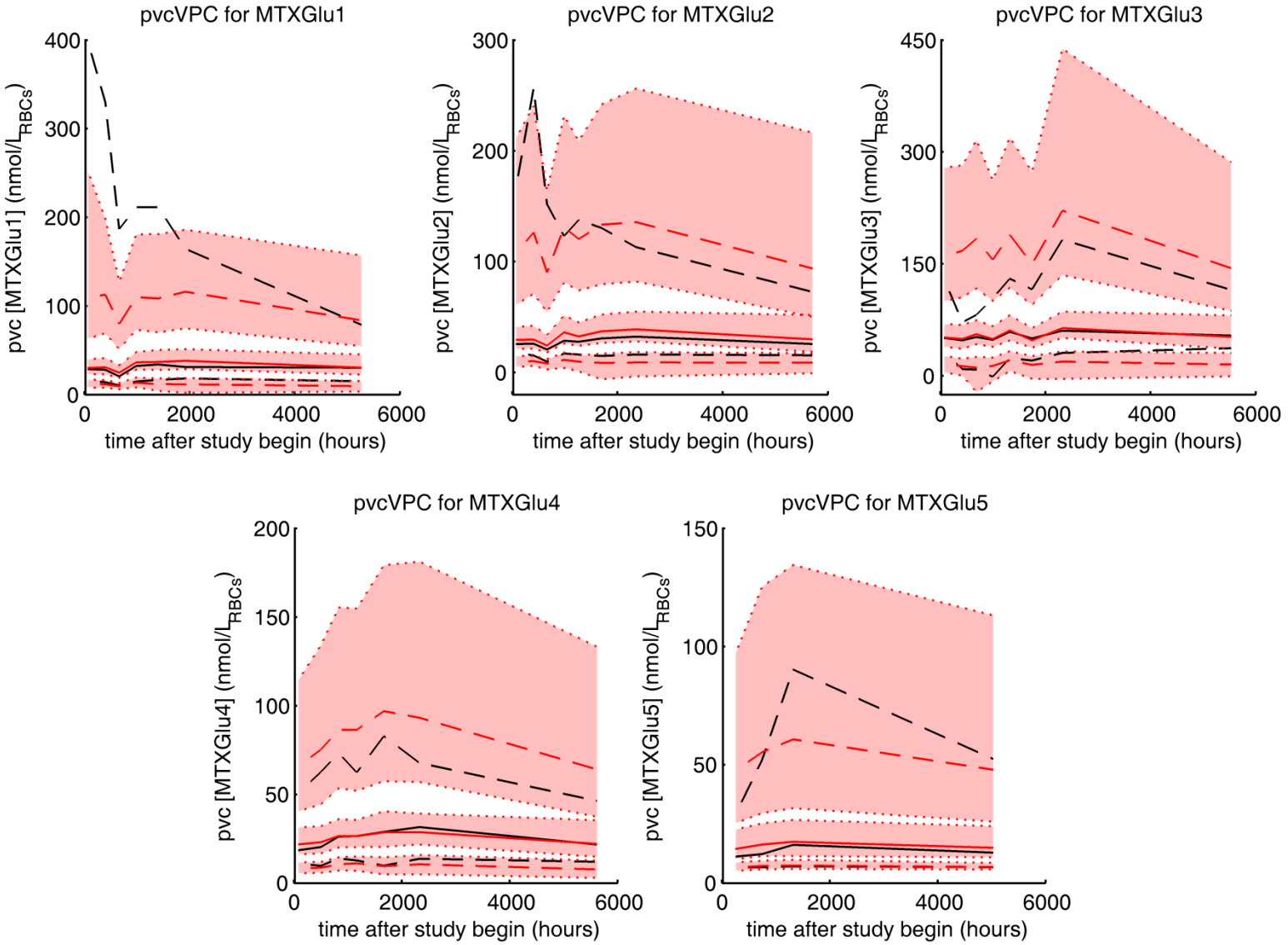
Comparing these plots with the corresponding figure in Chapter 5 (Figure 5.6) shows only subtle differences in the individual fits. However, the new model performs better on a global scale across all individuals and observations as can be seen from the AIC values: 15478.96 for the previous model developed in Chapter 5 *versus* 15452.96 for the final updated model.

The final updated model was furthermore assessed for its predictive performance based on pvcVPCs (Figure 6.4). The simulations underlying these plots were stratified based on the number of patients starting, stopping or receiving continuous therapy in the study, and were corrected for data below the limit of quantification (BLQ data), i.e. simulated MTXGluX concentrations below the limit of quantification (5 nmol/L_{RBCs}) were removed from the simulated data sets as was done in the original study. The 10^{th}, 50^{th} (median) and 90^{th} percentile of the prediction and variability corrected (pvc) observed and simulated MTXGluX concentrations were plotted against time after study begin as independent variable with an equal number of observations per bin. In addition, the 95% confidence intervals around the simulated percentiles were constructed.

following page:

Figure 6.3: *Individual fits for all five MTXGluX obtained with the updated parent-metabolite model for six representative patients: ID32 & ID34 starters, ID38 & ID40 stoppers, ID15 & ID25 patients on continuous therapy. Blue dots: observations, red line: individual prediction, grey line: population mean prediction.*


ID32
ID34
ID38
ID40
ID15
ID25
[Glu1]
[Glu2]
[Glu3]
[Glu4]
[Glu5]
time (weeks)


pvcVPC for MTXGlu1
pvc [MTXGlu1] (nmol/L_RBCs)
time after study begin (hours)
pvcVPC for MTXGlu2
pvc [MTXGlu2] (nmol/L_RBCs)
time after study begin (hours)
pvcVPC for MTXGlu3
pvc [MTXGlu3] (nmol/L_RBCs)
time after study begin (hours)
pvcVPC for MTXGlu4
pvc [MTXGlu4] (nmol/L_RBCs)
time after study begin (hours)
pvcVPC for MTXGlu5
pvc [MTXGlu5] (nmol/L_RBCs)
time after study begin (hours)

Prediction and variability correction was conducted to account for predictable variability arising from differences in other independent variables such as differences in dose and covariates between the patients [207]. When these corrections are applied the resulting pvcVPC plots provide a better reflection of the random variability in the predictions than classical uncorrected VPCs [207]. Note that this correction is done after the BLQ correction, and the prediction and variability corrected MTXGluX concentrations can be lower than the quantification limit.

Overall, the median model prediction show reasonably good agreement with the median of the data, with a slight tendency to over-predict the median MTXGlu2 concentrations. This means that the structural model is able to describe the MTXGluX concentrations in an average individual well. Notably, the best agreement is obtained for MTXGlu3, which is the most abundant MTXGluX. However, the model has difficulties in capturing the random BSV, as can be seen from the mismatch of the 10^{th} and 90^{th} percentiles and the large confidence interval especially around the upper percentile.

previous page:

Figure 6.4: *Prediction and variability corrected visual predictive checks (pvcVPCs) for all five MTXGluX based on the final updated model stratified for starters, stoppers and continuous therapy in the original data set. Median (solid lines), 10^{th} and 90^{th} percentiles (dashed lines) of the observed (black) and simulated (red) corrected MTXGluX concentrations with the 95% confidence interval around the simulated percentiles (shaded red areas).*

6.6. Discussion

6.6.1. Hypothesis testing

In this chapter the previously developed population PK model for MTX in RBCs was used to test four different hypotheses.

The first hypothesis test concerned the cleaving mechanism of γGH, i.e. whether the enzyme cleaves one, two, or one and two moieties of glutamate simultaneously. The previously developed model assumed cleaving of just one glutamate moiety and none of the alternative structural models tested here provided the apparent best fit to the data, although cleaving of two glutamate moieties was shown to occur *in vitro* by Rhee *et al.* [97]. However, it needs to be noted that the developed model is empirical in nature and the available data reflects a cross-section of MTXGluX concentrations in a random sample of RBCs rather than the kinetics in an individual RBC (see also Discussion in Chapter 5 Section 5.6.3). Therefore, it remains possible that the true cleaving mechanism of γGH cannot be observed in the available data based on this empirical model. It is also noted that the more complex models did not naturally collapse to the simpler models as the estimation algorithm was trapped in local minima. Nevertheless, using a variety of sets of initial estimates for the more complex models did not provide a fit that was superior to the previous model. In this light and based on the available data, the simple model, which assumes the cleaving of just one glutamate moiety, seems to describe the kinetics of intracellular MTX metabolism best. Yet, a formal identifiability analysis is required to assess whether the observed stability issues of the alternative models are due to a lack of structural (*a priori*) identifiability that is inherent in the models themselves, or result from a lack of information in the available data (deterministic non-identifiability). In the later case, changes in the design of future studies would be required to overcome the observed stability issues.

The second hypothesis test also considered structural aspects of the proposed model. Here, it was tested whether an additional loss of MTXPGs is

observable in the available data and supported by the model. Loss of MTXGluX could be due to two different reasons: 1) efflux transport via MDRT and 2) removal of RBCs from the circulation. MTXGluX is known to be a substrate for MDRT efflux pumps, while MTXPGs have a much lower affinity to these transporters. Nevertheless, Morrison and Allegra [8] found a model that included loss of long chain MTXGluX from human breast cancer cells to be superior to a model that only accounts for loss of MTXGlu1. If such a loss is due to active efflux transport, the rate constants of the loss of MTXPGs should differ from the rate constant of MTXGlu1 and also differ between the individual MTXGluX due to the differences in the affinities to the transporters. The second possible mechanism of loss is due to removal of RBCs from the circulation, i.e. death of RBCs. If this is the only loss mechanism affecting MTXGluX, all MTXGluX should be lost with the same rate constant when for simplification purposes RBC destruction is assumed to follow a first-order process rather than a finite lifespan. However, it was not possible to discriminate between these two possible mechanisms of MTXGluX loss based on the proposed model as the addition of further estimated loss parameters for MTXPGs resulted in additional identifiability issues in the model. Yet, the hypothesis that the loss of all MTXGluX is described by the same rate constant $k_{out,GluX} = k_{out,Glu1} = CL_{Glu1}/V_{Glu1}$ provided a better fit to the data than the original model which only allowed for the loss of MTXGlu1. Therefore, a general loss of all MTXGluX was supported in the model based on the available data and this updated model was used for the subsequent hypothesis tests on the significance of covariates, although the underlying mechanism(s) for this loss remain elusive.

The third hypothesis test addressed the question of whether genotypic covariates showed a significant effect on the intracellular PK of MTX in RBCs. Although a large number of polymorphisms affecting genes which are associated with enzymes and transporters relevant to the intracellular MTX PK have been identified and investigated for association with outcomes of MTX treatment, none of these polymorphisms have been tested for their significance

as covariates in a population PK model for MTX. In the data set used for analysis data on several polymorphisms relevant to the intracellular PK of MTX in RBCs were available, including polymorphisms affecting the influx and efflux transporters as well as the deglutamation enzyme γGH. These polymorphisms were tested independently as covariates on the corresponding parameter values in the updated population model. Although two of the tested polymorphisms (ABCC1 rs35592 and ABCC2 rs41488396) showed a notable reduction in the OFV (Table 6.5), none were found to have a statistically significant effect. Yet, it needs to be noted that this analysis only included genotypic information on 46 individuals (out of 48 patients in the whole data set) and some of the tested polymorphisms have a low frequency of carriers homozygous for the minor allele or multiple allele copies. Therefore, the lack of significance could be due to a lack of statistical power and a bigger study cohort would be required for a better assessment of the significance of these genotypic covariates. In addition, future studies should focus on the most promising covariate candidates seen in this analysis (ABCC1 rs35592 and ABCC2 rs41488396), and should be designed in such a way that a maximum of information on the affected parameter CL_{Glu1} is obtained, as this was the only parameter estimated that showed a notable change between the base model and the model with the genotypic covariates.

Nevertheless, the large number of known polymorphisms with a possible effect on the intracellular PK of MTX makes it difficult to assess the individual effect of a single polymorphism. Some polymorphisms potentially lead to an increase in the intracellular MTX concentrations, while others are believed to decrease these concentrations. If polymorphisms with opposing effects are present in the same individual only the net effect is observed, which can lead to insignificant results when testing these effects independently. This becomes even more relevant if the analysis is extended to include the pharmacodynamic effect of MTX, as polymorphisms have also been described for many of the enzymes involved in the postulated pathway of action of MTX. Therefore, it might be beneficial in future studies to use a combined genotypic index that

incorporates several polymorphisms according to their net effect as covariate in the model rather than individual polymorphisms, as was also suggested by Dervieux *et al.* [114]. For such a combined covariate analysis approach it would be relevant to know whether certain polymorphisms always occur in combination, i.e. are in linkage disequilibrium, and whether their effects are additive, synergistic or even antagonistic. This knowledge would facilitate choosing the structure of the corresponding covariate model, e.g. whether it needs to allow for an interaction between polymorphisms, and the choice of an appropriate covariate model structure is paramount to identify a true effect during the modelling analysis. Yet, despite the vast number of publications on polymorphisms relevant to MTX PK and its postulated pathway of action, such detailed information is still lacking to a large extent, which hampers the potential of identifying relevant genotypic covariates through a modelling based approach.

Finally, it was tested whether phenotypic covariates such as eGFR provided information relating to the predictable BSV component in an otherwise fixed model of the plasma PK of MTX. As the available data set did not contain MTX plasma concentrations, no BSV in the parameters describing the plasma PK of MTX could be estimated during model development in Chapter 5 and all individuals were assumed to exhibit the same plasma PK. However, this assumption risks bias in the estimates of the parameters describing the RBC PK. Therefore, the last hypothesis test assessed whether the addition of predictable BSV, described by a measure of eGFR as covariate on MTX plasma clearance (CL_1), results in a significant improvement of the model. As the eGFR values available in the data set were calculated according to the MDRD formula, BSA adjusted eGFR was included as covariate on CL_1 to account for the body size normalisation inherent in the MDRD formula. Including BSA adjusted eGFR on total plasma clearance already resulted in a significant improvement in the goodness of fit. The fit was improved even further if BSA adjusted eGFR was included as covariate on the renally cleared fraction of MTX only while allometrically scaled total body weight was used to

describe predictable BSV in the non-renally cleared fraction. These results are expected as differences in GFR only affect renal clearance of a drug, while differences in the hepatic capacity between individuals are related to differences in body size. Furthermore, it was found that a linear covariate model for BSA adjusted eGFR on renal MTX plasma clearance is superior to a non-linear model. Although this linear relationship between eGFR and renal MTX clearance is an indicator for a predominant first-order elimination process, this was not unexpected. It is unlikely that the secretion component of renal MTX clearance is saturated at the low doses of MTX used for the treatment of RA in the patients studied here, and a deviation from an apparent first-order elimination process would only be expected at higher MTX plasma concentrations than those present in these individuals. It needs to be noted, however, that the estimated covariate effect β of BSA adjusted eGFR on the renal component of MTX plasma clearance was 0.106. This estimate results in a predicted renal clearance of 0.72 L/hr in a typical individual with a BSA of 1.73 m^2 and an eGFR of 100 ml/min, which is only 10.6% of the population mean value if the covariate effect was not included (6.8 L/hr). It is likely that the estimate of β is biased as the RBC MTXGluX data might not be very informative with respect to the plasma kinetics of MTXGlu1.

Based on these results, the developed model was updated to include predictable BSV in the plasma PK of MTX described by BSA adjusted eGFR and allometrically scaled total body weight as covariates on renal and non-renal plasma clearance of MTX. Nevertheless, in future studies plasma samples should also be obtained so that the true BSV in the plasma PK (predictable based on covariates as well as unpredictable) can be estimated. This would then also allow for a truly unbiased estimation of the parameters that describe the intracellular PK of MTX.

6.6.2. Purpose of the model and application

Population PK models are developed for two main purposes: 1) to be able to use the model for extrapolation into different circumstances through

simulations, and/or 2) to provide an accurate description of the concentration-time course that can be used as a driver to describe a concurrent pharmacodynamic (PD) process.

The main aim of this work was to develop a population PK model for MTX measured in RBCs that is able to fulfil the second purpose, as the developed PK model is intended to form the basis for a full population PKPD model which would then allow testing whether RBC MTXGluX concentrations are suitable biomarkers for monitoring of MTX treatment in RA.

Based on the results obtain here, the updated model should be able to fulfil its designated purpose in future work. It does provide an accurate description of the intracellular PK of MTX for all individuals in this study which can be seen from the good individual fits that were obtained (Figure 6.3). Yet, a future PKPD model should best be developed based on the so-called IPP approach [208]. Using this approach the individual PK parameter estimates (also referred to as empirical Bayes estimates) are used as fixed values for each individual, i.e. the individually predicted PK profiles are retained, while only the PD related parameters are estimated. Therefore, this approach treats PK and PD as being independent processes as it does not include the PK data in the PD analysis. The only requirement with respect to the PK model is a good individual fit, which is obtained with the developed model for MTX measured in RBCs. Furthermore, no causal link between PD effects of MTX and MTX RBC concentrations is expected as RBCs are not located on the postulated pathway of action of MTX. This means that the underlying assumption of the IPP approach, independence between PK in RBCs and PD, is valid. In fact, the use of other PKPD modelling approaches which include PD information to estimate PK parameters, e.g. the simultaneous approach, could be problematic as under these circumstances the PD data will provide wrong information on the PK which can result in biased estimates for the PK parameters.

Nevertheless, the developed model is unable to fulfil the first potential purpose of a population PK model as evident from the misfit with respect to the predicted variability compared to the variability in the observations seen in

the pvcVPC plots in Figure 6.4. Although a good agreement between the median model predictions and the median observations was obtained, the variability in a population predicted by the model exceeds the variability observed in the data for all MTXPGs while the variability in MTXGlu1 is under-predicted as can be seen in the pvcVPC plots. This suggests that the BSV parameters in the model are not accurately estimated and do not reflect the true BSV in the corresponding structural parameters seen in the population.

One of the goals for including random effects such as BSV parameters in a non-linear mixed effect modelling analysis is to obtain an accurate estimate of the true variability in a population which then allows using the model for extrapolation purposes, i.e. simulations. This goal could not be achieved for this model and simulation based diagnostics such as the pvcVPCs in Figure 6.4 should be interpreted cautiously. The second purpose of including BSV terms in a population model is to eliminate the bias in the structural parameter estimates (fixed effects) that would occur if the repetitive measurements in an individual were treated as being independent of each other, i.e. not being correlated through the BSV terms.

Here, it was tested whether the structural model is stable with respect to changes in BSV on CL_{Glu1} as this BSV parameter affects all five MTXGluX via the loss parameter $k_{out,GluX} = k_{out,Glu1} = CL_{Glu1}/V_{Glu1}$. In two additional analyses BSV on CL_{Glu1} was fixed to ±25% of its estimated value and the remaining parameters were re-estimated. Neither of these runs resulted in differences in the estimates of the structural parameters (results not shown). Hence, including BSV in the model provides unbiased estimates of the fixed effects and is also required to obtain the good fit to the individual data. This means that the BSV terms in the model are non-ignorable, although the estimated values lack predictive potential. Future studies preferably on larger study cohorts are required to obtain a better estimation of the true BSV in the population that then also fulfils the goal of a good predictive performance.

6.7. Conclusion

Based on the available data the model supports a loss of all individual MTXGluX from RBCs when the loss is described by the same rate constant $k_{out,GluX} = k_{out,Glu1} = CL_{Glu1}/V_{Glu1}$, while simultaneous cleaving of more than one glutamate moiety by γGH is not supported in the model. However, future work is necessary to establish whether the observed stability issues are due to structural or deterministic identifiability issues in the proposed models.

In this small data set polymorphisms relevant to the intracellular MTX PK do not show a significant effect as genotypic covariates in the model when tested independently. Further studies on larger cohorts as well as additional information on interactions between the known polymorphisms are required for future work.

Although no plasma samples were available in the data set, it was possible to include predictable BSV in the plasma PK of MTX using phenotypic covariates such as eGFR, BSA and total body weight. BSA adjusted eGFR was found to have a significant linear covariate effect on the renal fraction of MTX plasma clearance, while allometrically scaled total body weight was able to describe variability in the non-renal fraction.

The final updated model is structurally stable and provides a good fit to the individual data. Therefore, it can be used in future work as the basis for a full PKPD model for MTX and MTXPGs measured in RBCs, and the IPP approach is recommended for the development of such a PKPD model. However, the model poorly estimates the variability seen in the population and therefore should not be used for extrapolation purposes such as simulation based predictions.

Future studies are required for a better quantification of BSV in the plasma PK of MTX as well as the true variability in MTXGluX concentrations.

PART IV

DISCUSSION & FUTURE PROSPECTS

Chapter 7: Discussion & Future Prospects

7.1. Synopsis of this thesis

In this thesis two different approaches were applied to model red blood cell (RBC) derived data. First, a bottom-up approach was applied to develop a model for the survival time of RBCs that was principally mechanism based. This model was subsequently applied to clinical data. The next part of this work focussed on describing the intracellular pharmacokinetics (PK) of methotrexate (MTX) in RBCs based on a mainly data driven, top-down modelling approach.

Both approaches are justified given the intended purposes of the developed models: quantification of the different physiological mechanisms involved in RBC destruction in the first case, and obtaining an accurate description of the concentration-time profile of MTX accumulation in RBCs in the second.

7.2. Discussion of the findings

7.2.1. A novel statistical model for RBC survival

The RBC survival model developed in Part II of this thesis has the intention of ultimately providing a better insight into RBC destruction mechanisms and how these are affected by pathological conditions such as chronic kidney disease (CKD). Thus, a mechanistic model was required, and a bottom-up approach was used to develop a novel statistical model for RBC survival. Here, "statistical" means that the model was developed in the frame work of survival data analysis, yet without being based on data. Instead, the underlying probability density function (pdf) was chosen based on prior knowledge of physiologically plausible mechanisms of RBC destruction.

To date, there is no reliable information available on the actual distribution of RBC lifespans in humans. However, it is known that physiological RBC destruction mechanisms include senescence, age-independent random destruction as well as death due to early or delayed failure. Therefore, a pdf that is able to account for these four mechanisms was

chosen as the underlying function in the model. In contrast, previously proposed RBC survival models that allow the RBC lifespan to vary within an individual mostly account only for senescence [75,76,78,79], and rarely for a combination of senescence and random destruction [80].

The developed model was shown to be able to replicate the shape of disappearance curves that are usually obtained with different RBC labelling methods. These include random labelling techniques where RBCs are labelled irrespectively of their age, e.g. using radioactive chromium (^{51}Cr), as well as cohort labelling techniques where the label is incorporated during the production of RBCs in the bone marrow over a period of time resulting in labelled RBCs of similar age, e.g. labelling with heavy nitrogen. Therefore, the model was thought to be in principle suitable for analysing RBC survival studies conducted in clinical practice.

The model was then extended to account for the flaws inherent in commonly used RBC labelling techniques to be able to obtain a more accurate description of the RBC lifespan. Notably, for the most commonly used RBC label, ^{51}Cr, the flaws included in the model did not only consist of the decay of the radioactive label and the random loss of label from the cells after dissociation of the chromium-haemoglobin complex, but also loss due to vesiculation of haemoglobin together with the bound label. Vesiculation of haemoglobin has only recently been reported in literature [68]. The developed model is one of the first where this mechanism of ^{51}Cr loss is accounted for when used to analyse the data of corresponding RBC survival studies.

It was tested whether full parameter estimation would be possible for this model under different labelling conditions based on theory of experimental design methodology, i.e. an information theoretical approach. Due to its mechanism based foundation, the developed model is of a complex nature and includes six fixed effect parameters. Although this is a rather large number of parameters to be estimated, it was nevertheless shown in Chapter 3 that full parameter estimation would be possible under ideal labelling conditions or under an intensive study design with the current labelling methods. However,

this analysis did not take random effects into account. If a random effect parameter would be included on each fixed effect parameter for a population analysis, then the total number of parameters to be estimated would increase to twelve. The deterministic identifiability of such a full population model requires further analysis, and the study design of future clinical studies should be optimised based on these results if the proposed model is to be used for data analysis in a full population setting in these studies.

The model can also be simplified by fixing certain components to render it applicable for data that does not provide sufficient information for full parameter estimation in a population setting. This approach was explored in Chapter 4 to estimate the survival of RBCs in healthy individuals as well as patients with CKD undergoing haemodialysis, as the available clinical data were not informative enough to support full parameter estimation in the model. Despite this simplification, the decrease in the apparent mean RBC lifespan in CKD patients compared to healthy controls was found to be similar to the decrease reported for the same data set based on an empirical analysis method [165]. This shows that applying the mechanism based model is not inferior to an empirical analysis. However, in contrast to the previous analysis [165], the mechanism based modelling analysis presented here allowed testing which underlying RBC destruction mechanism is more likely to result in the decreased RBC survival in CKD patients. The results imply that an increased random destruction is the more likely cause, rather than an accelerated senescence.

Nevertheless, although the developed model includes mechanistic principles of RBC destruction and was developed using a bottom-up approach, it is of semi-mechanistic nature as the underlying pdf was chosen arbitrarily. Further work is required to obtain the true RBC specific values for all parameters in the model, in healthy individuals as well as under different pathological conditions that affect RBC survival. Ideally, such future work should focus on recent RBC labelling methods such as biotin labelling [149,150]. This random labelling technique is believed to be less flawed than the ^{51}Cr

technique and would therefore be more informative for parameter estimation in the model in future studies.

7.2.2. A population PK model for MTX measured in RBCs

7.2.2.1. Model development

The main purpose of the RBC PK model for MTX developed in Part III was to provide an accurate description of the time course of RBC MTX concentrations. It is intended to form the basis of a full population pharmacokinetic-pharmacodynamic (PKPD) model that can be used to test the suitability of RBC MTX concentrations as a potential biomarker to monitor low-dose MTX treatment in rheumatoid arthritis (RA). Therefore, a compartmental PK model for MTX including its polyglutamated metabolites measured in RBCs was developed using a data driven top-down modelling approach. However, model development was not entirely based on the principle of parsimony, but also incorporated prior knowledge. Although the available data set did not contain plasma data, a model previously developed by Hoekstra *et al.* [91] was used to describe the plasma PK of MTX. Furthermore, the catenary structure of the parent-metabolite model was chosen based on prior knowledge of the mechanisms of intracellular MTX metabolism. Thus, the developed model is of semi-empirical as well is semi-mechanistic nature rather than being entirely empirical, i.e. data driven.

For simplification purposes and to reduce the computation time, all transport and enzymatic processes in the model were assumed to follow first-order kinetics. As this model is based on data of low-dose MTX treatment, this assumption is justifiable. The measured concentrations of all individual MTX species (MTXGluX) are much smaller than the reported K_m values of the corresponding enzymes or transporters. However, if the model were applied to data where MTX is administered in much higher doses, e.g. for cancer treatment, this assumption would not hold true. Therefore, this model cannot be used for extrapolation to high-dose MTX regimens as this would require the active processes to be described by Michaelis-Menten kinetics instead.

The catenary structure of the developed model resulted in stability issues during parameter estimation. To overcome this problem, the volume of distribution parameters of all five MTXGluX (V_{GluX}) as well as the rate constants describing the polyglutamation steps ($k_{FPGS1-4}$) were fixed. The corresponding values were obtained from the parent model for V_{GluX} and from the literature for $k_{FPGS1-4}$. Further work is warranted to formally assess whether the stability issues encountered during model development are due to a lack of structural identifiability (i.e. the model is *a priori* not globally identifiable due to its structure), or due to a lack of deterministic identifiability (i.e. the available data does not support full parameter estimation in the model). In the first case, it would be of interest to further investigate what structural changes are required to render the model structurally identifiable; whereas in the second case experimental design theory could be used to optimise the design of future clinical studies to provide data informative enough to support full parameter estimation.

Another interesting finding during model development related to the data below the limit of quantification (BLQ). When the parent drug MTXGlu1 was modelled by itself, a difference in the parameter estimates was obtained when using Stuart Beal's M3 method compared to the M6 method to handle BLQ data [189]. Thus, treating BLQ data as censored and computing the joint likelihood of true and censored data (M3 method) provided additional information that was valuable for parameter estimation in the parent model. However, this was not observed in the case of the parent-metabolite model, although the data set contained a considerable proportion of BLQ data, especially for the highest MTX polyglutamate, MTXGlu5 (45%). Using the M6 method, all BLQ observations other than the first or last in a sequence of decreasing or increasing concentrations are removed from the data set, while the remaining BLQ observations are replaced by a value equal to half the limit of quantification. Removing such a large number of observations, even if they are censored data, would be expected to decrease the information inherent in the data set and to lead to less accurate parameter estimates [209]. Apparently

this was not the case here, as the parameter estimates obtained using the M3 method did not differ significantly from those obtained using the M6 method for the parent-metabolite model. A possible explanation for this might lie within the catenary model structure. Within the chain of RBC compartments, an observation of a particular MTXGluX also contains information on the other MTXGluX species. Thus, for a particular observation it might be equally informative to have a true measurement available for a short chained MTXGluX than to compute the likelihood of a censored observation for a longer chained MTXGluX that was BLQ at this observation. A simulation-estimation study would be a straight forward approach in future work to assess this interesting finding and to test the hypothesised explanation.

7.2.2.2. Hypothesis testing

The developed RBC PK model for MTX was used to test different hypotheses related to mechanistic aspects of the intracellular metabolism of MTX as well as to assess the significance of genotypic and phenotypic covariates.

The underlying mechanism of MTX deglutamation via the enzyme γ-glutamyl hydrolase (γGH) was assessed. Although it was reported in the literature that γGH is able to cleave two glutamate moieties simultaneously [97], a model structure that only allowed for cleavage of a single glutamate was found to be superior in this work. It was also assessed whether loss of each MTXGluX is supported in the model. This hypothesis was of two fold interest. Firstly, Morrison and Allegra [8] found that longer chained MTXGluX are likely to leave human breast cancer cells *in vitro*, although the multi drug resistance transporters involved in the efflux of MTX have a much lower affinity to MTX polyglutamates compared to the parent drug. Secondly, as RBCs were used as the matrix for the MTX PK data analysed here, destruction of RBCs should also result in a loss of all MTXGluX from the system. In fact, based on the developed model, loss of all MTXGluX from the system was found to be superior to loss of only MTXGlu1. However, further stability issues

became apparent during the structural hypothesis tests and not all postulated models were identifiable. Therefore, it was not possible to differentiate between loss of MTXGluX due to transport or due to RBC destruction. This again highlights the need for a formal identifiability analysis of all proposed model structures.

Polymorphisms affecting genes coding for enzymes and transporters relevant to the PK or PD of MTX have gained an increasing interest as potential biomarkers to explain some of the variability seen between patients in the efficacy and toxicity of MTX treatment [106]. However, to date no single polymorphism has emerged that would be suitable as predictor for the variability seen between individuals. This is largely due to a lack of evidence from association studies, as the observed results can often not be confirmed in replicate studies. In this study, it was tested whether genotypes that were considered relevant to the intracellular PK of MTX are significant covariates in the developed population model. Genotypic information available in the data set included polymorphisms relevant to the transporters facilitating influx and efflux of MTX as well as γGH. However, none of the tested polymorphisms provided a significant improvement in the goodness of fit when tested as a covariate on the corresponding PK parameters in the model. This result was not surprising as the available data set only contained genotypic information for 46 individuals. Given the low frequency of some of the alleles in the general population such a small data set is likely to be underpowered to show a significant effect. In addition, if different polymorphisms with opposing effects on the intracellular MTX concentrations occur in the same individual, then the observed net effect may be too small to show a significant effect when only a single polymorphism is tested as covariate as it was done here. Therefore, a larger sample size is required to provide enough power to reassess the significance of genotypic covariates in future studies. Such an analysis should also take into account the opposing effects of different polymorphisms by using genotypic indices rather than single polymorphisms as covariates in the model.

No plasma data were available in this analysis and the parameters describing the plasma PK of MTX in the developed model were fixed to population values obtained from literature [91]. Nevertheless, it was tested whether predictable between subject variability (BSV) in the otherwise fixed plasma PK of MTX could be accounted for by including phenotypic covariates in the model. Estimated glomerular filtration rate (eGFR) adjusted for body surface area (BSA) was tested as a covariate on plasma clearance of MTX based on three different covariate model structures. Out of these, BSA adjusted eGFR as covariate on renal clearance together with allometrically scaled total body weight as a covariate on non-renal clearance was found to be the best covariate model and resulted in an overall improvement in the model based on change in objective function value. However, in future studies MTX plasma concentrations should also be obtained to allow for quantification of random as well as predictable BSV in the plasma PK of MTX.

7.3. Future Prospects

7.3.1. RBC survival

With the currently available labelling methods for RBCs it is only possible to obtain a relative measure of the mean RBC lifespan as all methods are inherently flawed. Such a relative lifespan value is sufficient when comparing study groups, e.g. healthy controls *versus* CKD patients, or to assess RBC survival before and during treatment, e.g. with recombinant human erythropoietin. Yet, knowledge of the underlying RBC lifespan distribution and quantification of the processes involved in RBC destruction would be desirable to obtain a better insight into how pathological conditions such as CKD, sickle cell anaemia or haemolytic diseases affect the survival of RBCs. In combination with more informative labelling techniques, a semi-mechanistic model for RBC survival as was developed in this thesis might prove useful to obtain a better understanding of these pathologies in future studies.

In addition, knowledge of RBC survival is crucial when dealing with RBC derived clinical biomarkers, such as glycated haemoglobin (HbA_{1c}). HbA_{1c} is

the most commonly used biomarker for glycaemic control in diabetic patients, and it has also been suggested that elevated HbA_{1c} concentrations are sufficient to diagnose diabetes [210]. However, the extent of glycation of haemoglobin does not only depend on blood glucose concentrations but also the lifespan of RBCs. A reduced RBC survival, for example in patients with CKD, results in lower HbA_{1c} concentrations [124], which can lead to a false assumption of adequate glycaemic control or an incorrect negative diagnosis for diabetes. The use of mathematical models to describe the relationship between HbA_{1c} and RBC survival would allow interpretation of corrected HbA_{1c} values in patients with impaired RBC survival. Several such attempts have been proposed in literature [79,211,212]; however all employ simplified descriptions of RBC survival based on top-down modelling approaches. These empirical models of RBC survival might not hold true in a different population, such as CKD patients, and the limitations associated with RBC labelling methods are usually not taken into account. By using a mechanism based modelling approach these limitations might be overcome. In future work, the RBC survival model proposed here could be extended by including the mechanism of intracellular glycation of haemoglobin similar to the approach presented by Alskär *et al.* [213] which was based on an empirical RBC survival model [80]. Such an extended model could then be evaluated based on RBC survival data and HbA_{1c} measurements obtained in additional clinical studies.

Similarly, the proposed RBC survival model could also form the basis of a more mechanism focussed model for the intracellular PK of MTX in RBCs in the future. The empirical RBC PK model developed here did not *a priori* take into account the destruction of RBCs and subsequent loss of MTX from the system. However, as mentioned above, it became apparent during the hypothesis testing analysis that a loss of MTXGluX from RBCs is supported even under this empirical model structure. Yet, it was not possible to determine whether this loss can be attributed to RBC destruction only or includes a transporter mediated efflux process, as the empirical model suffers from stability issues. Incorporating RBC survival in a mechanistic model for the

intracellular PK of MTX could facilitate the delineation of these processes in future work.

7.3.2. RBCs as matrix for pharmacokinetic data

When dealing with RBC derived PK data, such as the MTX data in this thesis, it is important to bear in mind the specific physiological characteristics of RBCs. As mature RBCs are anucleate cells, they lack the ability to perform cell division and self renewal, and only have a limited enzyme and transporter capacity. This means that intracellular metabolism as well as influx and efflux processes in RBCs differ considerably from other cell lines, as was seen for the RBC MTX PK when compared to human breast cancer cells as analysed by Morrison and Allegra [8] or lymphocytes as studied by Panetta *et al.* [90]. Furthermore, RBCs normally survive for several months in the human circulation. This survival time should be accounted for when RBC derived PK data is analysed.

It is also important to realise that RBC derived PK data also contains information on RBC survival. Drugs that accumulate in RBCs could be used to determine RBC survival by analysing their RBC PK profile. Such an analysis should ideally be undertaken using drugs that provide cleaner data than MTX does. Thus, the ideal drug should accumulate intracellular preferably without the ability to leave the cells through either active transport or passive diffusion. Furthermore, it should not show an extensive intracellular metabolism to avoid stability issues as were noticed for the MTX RBC PK model, and should not be toxic to the cell. If the observed rate constant of loss from the system is the same for all tested drugs, then it is likely that this is a system parameter related to RBC destruction rather than a drug specific parameter.

In fact, a strong indicator for such a system related RBC loss parameter could be obtained if future studies on the MTX RBC PK verified that all MTXGluX are lost from the system at the same rate. Another candidate for such an analysis is the anti-malaria drug mefloquine, which accumulates in RBCs due to binding to haemoglobin and the cell membrane [214]. The slow loss of

mefloquine from the RBCs observed by San George *et al.* [214] could be quantified in further *in vitro* experiments and thus accounted for in an RBC PK model for mefloquine as a fixed parameter. Clinical data obtained in *in vivo* studies would then allow estimating a loss of mefloquine from the system that is potentially specific for RBC destruction.

7.3.3. RBC MTX concentrations as biomarker for low-dose MTX treatment

The suitability of RBC MTX concentrations as a biomarker to monitor low-dose MTX treatment has not only been discussed in literature for the treatment of RA [114,117], but also for the treatment of other inflammatory autoimmune diseases. These include psoriasis [215] and inflammatory bowel disease [216], where MTX is used as second line immunosuppressant. As for RA, the doses of MTX required to reach therapeutic effect in these conditions incur significant variability between patients. In addition, MTX plasma concentrations lack a clear correlation with clinical outcomes of MTX treatment and are considered unsuitable as a biomarker, not only in RA but also in psoriasis as well as inflammatory bowel disease. As an alternative, RBC MTX concentrations could be useful predictors for MTX efficacy and/or toxicity as they are a measure of cumulative exposure to the drug due to the intracellular accumulation of MTX, similar to the area under the curve of a plasma concentration time curve or measuring HbA_{1c} as marker of cumulative blood glucose exposure. Plasma concentrations of a drug, or blood glucose concentrations for that matter, only give an indication of the momentary situation, i.e. whether the drug concentrations are inside or outside of the therapeutic window or whether the patient is hyper- or hypoglycaemic. Measures of cumulative exposure, on the other hand, often show a better correlation with long-term effects, e.g. treatment success or diabetic complications.

Although the population RBC PK model for MTX developed in the third part of this thesis was built based on clinical data obtained in patients with RA, it can also be applied to other clinical data of low-dose MTX treatment. Thus, in

future work this model could not only facilitate the quest for a biomarker to monitor MTX treatment in RA, but also in other inflammatory autoimmune diseases. For this purpose, the developed population PK model should be extended towards a full population PKPD model, where the MTX RBC concentrations are linked with PD outcomes such as decrease in disease activity or occurrence of adverse effects. At the moment, it is unclear whether the PK of the parent drug MTXGlu1, of any of the individual polyglutamated MTX metabolites (MTXGlu2 – MTXGlu5), or of a particular combination of these MTX species is the main driving force of the PD effects of MTX. Using a population PKPD modelling approach, the PK profile of each MTXGluX and also their combinations can be tested to explore the driving force for the PD effect. This would then allow determination of which, if any, of these is the most suitable biomarker for a particular clinical outcome of MTX and could therefore be used to monitor MTX treatment.

Nevertheless, it needs to be considered that RBCs are not located on the postulated causal pathway of MTX action. In future work RBC MTX concentrations might be shown not to have a clinically relevant correlation with clinical outcomes of MTX treatment, and therefore not to be suitable biomarkers for low-dose MTX treatment in RA or other inflammatory autoimmune diseases. In this case, focus should be given to alternative biomarkers that could be purported to have a causal relationship with clinical outcomes of MTX treatment. Recently, van Haandel *et al.* [217] suggested MTX concentration measured in peripheral blood mononuclear cells (PBMC) as an alternative biomarker for MTX treatment in juvenile arthritis. PBMCs are white blood cells with a single nucleus, i.e. macrophages, monocytes and lymphocytes. These cells play an important role in the immune response and also autoimmune diseases. In contrast to RBCs, PBMC are directly involved in the anti-inflammatory effect of MTX, and intracellular MTX concentrations in PBMCs are thus more likely to have a causal relationship with clinical outcomes of MTX treatment. MTX has been shown to inhibit the proliferation of PBMCs [218] as well as to decrease the production of pro-inflammatory

cytokines and increase the production of anti-inflammatory cytokines by PBMCs [219]. Thus, under the assumption that these effects are directly related to the intracellular accumulation of MTX, PBMC MTX concentrations are promising candidates as biomarkers to monitor MTX treatment in inflammatory autoimmune diseases.

In the case of RA, it would also be possible to obtain mononuclear cells directly from the synovial fluid instead of extracting them from peripheral blood samples as it was done in the study by van Haandel *et al.* [217]. Assuming that MTX acts locally in the joints, i.e. at the primary site of pathology, this would allow measuring intracellular MTX concentrations directly at the site of action and in the postulated effect compartment, the mononuclear cells. For a PKPD analysis, this is a highly unusual, but probably advantageous situation that warrants further exploration in future studies.

7.4. Conclusion

Two different approaches to model RBC derived clinical data were successfully applied in this thesis.

The bottom-up approach used in Part II resulted in the development of a novel statistical model for the survival time of RBCs that is able to account for plausible physiological mechanisms of RBC destruction, and the model was successfully applied to clinical data.

The population PK model for RBC MTX concentrations developed using a top-down approach in the third part of this thesis is capable of describing the time course of accumulation of MTX and its polyglutamated metabolites inside RBCs, and can be used as the basis for a full PKPD model for low-dose MTX treatment in RA in future work.

Ultimately, both models will require updating and refinement as more data and mechanistic knowledge becomes available in the future. Model development and modelling of clinical data should thus be seen as dynamic and iterative processes of learning and confirming [220].

PART V

APPENDICES

These appendices contain additional material related to the individual chapters in this thesis.

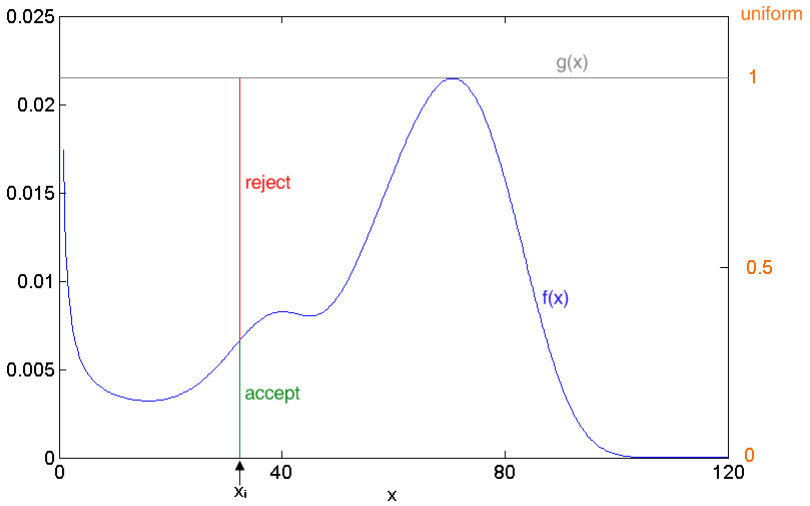
Some of these materials have been published in peer-reviewed publications or were included in my Postgraduate Certificate thesis as indicated accordingly.

Appendix 1: Appendix to Chapter 1

A.1.1. *Rejection Sampling Method*

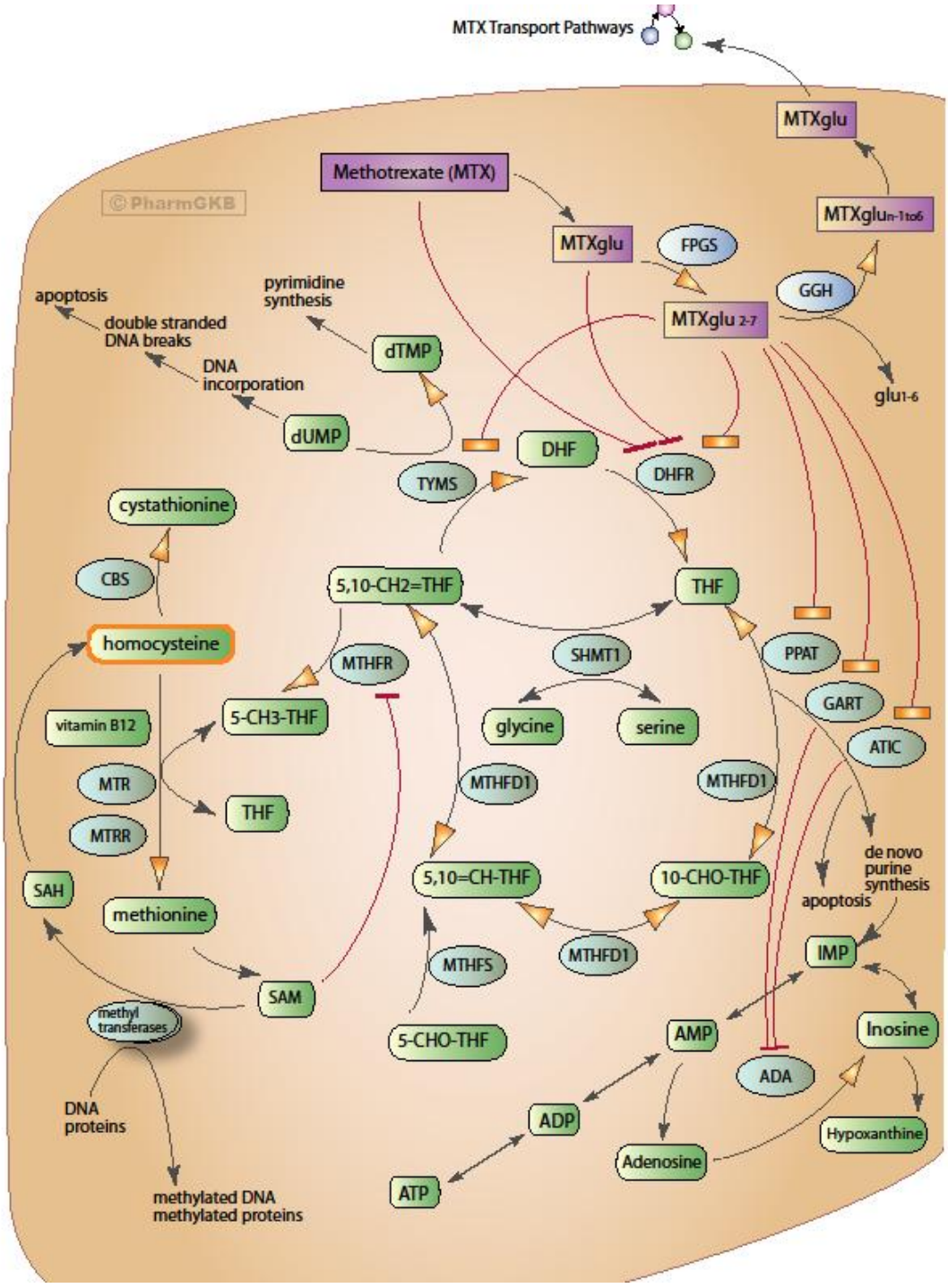
Given a known probability density function (pdf) $f(x)$ with a corresponding cumulative distribution function $F(x)$, which is not easily invertible, a sample of x-values that yields a similar pdf can be derived by the following steps [81,82,84,85]:

1) Find another function $g(x)$ that lies above $f(x)$ everywhere, i.e. $f(x) \leq g(x)$. The simplest case is to draw a straight line through the maximum of $f(x)$, if $f(x)$ has an absolute maximum $f_{max}(x)$ in the desired interval of x.
2) Generate a random value x_i.
3) Generate a random uniform number u_i distributed between 0 and 1.
4) Evaluate $f(x)$ for the value x_i.
5) Calculate the ratio r_i of $f(x_i)$ and $g(x_i)$: $r_i = f(x_i)/g(x_i)$
6) Accept x_i if $u_i \leq r_i$, otherwise reject it.
7) Repeat steps 2 – 6 until the desired number of x_i has been accepted or the maximum number of repetitions has been reached.



App Fig. 1.1.1: *Illustration of the rejection sampling method for RBC lifespans [81].*

A.1.2. *Postulated pathway of action of MTX and MTXPGs*

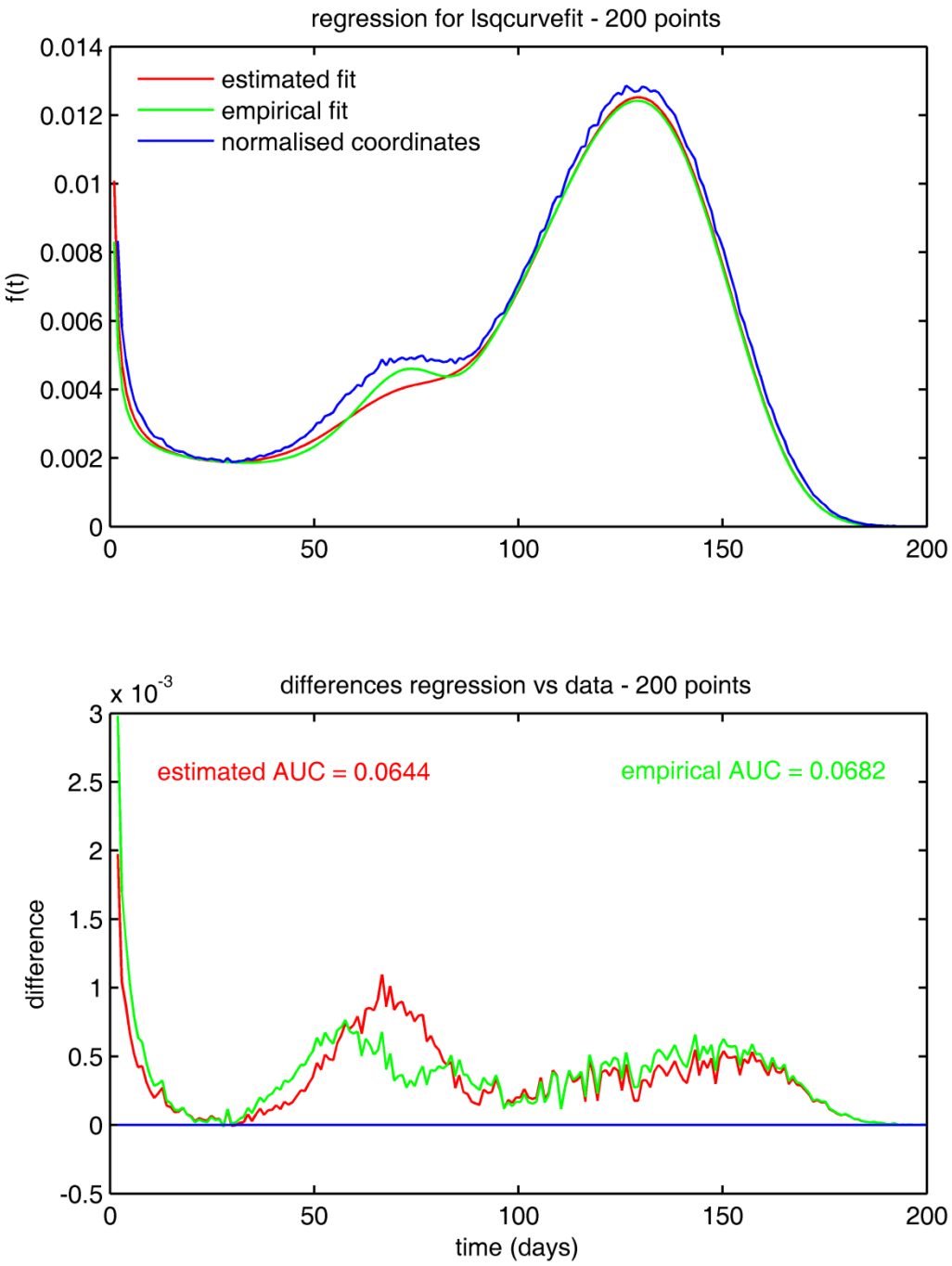


App Fig. 1.2.1: *Illustration of the pathway of action of MTX and MTXPGs [221].*

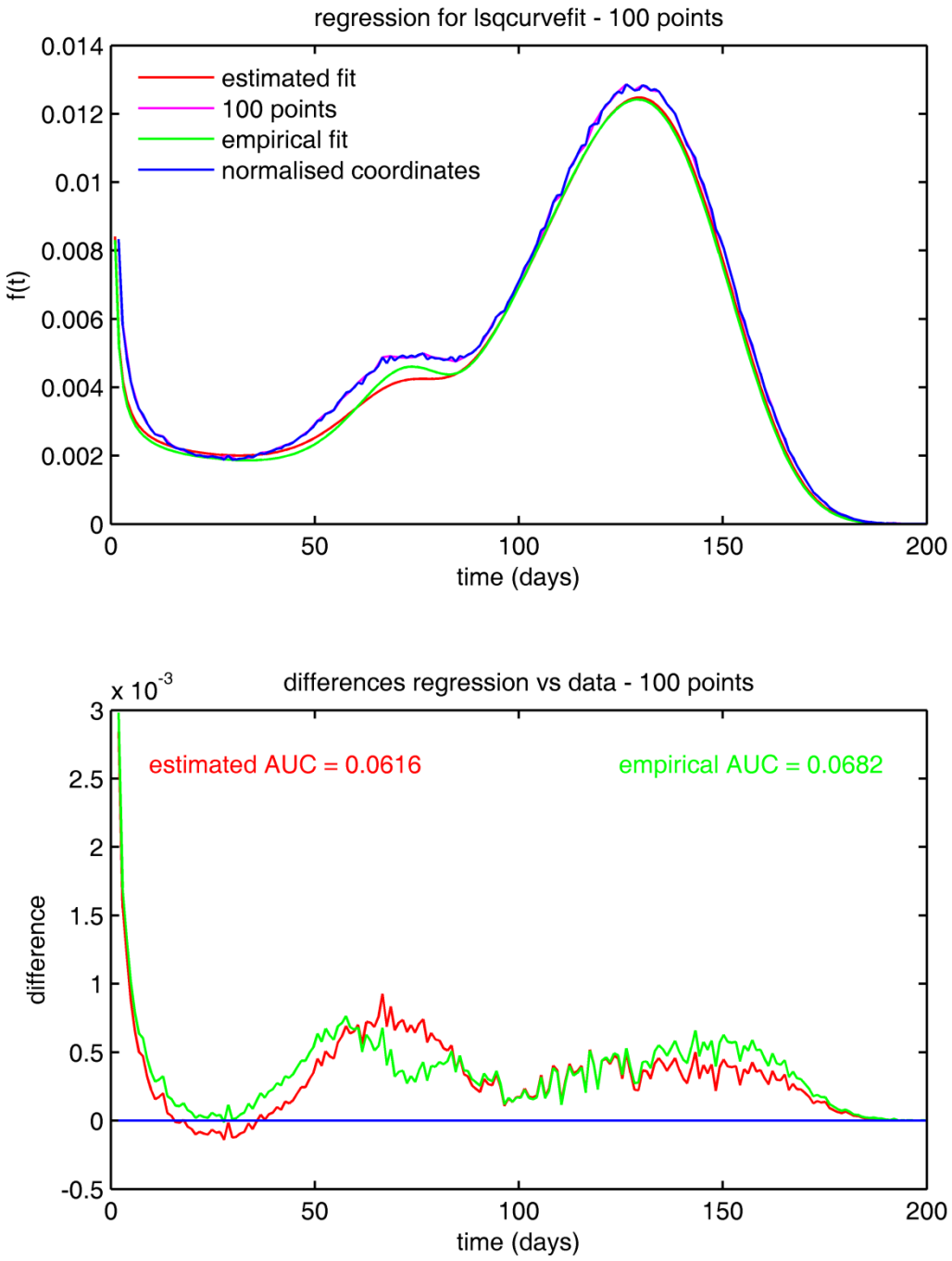
App Fig. 1.2.1 was obtained with permission of the copyright owner, the PharmGKB online database (http://www.pharmgkb.org/pathway/PA2039#, accessed on 08.06.2012) [222]. A detailed description of the figure is given online.

Appendix 2: Appendix to Chapter 2

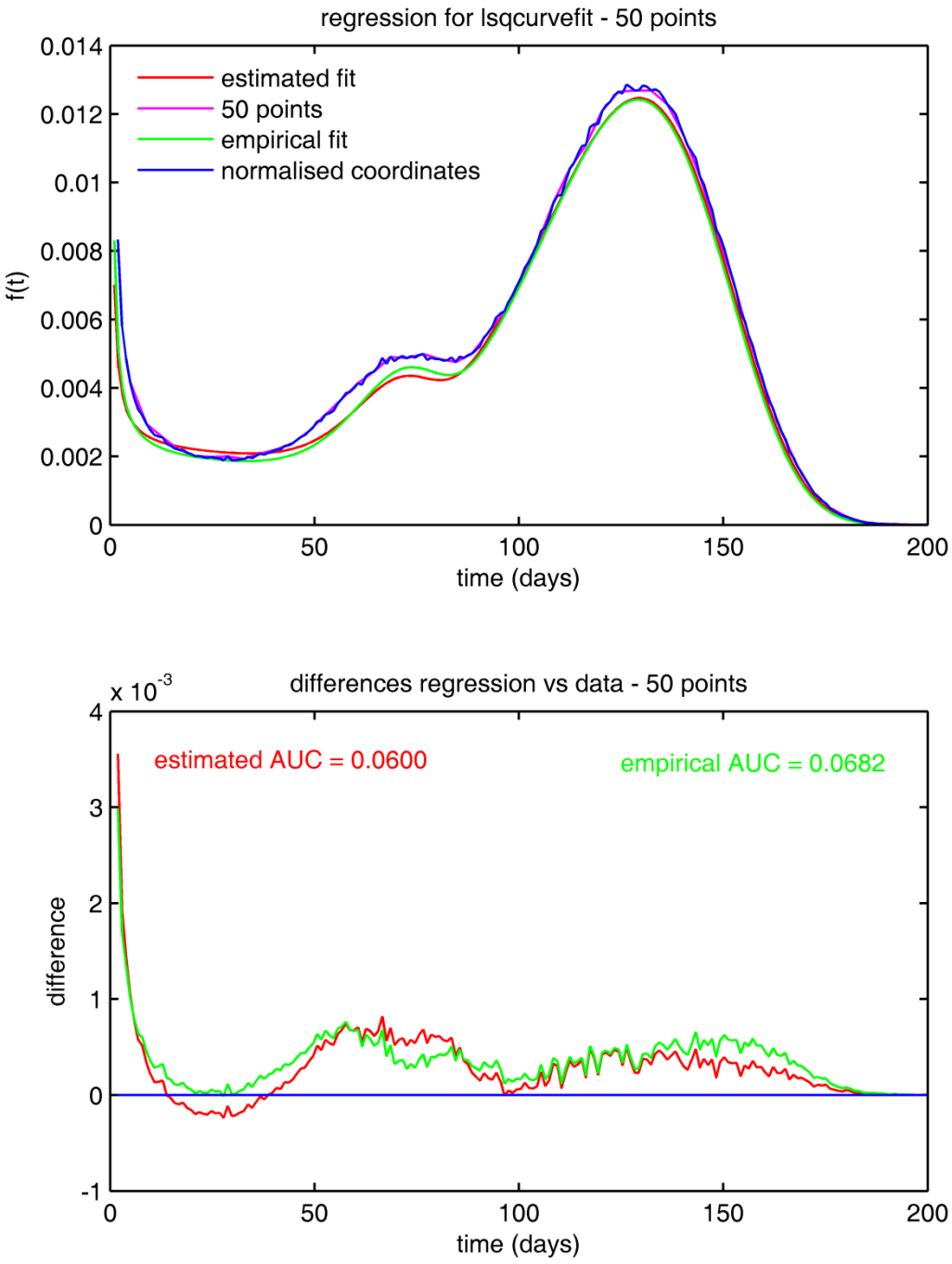
A.2.1. *Estimation of RBC specific parameters – Reduced data sets*



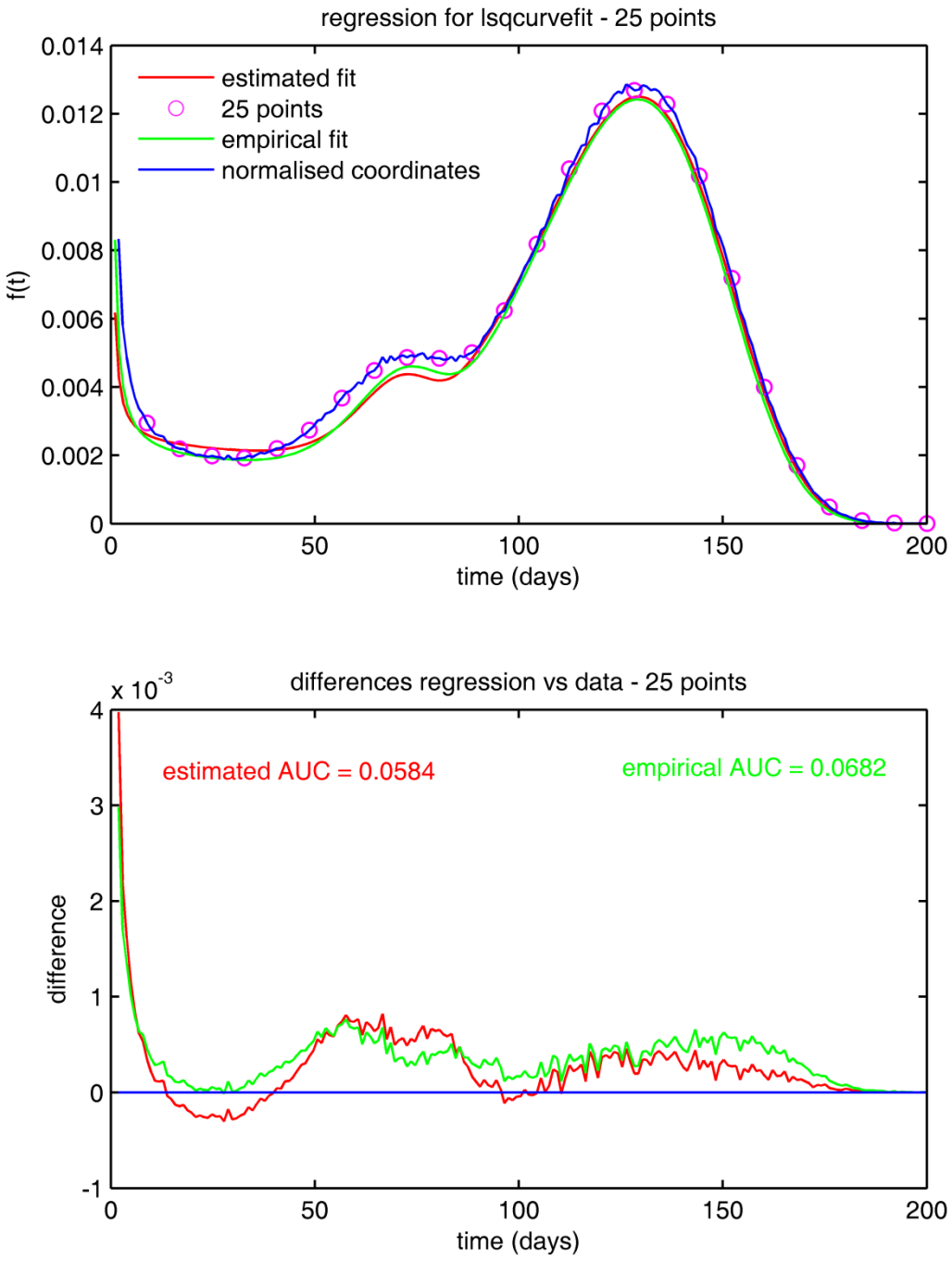
App Fig. 2.1.1: *Upper panel: Profile of the 200 normalised histogram coordinates (blue) compared with the empirically fitted pdf (green) and the estimated pdf (red) based on all 200 points. Lower panel: Difference plots for empirical and estimated pdf fits with area between difference curve and null (AUC)*



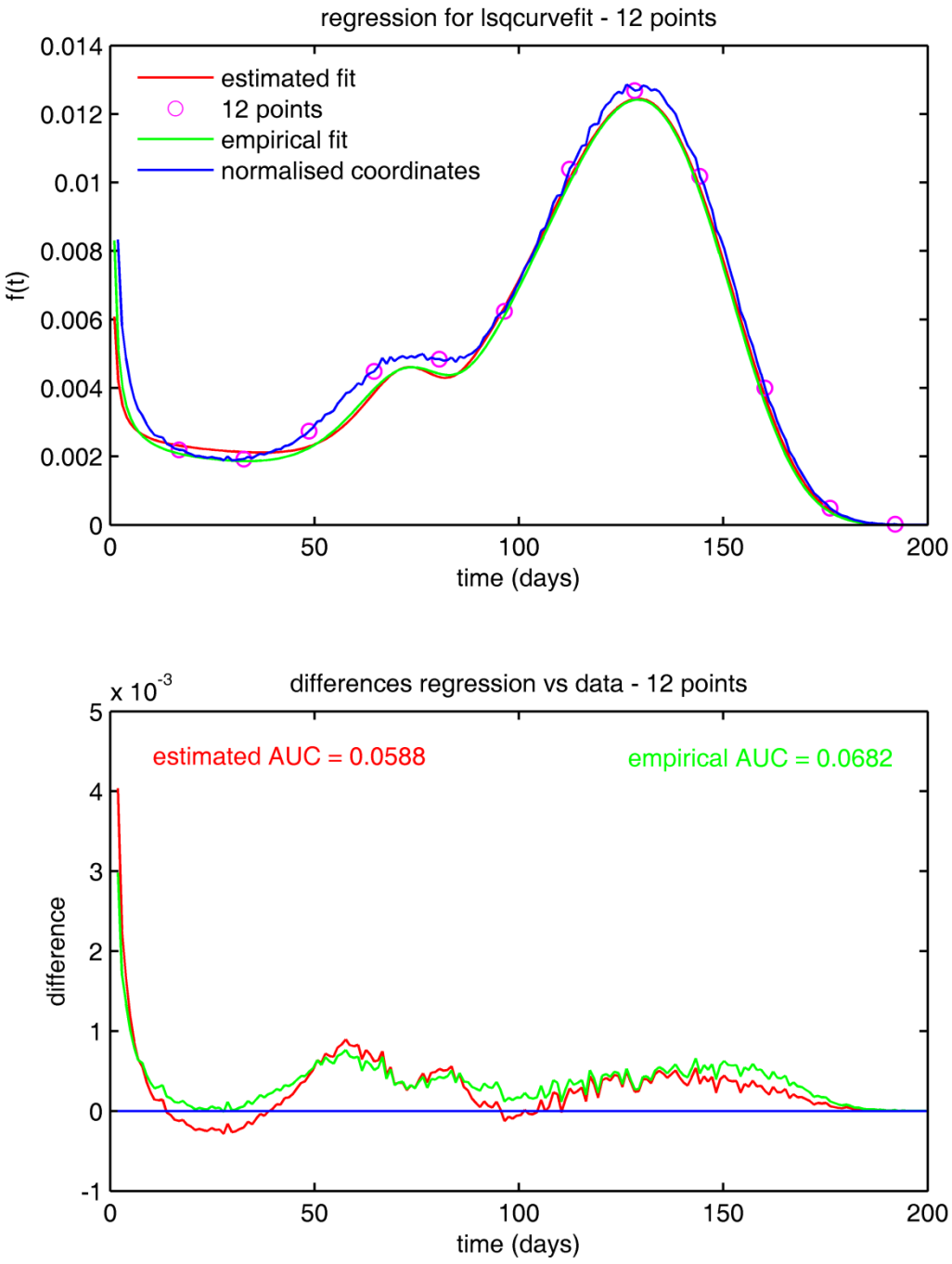
App Fig. 2.1.2: *Upper panel: Profile of the 200 normalised histogram coordinates* (*blue*) *compared with the empirically fitted pdf* (*green*) *and the estimated pdf* (*red*) *based on 100 points* (*pink*). *Lower panel: Difference plots for empirical and estimated pdf fits with area between difference curve and null* (*AUC*).



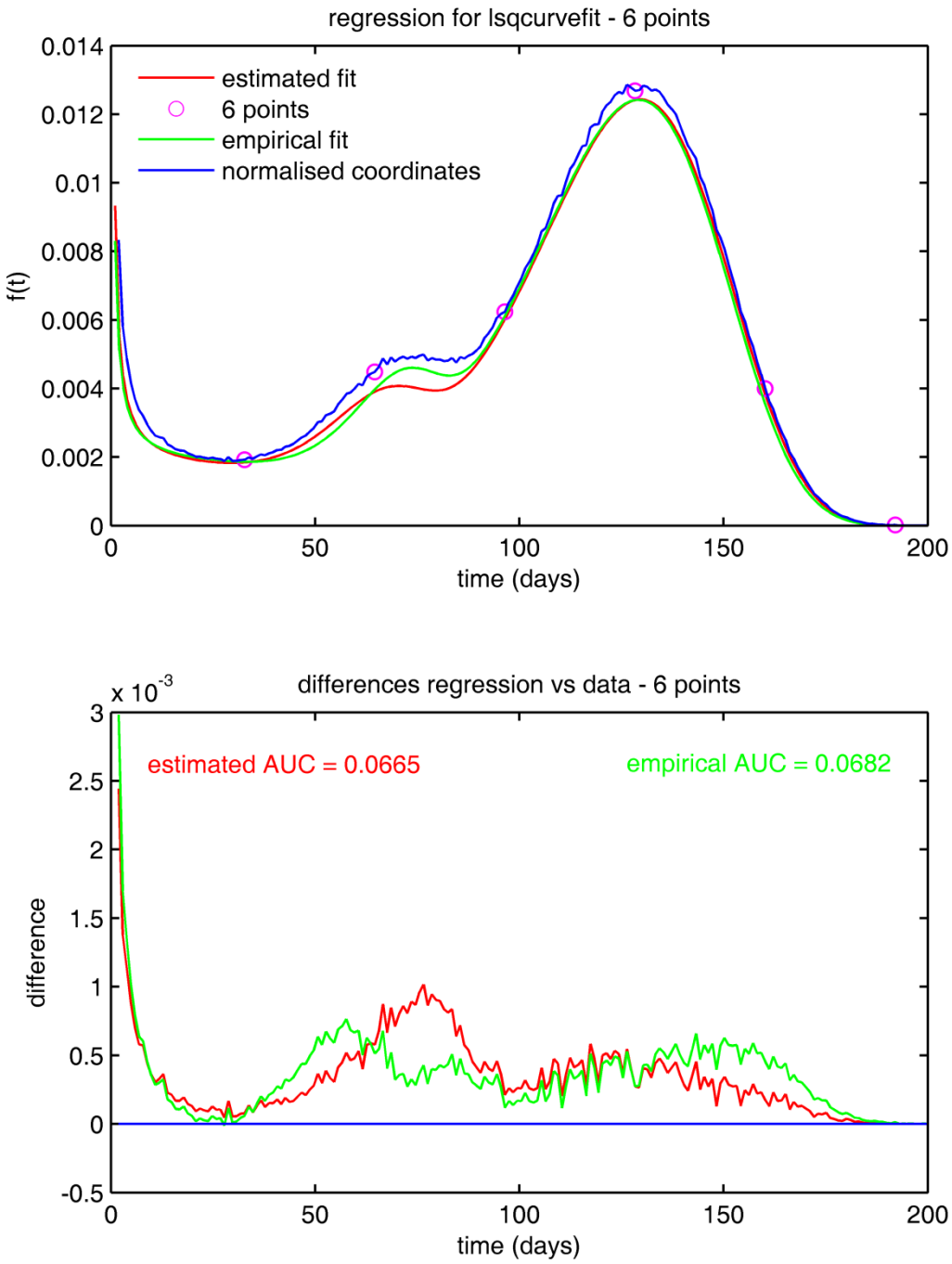
App Fig. 2.1.3: *Upper panel: Profile of the 200 normalised histogram coordinates (blue) compared with the empirically fitted pdf (green) and the estimated pdf (red) based on 50 points (pink). Lower panel: Difference plots for empirical and estimated pdf fits with area between difference curve and null (AUC).*



App Fig. 2.1.4: *Upper panel: Profile of the 200 normalised histogram coordinates (blue) compared with the empirically fitted pdf (green) and the estimated pdf (red) based on 25 points (pink). Lower panel: Difference plots for empirical and estimated pdf fits with area between difference curve and null (AUC).*



App Fig. 2.1.5: *Upper panel: Profile of the 200 normalised histogram coordinates (blue) compared with the empirically fitted pdf (green) and the estimated pdf (red) based on twelve points (pink). Lower panel: Difference plots for empirical and estimated pdf fits with area between difference curve and null (AUC).*



App Fig. 2.1.6: *Upper panel: Profile of the 200 normalised histogram coordinates (blue) compared with the empirically fitted pdf (green) and the estimated pdf (red) based on six points (pink). Lower panel: Difference plots for empirical and estimated pdf fits with area between difference curve and null (AUC).*

A.2.2. *Improved MATLAB® code for the RBC lifespan model*

```
% RBC lifespan model
% 02/09/2009
% Julia Korell


clc
clear all


% Characteristics of experiment
Tmax = 200;         % how long the experiment will run for
Tbirth = 1;         % how long birth will occur for
prod_rate = 1000;   % daily production rate of RBCs
t = 1:1:Tmax;


% load previously generated RBC lifespan values (~1.8Mio.)
load LS_sample_RBCs.mat     % R

all_LS = [];
all_BD = [];
birthdates = 1:Tbirth;      % days of birth


for ij = 1:prod_rate
    % sampling LS for each day with replacement
    lifespan = randsample(R,Tbirth,true);
    all_LS = [all_LS;lifespan'];    % LS matrix
    all_BD = [all_BD;birthdates];   % birthdate matrix
    if rem(ij,100) == 0
       disp(ij)
    end
end

% death of each RBC
clocklife = all_LS + all_BD;


for ijk = 1:Tmax
       for ij = 1:prod_rate
           for jk = 1:Tbirth
               % check for being born
               if jk > ijk
                    alive(ij,jk) = 0;
               % check for being dead
               else if clocklife(ij,jk) < ijk
```

```
                        alive(ij,jk) = 0;
                    else
                        alive(ij,jk) = 1;
                    end
                end
            end
        end

     % living RBCs per day
     total_alive(ijk) = sum(sum(alive));
     if rem(ijk,100) == 0
        disp(ijk)
    end
end


figure(1)
plot(t,total_alive)
xlabel('time (days)')
ylabel('number of RBCs')


% lifespan distribution in experiment
LS_dist = reshape(all_LS,1,numel(all_LS));

figure(2)
hist(LS_dist,100)
xlabel('lifespan (days)')
ylabel('number of sampled lifespans')
```

A.2.3. Exchange algorithm implemented in MATLAB®

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%                 GENERALIZED EXCHANGE ALGORITHM                       %
%                    04/03/2010  Julia Korell                          %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [XOPT OBJ] = exchange(FUN,XSPACE,NX,options,varargin)

% EXCHANGE finds those values XOPT that maximize the objective
% function FUN over the discrete space XSPACE
%
% Required input arguments:
%   FUN     objective function (needs to be specified as @FUN)
%   XSPACE  discrete space to optimize over
%   NX      number of values to optimize for
%
% Available options (define in structure options):
%   maxit   number of maximum iterations (default = 1e4)
%
% Define problem dependent variables V1,V2... as
%   XOPT=EXCHANGE(@FUN,XSPACE,NX,options,V1,V2,...)
% Pass options=[] to use default settings.
%
% Display of intermediate results can be switched on/off in
% lines 70/71.
% Display of counter of iterations can be switched on/off or
% modified in lines 82/83.


tic

if nargin < 3, error('requires at least three input arguments');
end
if nargin < 4
    options = struct('maxit',1e4);
else
    if ~isfield(options,'maxit')
        options.maxit = 1e4;
    end
end


maxit = options.maxit;      % maximum number of iterations
iter = 0;
numex = 0;
```

```
x0 = XSPACE(1:NX);
OBJ = feval(FUN,x0,varargin{:});


for i = 1:1:maxit

    if i == 1
        xex(i,:) = x0(randperm(length(x0)));
        X = x0;
    else
        Xold = xex(i-1,:);
        xex(i,:) = Xold(randperm(length(Xold)));
        X = xex(i,:);
    end

     % exchange time points and evaluate objective function
    for j = 1:1:length(X)
        for k = 1:1:length(XSPACE)
            xex(i,j) = XSPACE(k);
            OBJ_new = feval(FUN,xex(i,:),varargin{:});

     % accept new value if OBJ_new is bigger than the previous
            if OBJ_new > OBJ
                    OBJ = OBJ_new;
%                 disp('updated objective function value')
%                 disp(OBJ)
                numex = numex+1;
                break
            else
                xex(i,j) = X(j);
            end
        end
    end

    iter = iter+1;
    if rem(iter,100)==0
        disp('iteration')
        disp(iter)
    end

    if i>1            % no further changes
        if sort(xex(i,:),2) == sort(xex(i-1,:),2)
            break
        end
    end
end

XOPT = sort(xex(i,:));      % optimal x-values
```

```
tElapsed = toc;

% Display results
if iter == maxit
    disp ('maximum number of iterations reached')
    disp ('time elapsed')
    disp (tElapsed)
else
    disp('number of iterations')
    disp(iter)
    disp('number of exchanges')
    disp(numex)
    disp('time elapsed (sec)')
    disp(tElapsed)
    disp('optimal x-values')
    disp(XOPT)
    disp('objective function value')
    disp(OBJ)
end
```

Appendix 3: Appendix to Chapter 3

A.3.1. *Equivalent functions of survival time for the RBC model*

According to the introduction on survival analysis in this thesis (Section 1.1.3), the survival time can be expressed as different mathematical functions which are all mathematically equivalent, such as the survival function $S(t)$, the hazard function $h(t)$, and the probability density distribution (pdf) $f(t)$.

Here, the equations of these functions with respect to the proposed RBC survival model are presented, where *FW* stands for Flexible Weibull and *RAW* for Reduced Additive Weibull, describing the two distributions underlying the combined model [83]. Note, the full notation of the hazard function is not shown due to its high level of complexity.

$$\begin{aligned} f(t) &= m \cdot f_{FW}(t) + (1-m) \cdot f_{RAW}(t) \\ &= m \cdot \left(exp\left(-exp(s_1 \cdot t - s_2/t) - c \cdot t\right) \cdot \left((s_1 + s_2/t^2) \cdot exp(s_1 \cdot t - s_2/t) + c \right) \right) + \\ &\quad (1-m) \cdot exp\left(-(r_1 \cdot t)^{r_2} - (r_1 \cdot t)^{1/r_2} \right) \cdot \left((r_1 \cdot r_2)(r_1 \cdot t)^{r_2 - 1} + (r_1/r_2)(r_1 \cdot t)^{(1/r_2)-1} \right) \end{aligned}$$

App Eq. 3.1.1: *Probability density function f(t) of the RBC lifespan model proposed in Chapter 2.*

$$\begin{aligned} S(t) &= m \cdot S_{FW}(t) + (1-m) \cdot S_{RAW}(t) \\ &= m \cdot \left(\exp\left(-\exp(s_1 \cdot t - s_2/t) - c \cdot t\right) \right) + (1-m) \cdot \left(\exp\left(-(r_1 \cdot t)^{r_2} - (r_1 \cdot t)^{1/r_2} \right) \right) \end{aligned}$$

App Eq. 3.1.2: *Survival function S(t) of the RBC survival model proposed in Chapters 3&4.*

$$h(t) = \frac{f(t)}{S(t)} = \frac{m \cdot f_{FW}(t) + (1-m) \cdot f_{RAW}(t)}{m \cdot S_{FW}(t) + (1-m) \cdot S_{RAW}(t)}$$

App Eq. 3.1.3: *Corresponding hazard function h(t) of the proposed model.*

A.3.2. *Loss of label due to vesiculation, decay and elution*

The following equations describe the loss of a label based on vesiculation, radioactive decay and elution only. Based on these equations the loss of label in addition to the loss due to death of RBCs has been included into the model in Chapter 3 (Sections 3.4.2 & 3.4.3).

$$N_v(t) = N_0 - k_v \cdot t^v$$

App Eq. 3.2.1: *Loss of label due to vesiculation.*

$$N_d(t) = N_0 \cdot 2^{-t / t_{1/2Cr}}$$

App Eq. 3.2.2: *Loss of label due to radioactive decay, e.g. of ^{51}Cr.*

$$N_e(t) = N_0 \cdot 2^{-t / t_{1/2el}}$$

App Eq. 3.2.3: *Loss of label due to elution.*

where N_0 is the total number of labelled RBCs at time 0, and $N_x(t)$ is the remaining number of labelled cells at time t after loss due to the individual process vesiculation $N_v(t)$, decay $N_d(t)$ or elution $N_e(t)$, while k_v , v, $t_{1/2Cr}$ and $t_{1/2el}$ are defined as in Chapter 3.

A.3.3. *Final MATLAB® codes for the RBC survival model*

The following code describes the final RBC survival model based on the proposed survival function and can account for flaws associated with certain RBCs labelling methods, such as reuse of the label, radioactive decay, vesiculation and elution.

A.3.3.1. General code for the survival function

```
function smix = survival_mix(para,t)

% survival function of the mixed model
% parameter vector para needs 6 input items, ordered according
% to a,b,x,y,z,p

a = para(1);
b = para(2);
x = para(3);
y = para(4);
z = para(5);
p = para(6);


smix = p.*(exp(-exp(x.*t-y./t)-z.*t))+(1-p)*(exp(-((a.*t).^b) -
(a.*t).^(1/b)));
```

A.3.3.2. Code for a cohort labelling method

```
function Nlabel = survival_model_cohort_reuse(t,a,b,x,y,z,p,
Tbirth, prodrate, reuse_fraction,P)

% Function for the survival model evaluated for teval
% For a cohort labelling method with reuse
% 07/04/2010
% Julia Korell

% Calls "survival_mix"

%%% SETTINGS %%%
% t = time course of simulation, needs to be continuous over
% time!!!
% Tbirth = time until production occurs
% prodrate = constant daily production rate
% reuse_fraction = fraction of lost label reused on following
% day
% P = controls vesiculation (0=none, 1=linear, 2=increasing)

para = [a b x y z p];
Tmax  = max(t);             % max. time for simulation
prod = zeros(1,Tmax);
```

```
prod(1:Tbirth) = prodrate;
median_LS = 115;        % median RBC lifespan in days
t80 = median_LS;        % time of 80% loss due to vesiculation

if P == 0
    kves = 0;           % no vesiculation
else
    kves = 0.2/(t80^P); % vesiculation rate per day (loss of Hb)
end

%%% SIMULATION %%%

for i = 1:1:Tmax        % i = day of birth

    if i==1
        reuse(i) = 0;   % no reuse on first day
    else                % loss from day before
        loss(i) = sum(N_ves(1:i-1,i-1))-sum(N_ves(1:i-1,i));
        reuse(i) = loss(i)*reuse_fraction;     % fraction reused
    end

    ti = t-i;           % time course of survival
        S(1:(i-1)) = 0;     % not born yet
    S(i:Tmax) = survival_mix(para,ti(i:Tmax));
                                        % survival of fraction
    N(i,:) = (prod(i)+reuse(i))*S;

    % Vesiculation
    ves(1:(i-1)) = 0;
    ves(i:Tmax) = (1-kves*(ti(i:Tmax)).^P);
    N_ves(i,:) =  N(i,:).*ves;

end

Nlabel = sum(N_ves);   % #RBCs corrected for vesiculation
```

A.3.3.1. Code for a random labelling method

```
function Nlabel_vde = survival_model_random_loss(t,a,b,x,y,z,p,
Tbirth,prodrate,half_life_decay,half_life_elution,P)

% Function for the survival model with a random label evaluated
% for time t
% Accounting vesiculation, decay & elution to occur
% 08/04/2010
% Julia Korell

% RETURNS ONLY DISAPPEARANCE OF LABEL (Tbirth:Tmax)!!!
% Calls "survival_mix"

%%% SETTINGS %%%

% t = time course of simulation, needs to be continuous over
% time!!!
```

```
% Tbirth = time until production occurs = day of random
% labelling (Tlabel)!
% prodrate = constant daily production rate
% half_life_decay = radioactive decay of label (0=none)
% half_life_elution = elution of label (0=none)
% P = vesiculation (0=none, 1=linear, 2=increasing with age)

para = [a b x y z p];
Tmax  = max(t);          % max. time for simulation
prod = zeros(1,Tmax);
prod(1:Tbirth) = prodrate;
median_LS = 115;         % median RBC lifespan in days
t80 = median_LS;         % time of 80% loss due to vesiculation
Tlabel = Tbirth;         % time of labelling = stop of production
t_lab = 0:Tmax-Tlabel;   % days after labelling


if half_life_decay == 0
    decay = 1;           % no decay of label
else
    decay = 2.^(-(t_lab./half_life_decay));   % decay per day
end

if half_life_elution == 0           % no elution of label
    elution = 1;
else
    elution = 2.^(-(t_lab./half_life_elution));
                                                % elution per day
end

if P ==0
    kves = 0;            % no vesiculation
else
    kves = 0.2/(t80^P); % vesiculation rate per day (loss of Hb)
end


%%% SIMULATION %%%

for i = 1:1:Tmax
    ti = t-i;

    S(1:(i-1)) = 0;                                % not born yet
    S(i:Tmax) = survival_mix(para,ti(i:Tmax));
                                              % survival of fraction

    N(i,:) = prod(i)*S;

    % Vesicualtion
    ves(1:(i-1)) = 0;
    ves(i:Tmax) = (1-kves*(ti(i:Tmax)).^P);
    N_ves(i,:) =  N(i,:).*ves;
end

Ntot = sum(N);           % summing up cells per day
Ntot_ves = sum(N_ves);   % #RBCs corrected for vesiculation
```

```
% Disappearance of label
Nlabel_ves = Ntot_ves(Tlabel:Tmax);
Nlabel_vd = Nlabel_ves.*decay;
Nlabel_vde = Nlabel_vd.*elution;
                                % observed disappearance of label
```

A.3.4. *Hypercube design*

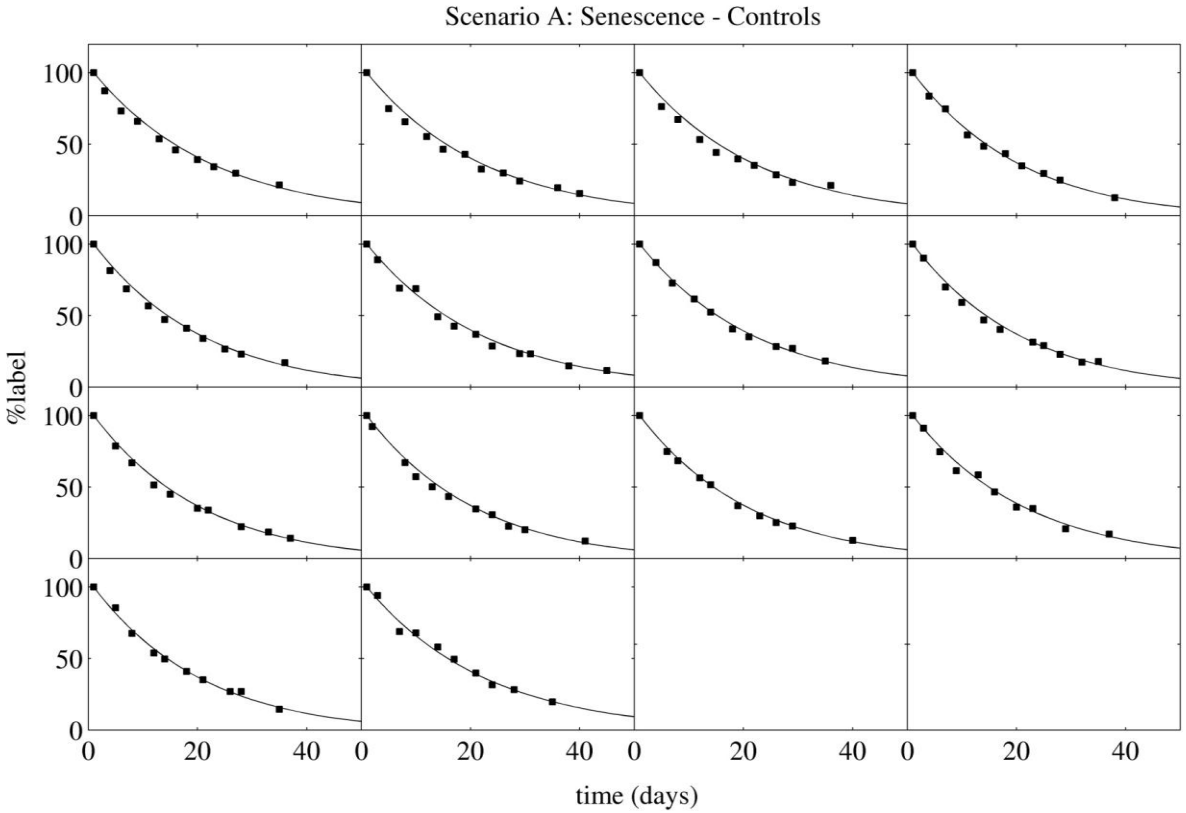
The following table presents the 2^6 possible combinations of the lower (L) and upper (H) point estimates giving the vertexes in the hypercube design space for parameter estimation.

App Tab. 3.4.1: *Vertexes of the hypercube design space for the robust optimal design.*

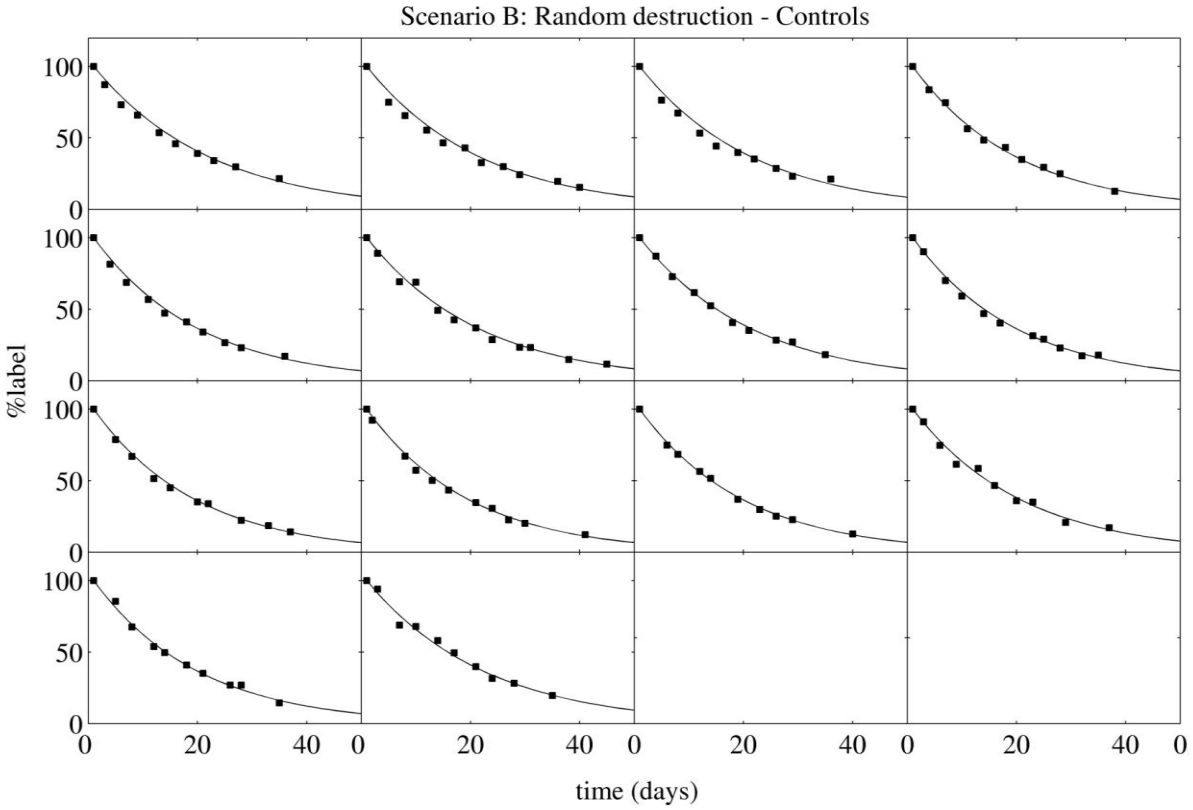
Vertex	r_1	r_2	s_1	s_2	c	m	*Vertex*	r_1	r_2	s_1	s_2	c	m
1	H	L	L	L	L	L	33	L	L	L	L	L	L
2	H	L	L	L	L	H	34	L	L	L	L	L	H
3	H	L	L	L	H	L	35	L	L	L	L	H	L
4	H	L	L	L	H	H	36	L	L	L	L	H	H
5	H	L	L	H	L	L	37	L	L	L	H	L	L
6	H	L	L	H	L	H	38	L	L	L	H	L	H
7	H	L	L	H	H	L	39	L	L	L	H	H	L
8	H	L	L	H	H	H	40	L	L	L	H	H	H
9	H	L	H	L	L	L	41	L	L	H	L	L	L
10	H	L	H	L	L	H	42	L	L	H	L	L	H
11	H	L	H	L	H	L	43	L	L	H	L	H	L
12	H	L	H	L	H	H	44	L	L	H	L	H	H
13	H	L	H	H	L	L	45	L	L	H	H	L	L
14	H	L	H	H	L	H	46	L	L	H	H	L	H
15	H	L	H	H	H	L	47	L	L	H	H	H	L
16	H	L	H	H	H	H	48	L	L	H	H	H	H
17	H	H	L	L	L	L	49	L	H	L	L	L	L
18	H	H	L	L	L	H	50	L	H	L	L	L	H
19	H	H	L	L	H	L	51	L	H	L	L	H	L
20	H	H	L	L	H	H	52	L	H	L	L	H	H
21	H	H	L	H	L	L	53	L	H	L	H	L	L
22	H	H	L	H	L	H	54	L	H	L	H	L	H
23	H	H	L	H	H	L	55	L	H	L	H	H	L
24	H	H	L	H	H	H	56	L	H	L	H	H	H
25	H	H	H	L	L	L	57	L	H	H	L	L	L
26	H	H	H	L	L	H	58	L	H	H	L	L	H
27	H	H	H	L	H	L	59	L	H	H	L	H	L
28	H	H	H	L	H	H	60	L	H	H	L	H	H
29	H	H	H	H	L	L	61	L	H	H	H	L	L
30	H	H	H	H	L	H	62	L	H	H	H	L	H
31	H	H	H	H	H	L	63	L	H	H	H	H	L
32	H	H	H	H	H	H	64	L	H	H	H	H	H

Appendix 4: Appendix to Chapter 4

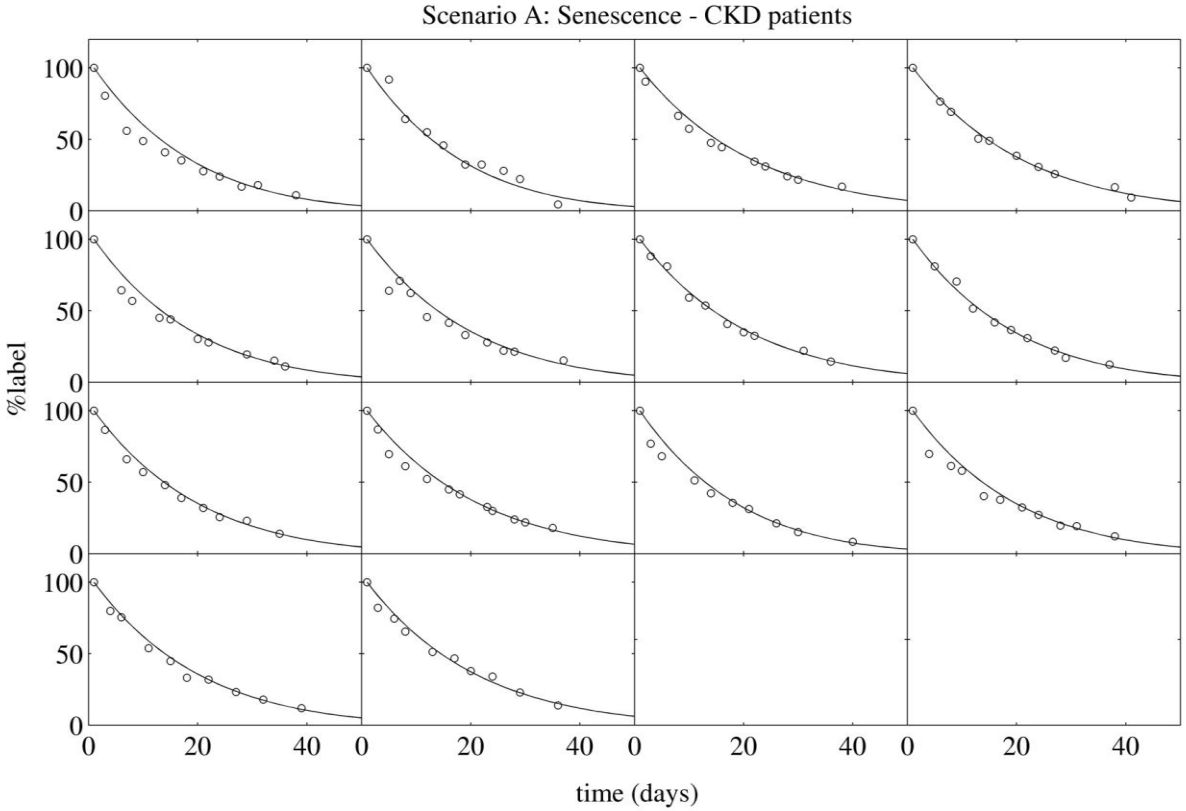
A.4.1. *Individual fits obtained in the two-stage approach [223]*



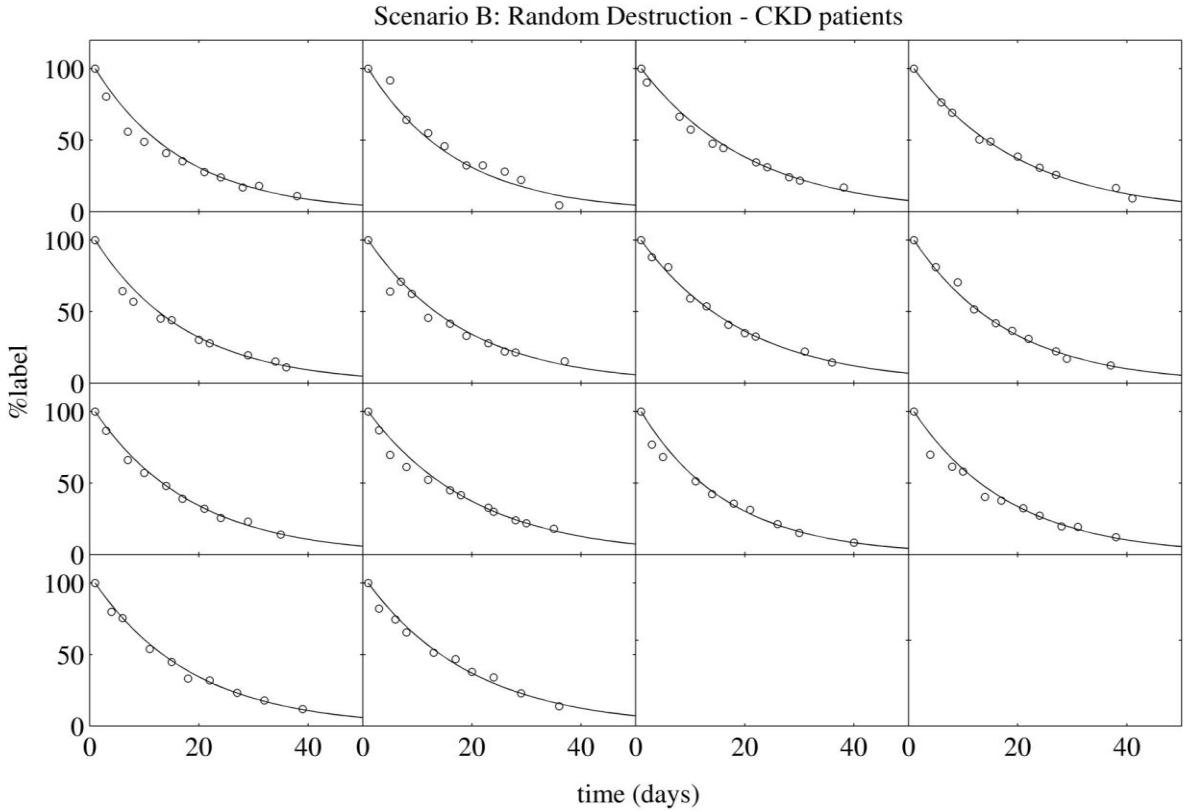
App Fig. 4.1.1: *Individual fits obtained in the two-stage approach for estimating senescence in controls.*



App Fig. 4.1.2: *Individual fits obtained in the two-stage approach for estimating random destruction in controls.*



App Fig. 4.1.3: *Individual fits obtained in the two-stage approach for estimating senescence in CKD patients.*



App Fig. 4.1.4: *Individual fits obtained in the two-stage approach for estimating random destruction in CKD patients.*

A.4.2. *Additional analysis – Testing rHuEPO as covariate*

A.4.2.1. Materials & Methods

Out of the 14 patients with chronic kidney disease (CKD) on haemodialysis studied by Vos *et al*. [165], eleven received human recombinant erythropoietin (rHuEPO), two of those with additional iron supplementation, while three patients received iron supplementation alone without administration of rHuEPO or any other erythropoiesis stimulating agent. It should be noted that there is a high concordance between CKD and rHuEPO and although not 100% the results delineating these components should be viewed with caution.

Based on the population approach using MONOLIX® 1.1 that was described in Chapter 4 (Section 4.4.4), it was tested whether the application of rHuEPO is a significant covariate on RBC survival. Two models were considered: 1) including rHuEPO on its own as a covariate (rHuEPO model); and 2) including rHuEPO and CKD as covariates together (rHuEPO & CKD model). Significance of the covariates compared to the base model without covariate effect was assessed globally based on the likelihood ratio test (LRT) and locally based on the Wald test for the corresponding covariate coefficients β_{EPO} and β_{CKD} (please refer to Section 4.4.4, Equation 4.10 for the implementation of the covariate model in MONOLIX®) .

As before, it was also investigated based on the objective function value (OFV) which destruction mechanism is preferred for these new models: senescence (Scenario A) or random destruction (Scenario B).

A.4.2.2. Results

An overview of the results for this analysis is given in App Tab. 4.2.1. With respect to the preference of destruction mechanism, again estimating random destruction is preferred over estimating senescence based empirically on the higher OFVs across all models tested in this additional analysis for Scenario B.

Inclusion of rHuEPO as a covariate was found to be significant based on both, Wald test as well as LRT. However, CKD was found not to be a significant covariate, both locally as well as globally, once rHuEPO had been added as covariate in the model.

These results are discussed further in Chapter 4 (Section 4.7.1).

App Tab. 4.2.1: *Fixed and random effect parameter estimates (%SE) for the population approach – Testing rHuEPO as covariate.*

Parameter	Estimates			Description
Scenario A: *Estimating* $\hat{\bar{s}}_2$				
	Base model	rHuEPO model	rHuEPO & CKD model	
$\hat{\bar{s}}_2$ (days)	162.49 (6.9)	200.45 (5.6)	199.89 (6.3)	population mean estimate
β_{CKD}	-	-	0.0493 (304 *ns*)	covariate factor of CKD
β_{EPO}	-	-0.5399 (16.2)	-0.4984 (30.9)	covariate factor of EPO
Ω	0.1117 (34.1)	0.0385 (39.0)	0.0400 (40.3)	between subject variance
$\bar{L}$ (days)	62.8	69.21	69.12	mean RBC lifespan
$\bar{L}_\beta$ (days)	-	53.81	53.55	mean RBC lifespan with covariate effect(s)
σ^2_{add}	2.96 (12.5)	2.54 (12.2)	3.14 (12.9)	additive error (variance)
CV_{prop}	0.0251 (12.5)	0.0295 (12.2)	0.0243 (12.9)	proportional error (coefficient of variation)
LL	-780.10	-769.65	-769.62	log likelihood

(*ns*) = Parameter not significant based on Wald test

App Tab. 4.2.1: *Continued.*

Parameter	Estimates			Description
Scenario B: *Estimating* $\hat{\bar{c}}$				
	Base model	rHuEPO model	rHuEPO & CKD model	
$\hat{\bar{c}}$ (days^{-1})	0.0133 (7.3)	0.0109 (6.4)	0.0107 (6.5)	population mean estimate
β_{CKD}	-	-	0.0113 (310.3 *ns*)	covariate factor of CKD
β_{EPO}	-	0.5449 (17.6)	0.5036 (33.1)	covariate factor of EPO
Ω	0.1296 (18.6)	0.0496 (37.1)	0.0504 (36.9)	between subject variance
$\bar{L}$ (days)	56.0	62.74	63.15	mean RBC lifespan
$\bar{L}_{\beta}$ (days)	-	44.78	44.79	mean RBC lifespan with covariate effect(s)
σ^2_{add}	2.27 (12.4)	2.25 (12.7)	2.18 (12.6)	additive error (variance)
CV_{prop}	0.0234 (12.4)	0.0240 (12.7)	0.0247 (12.6)	proportional error (coefficient of variation)
LL	-752.12	-742.10	-742.03	log likelihood

(*ns*) = Parameter not significant based on Wald test

A.4.3. *Additional analysis for patients on peritoneal dialysis*

A.4.3.1. Materials & Methods

Vos *et al*. studied an additional cohort of five CKD patients undergoing peritoneal dialysis (PD) [165], but did not recruit age and sex matched controls for these patients All PD patients received rHuEPO treatment, one with additional iron supplementation.

Here, the RBC survival in these patients was estimated based on the two-stage approach described in Chapter 4 (Section 4.4.3). The resulting apparent mean RBC lifespan was compared with the results of the healthy controls as well as the CKD cohort undergoing haemodialysis (HD) based on a two-tailed t-test assuming unequal variances. Again, it was also tested whether an accelerated senescence (Scenario A) or an increased random destruction (Scenario B) is preferred as underlying destruction mechanism in these patients.

A.4.3.2. Results

App Tab. 4.3.1 shows an overview of the results for this analysis. The calculated apparent mean RBC lifespan falls between those estimated for controls and HD patients for both scenarios. No statistically significant difference compared to either of these two groups was observed.

In two out of the five patients estimation of senescence was preferred, while the remaining three had a better fit when estimating random destruction. Due to the small number of patients in this cohort, the slight preference towards random destruction is less conclusive as for the controls and HD patients, and the PD patients were not included in the population analysis in Chapter 4.

The results of this two-stage analysis are discussed further in Chapter 4 (Section 4.7.2).

App Tab. 4.3.1: *Results for the PD patients using the two-stage approach.*

	PD group	p value *vs.* controls	p value *vs.* HD patients	Description
Scenario A: *Estimating senescence* $\bar{s}_2$				
$\bar{s}_2$ (days)	170.00	0.0529	0.4959	population mean
Ω_{s2} (days^2)	4059.76			BSV of population mean*
$\bar{L}_{s_2}$ (days)	63.2	0.0663	0.4960	mean RBC lifespan
$\Omega_{L,s2}$ (days^2)	120.1			BSV of mean lifespan*
Scenario B: *Estimating random destruction* $\bar{c}$				
$\bar{c}$ (days^{-1})	0.0146	0.1252	0.4582	population mean
Ω_c (days^{-2})	3.44 x 10^{-5}			BSV of population mean*
$\bar{L}_c$ (days)	55.4	0.0976	0.4483	mean RBC lifespan
$\Omega_{L,c}$ (days^2)	195.2			BSV of mean lifespan*

* BSV = between subject variance

Appendix 5: Appendix to Chapter 5

A.5.1. *Imputation of missing covariates*

In the data set of the oral study, height was not recorded for eight individuals; however this covariate was required for the calculation of BMI in this analysis. Therefore, multiple linear regression was conducted in the remaining 40 patients and a single value for height was imputed for these eight individuals based on the parameter values from this regression analysis.

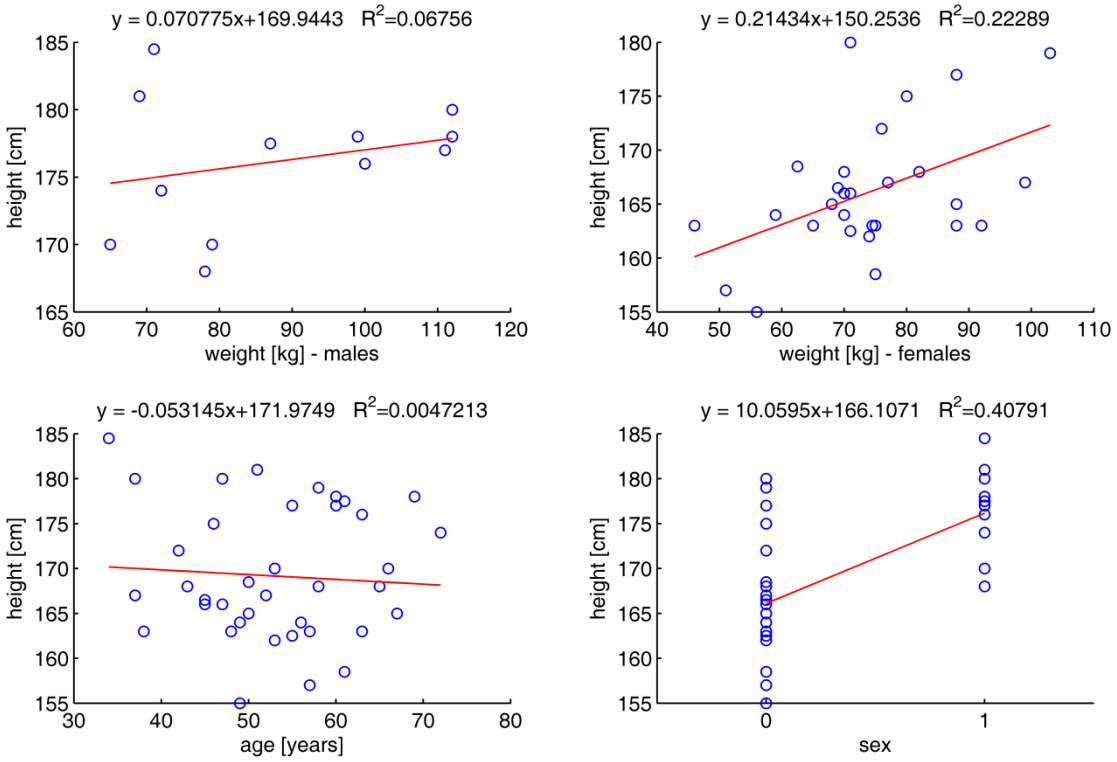
Firstly, the correlation between height and other covariates in the data set was determined to obtain reasonable regression variables. For this, a simple linear regression was conducted in MATLAB®. Weight, sex (0 = females, 1 = males) and age were chosen as initial predictive variables. App Fig. 5.1.1 shows the corresponding regressions of height against these covariates. Note, that for weight, the regression analysis was split into males and females (i.e. resulting in a multiple linear regression analysis), as a difference between both sexes in the relationship of height with weight was expected *a priori*.

Based on these results, weight and sex were chosen as regression variables and a multiple linear regression was conducted without (App Fig. 5.1.2 and App Fig. 5.1.2) and with an interaction (App Fig. 5.1.3 and App Fig. 5.1.3) between weight and sex.

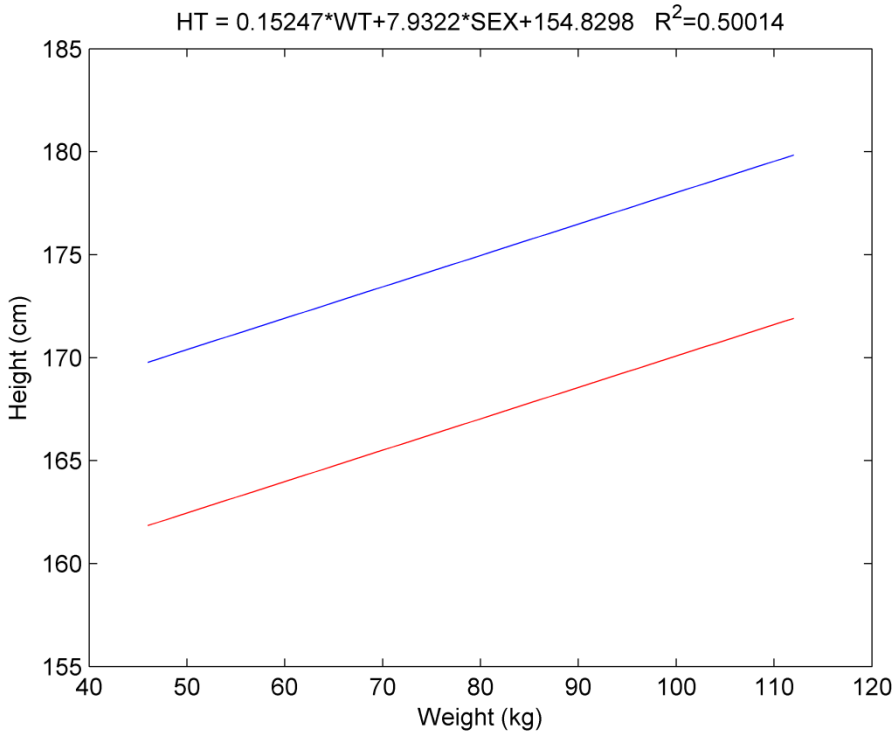
The regression including an interaction term between sex and weight resulted in a better correlation (R^2 = 0.52 compared to R^2 = 0.50 without interaction). Therefore, the regression with interaction was used to impute the missing values for height in the data set:

$$HT = 0.21 \cdot WT + 19.69 \cdot SEX - 0.14 \cdot WT \cdot SEX + 150.25$$

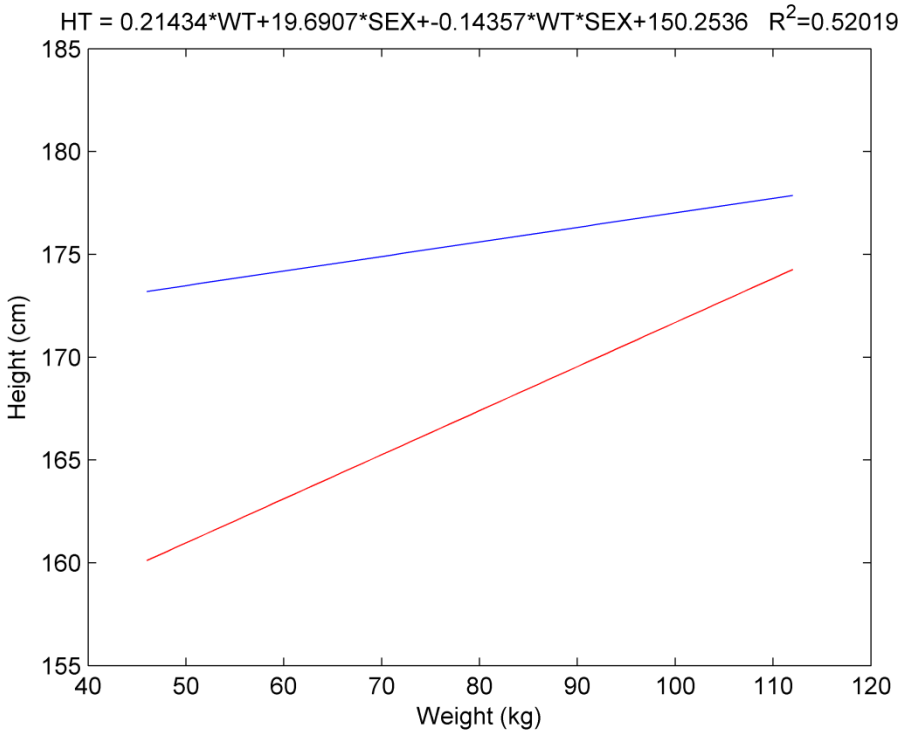
App Eq. 5.1.1: *Imputation of height (HT) based on weight (WT) and sex, where SEX = 0 for females and SEX = 1 for males.*



App Fig. 5.1.1: *Correlation between height and other covariates in the data set.*



App Fig. 5.1.2 *&* ***App Eq. 5.1.2:*** *Regression of height (HT) against weight (WT) and sex without interaction. Blue = males (SEX = 1), red = females (SEX = 0).*



App Fig. 5.1.3 & App Eq. 5.1.3: *Regression of height (HT) against weight (WT) and sex with interaction. Blue = males (SEX = 1), red = females (SEX = 0).*

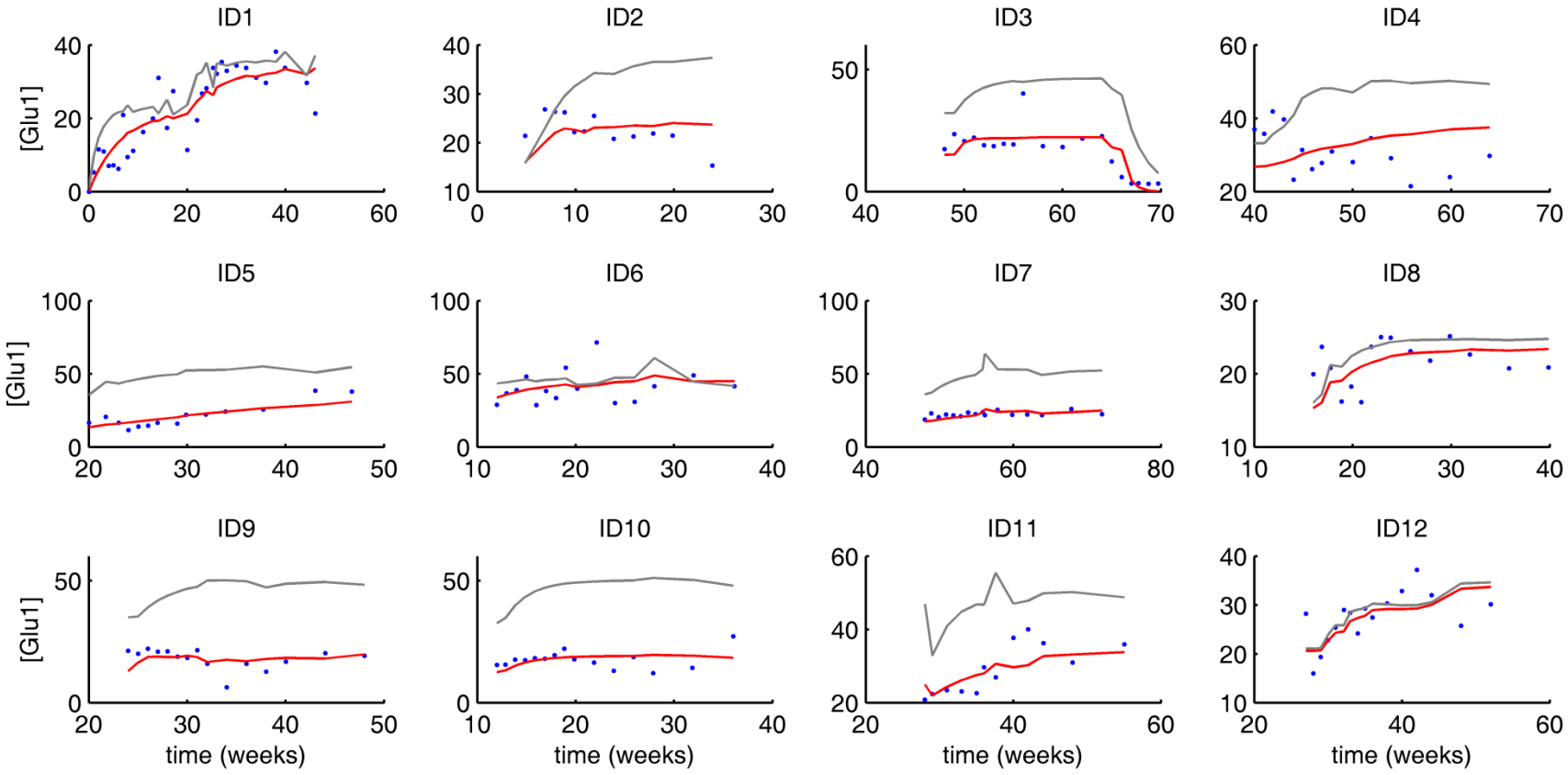
A.5.2. *NONMEM® code for the final parent drug model*

```
;RBC MTXGlu1 parent model
;combined data set BASED ON AVERAGE of replicate measurements
;M3 to handle BLQ: compute joint LH
;PK model - parent:
;2cmpt PK model for MTX based on article by Hoekstra et al.
;(2004)
;combined error model
;first-order uptake and elimination from RBCs, elimination coded
;as clearance
;LBW, [Hb] and MCV as covariates on VRBCs
;no BSV on CLRBCs


$PROBLEM   MTXGlu1 MBP41 M3


$INPUT     ID TIME AMT ROUTE II ADDL DVID CMTX OCC DV=AVE MDV
           EVID BLQ STUDY STARTSTOP HB MCV HEIGHT WEIGHT AGE SEX

$DATA      MTX_data.csv IGNORE=#
           IGNORE=(DVID.GE.2,OCC.EQ.2)

$SUBROUTINE ADVAN5


$MODEL     NCOMP=4
      COMP=(ABS,DEFDOSE)
      COMP=(PLASMA)
      COMP=(PERIPH)
      COMP=(RBCGLU1) ; Parent drug MTXGlu1 in RBCs

$PK
      ; COMPARTMENTS
      IF(DVID.EQ.0) THEN
            CMT=1      ; dose => absorption compartment
      ELSEIF(DVID.EQ.1) THEN
            CMT=4      ; observation => parent MTXGlu1 in RBCs
      ENDIF

      ; LIMIT OF QUANTIFICATION
      LOQ = 5          ; limit of quantification (nmol/LRBCs)

      ; PLASMA PK FIXED: mean parameter estimates Hoekstra et al
      IF(ROUTE.EQ.0) THEN
            F1 = 0.7               ; oral dosing
            K12 = 0.87             ; (/hrs) oral absorption
            ALAG1 = 0.36           ; lag-time after oral
      ELSEIF(ROUTE.EQ.1) THEN
            F1 = 1                 ; sc dosing
            K12 = 0.36             ; (/hrs) sc absorption
            ALAG1 = 0.06           ; lag-time after sc
      ENDIF
```

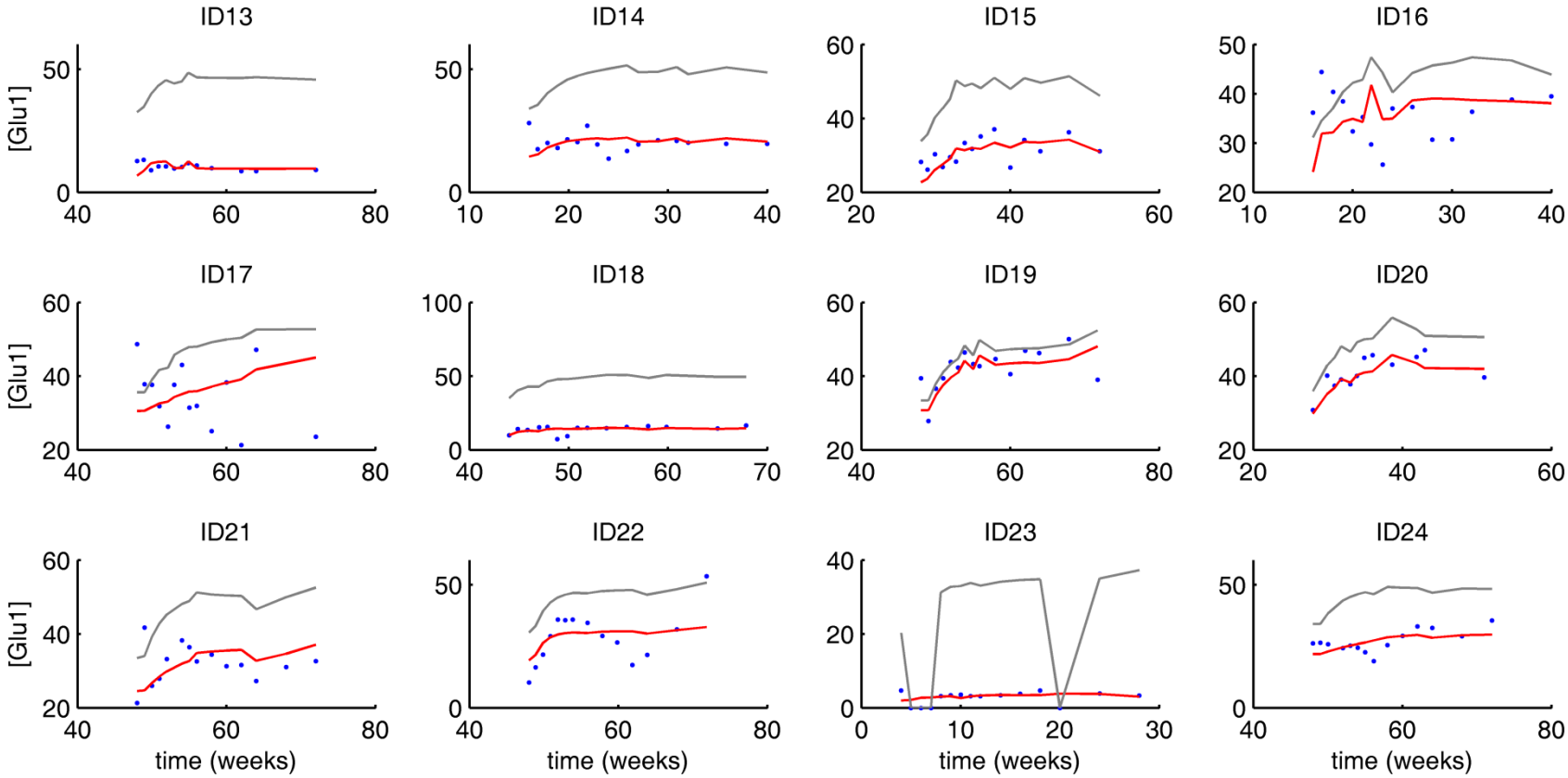
```
      K23 = 0.81       ; (/hrs) distribution into peripheral
      K32 = 0.55       ; (/hrs) redistribution from peripheral

      CL = 8.4         ; (L/hrs)
      V2 = 9.6         ; (L)
      K20 = CL/V2      ; (/hrs) elimination from plasma

      ; COVARIATES
      IF(NEWIND.LE.1) THEN
            BMI = WEIGHT/((HEIGHT/100)**2)         ; BMI
      ENDIF

      IF(NEWIND.LE.1.AND.SEX.EQ.0) THEN            ; LBW females
            LBW = (9270*WEIGHT)/(8780+244*BMI)
      ENDIF

      IF(NEWIND.LE.1.AND.SEX.EQ.1) THEN            ; LBW males
            LBW = (9270*WEIGHT)/(6680+216*BMI)
      ENDIF

      ; RBC PK PARAMETERS FOR PARENT MTXGLU1
      TVK24 = THETA(1)                             ; kin
      TVCL4 = THETA(2)                             ; CLGlu1
      TVV4 = THETA(3)*LBW/45*(MCV/90)*(HB/130)     ; VGlu1

      ; RUV PARAMETERS
      CV21 = THETA(4)        ; CV2prop MTXGlu1
      ADV1 = THETA(5)        ; ADV MTXGlu1

      ; BSV ON RBC PK PARAMETERS FOR PARENT MTXGLU1
      K24 = TVK24*EXP(ETA(1))
      CL4 = TVCL4
      V4 = TVV4*EXP(ETA(2))
      K40 = CL4/V4

$ERROR
      ; CONCENTRATIONS
      CP = A(2)/V2           ; MTXGlu1 plasma concentration
      CGLU1 = A(4)/V4        ; MTXGlu1 concentration in RBCs

      ; RUV
      SD1 = SQRT(CV21*CGLU1**2+ADV1)   ; combined error MTXGlu1

      ; PARENT MTXGLU1

            F_FLAG=0                   ; observation
            Y=CGLU1+SD1*ERR(1)
            IPRED=CGLU1
            IRES=DV-CGLU1
            IWRES=IRES/SD1


      IF(DVID.EQ.1.AND.BLQ.EQ.1) THEN

            F_FLAG=1                   ; BLQ data
            DUM1 = (LOQ-CGLU1)/SD1
```

```
            CUMD1 = PHI(DUM1)

            Y=CUMD1
            IRES=Y-CGLU1
            IWRES=IRES/SD1

      ENDIF


$THETA   ; Fixed effects
      (0, 1.35E-04)    ; TVK24
      (0, 7.05E-04)    ; TVCL4
      (0, 0.287)       ; TVV4
      (0, 0.0412)      ; CV21 MTXGLU1
      (0, 12.8)        ; ADV1 MTXGLU1


$OMEGA   ; Between subject variability
      0.405      ; BSVK24
      1.210      ; BSVV4


$SIGMA  ; Residual unexplained variability
      1  FIX  ; EPS


$EST  NOABORT MAXEVAL=9990 SIG=3 PRINT=1 METHOD=1 INTERACTION
      LAPLACIAN NUMERICAL SLOW

$COVARIANCE
```

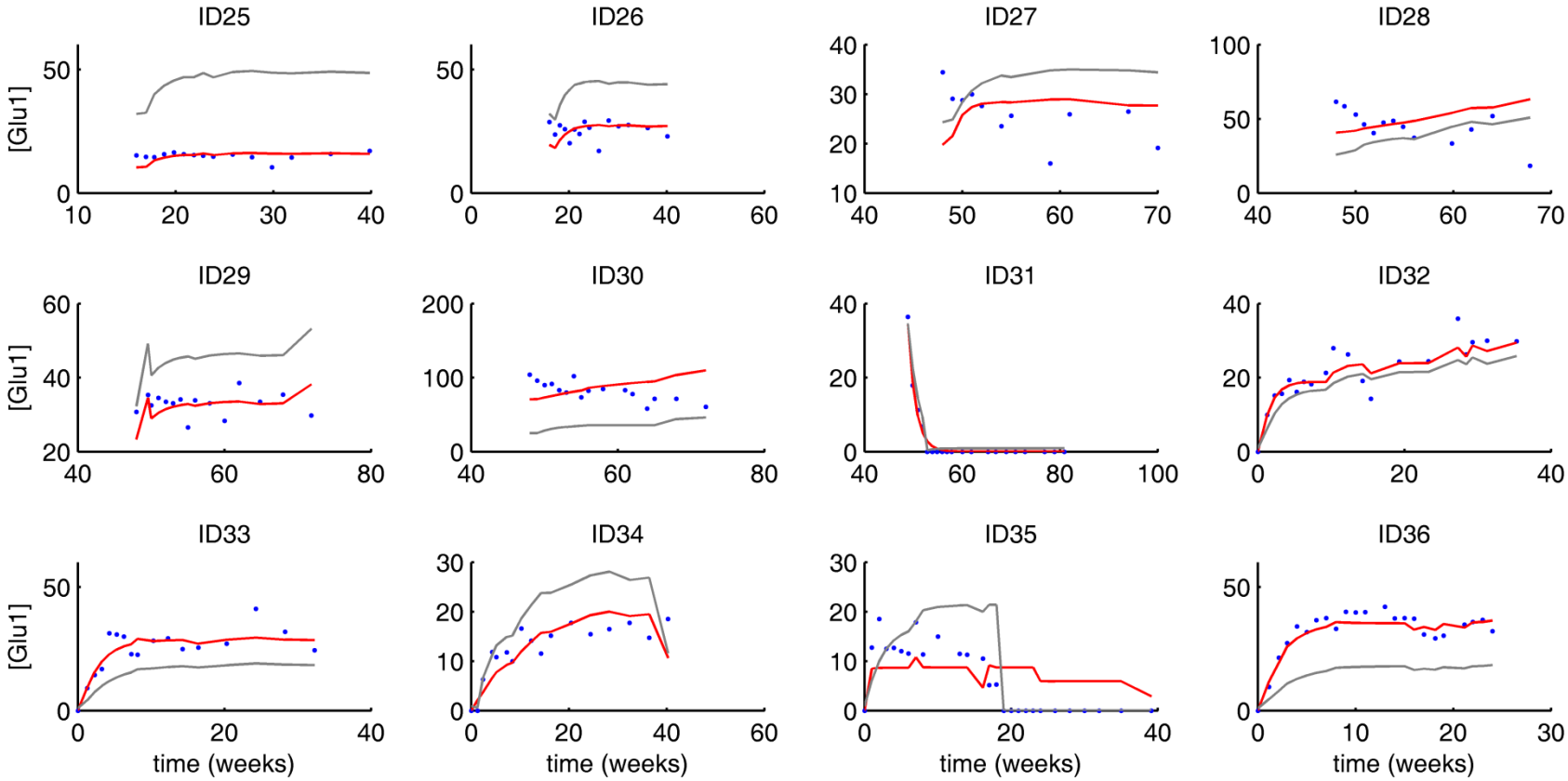
A.5.3. *Individual fits for the final parent model*



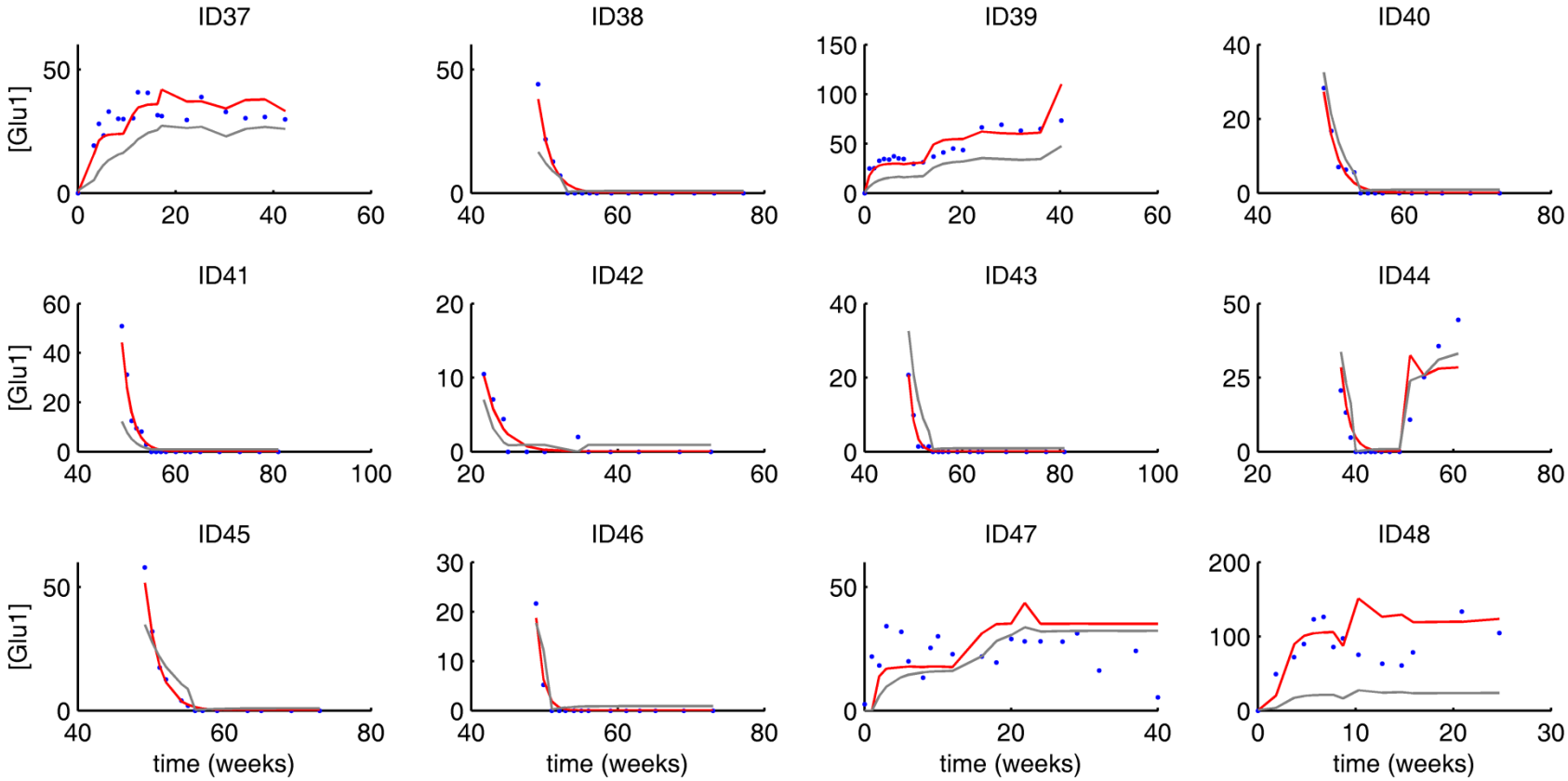
App Fig. 5.3.1: *Individual fits obtained with the final parent model for MTXGlu1: ID1 to ID12.*



App Fig. 5.3.2: *Individual fits obtained with the final parent model for MTXGlu1: ID13 to ID24.*



App Fig. 5.3.3: *Individual fits obtained with the final parent model for MTXGlu1: ID25 to ID36.*



App Fig. 5.3.4: *Individual fits obtained with the final parent model for MTXGlu1: ID37 to ID48.*

A.5.4. *NONMEM® code for the final reduced parent-metabolite model*

```
;RBC MTXPGs Simultaneous parent-metabolite model
;combined data set BASED ON AVERAGE of replicate measurements
;M6 to handle BLQ

;2cmpt PK model for MTX based on article by Hoekstra et al.
;(2004)
;combined error model
;first-order uptake and elimination from RBCs, elimination coded
;as clearance

;4 metabolites = MTXGlu2 - MTXGlu5
;first-order FPGS fixed to values obtained by Morrison & Allegra
;(1987) Model I, no BSV
;first-order gGH estimated with BSV on gGH3-5
;gGH cleaving one terminal glutamate moiety only
;no elimination of metabolite
;same VRBCs for all MTXPGs fixed to parent value, with BSV on
;VGlu1, VGlu2 & VGlu5
;only MCV as covariate on VRBCs
;with BSV on CLGlu1


$PROBLEM    PMGlu15 fixedFPGS&V M6

$INPUT     ID TIME AMT ROUTE II ADDL DVID CMTX OCC DV=AVE MDV
EVID BLQ STUDY STARTSTOP RBC HB MCV HEIGHT WEIGHT AGE SEX

$DATA      MTX_data.csv IGNORE=#
           IGNORE=(OCC.EQ.2,BLQ.EQ.1)

$SUBROUTINE ADVAN5


$MODEL      NCOMP=8
      COMP=(ABS,DEFDOSE)
      COMP=(PLASMA)
      COMP=(PERIPH)
      COMP=(RBCGLU1) ; Parent drug MTXGlu1 in RBCs
      COMP=(RBCGLU2) ; Metabolite MTXGlu2 in RBCs
      COMP=(RBCGLU3) ; Metabolite MTXGlu3 in RBCs
      COMP=(RBCGLU4) ; Metabolite MTXGlu4 in RBCs
      COMP=(RBCGLU5) ; Metabolite MTXGlu5 in RBCs


$PK
      ; COMPARTMENTS
      IF(DVID.EQ.0) THEN
            CMT=1      ; dose => absorption compartment
      ELSEIF(DVID.EQ.1) THEN
            CMT=4      ; observation => parent MTXGlu1 in RBCs
      ELSEIF(DVID.EQ.2) THEN
            CMT=5      ; observation => MTXGlu2 in RBCs
```

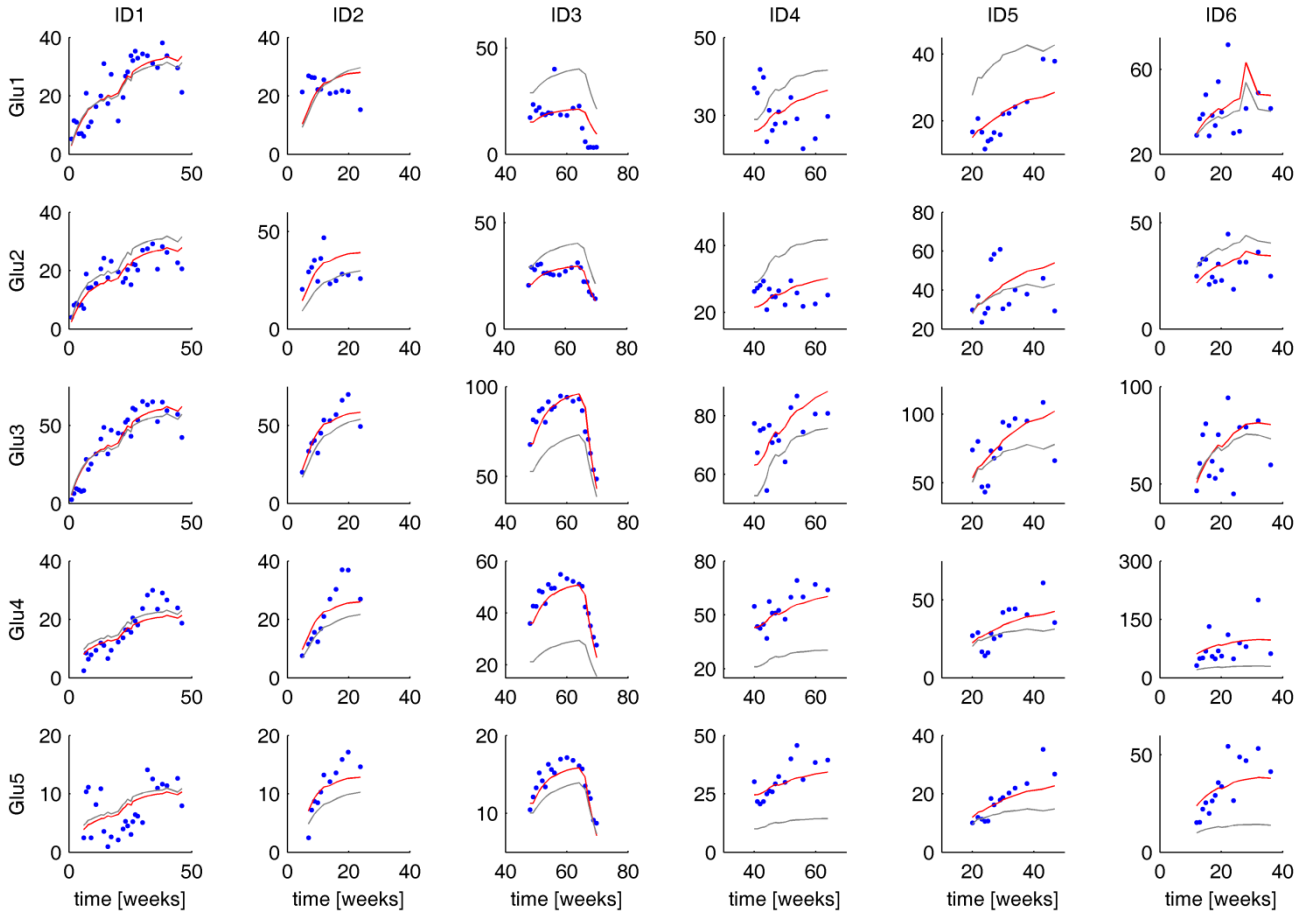
```
ELSEIF(DVID.EQ.3) THEN
      CMT=6       ; observation => MTXGlu3 in RBCs
ELSEIF(DVID.EQ.4) THEN
      CMT=7       ; observation => MTXGlu4 in RBCs
ELSEIF(DVID.EQ.5) THEN
      CMT=8       ; observation => MTXGlu5 in RBCs
ENDIF


; LIMIT OF QUANTIFICATION
LOQ = 5           ; limit of quantification (nmol/LRBCs)


; PLASMA PK FIXED: mean parameter estimates Hoekstra et al

IF(ROUTE.EQ.0) THEN
      F1=0.7                ; oral dosing
      K12 = 0.87            ; (/hrs) oral absorption
      ALAG1 = 0.36          ; lag-time after oral
ELSEIF(ROUTE.EQ.1) THEN
      F1=1                  ; sc dosing
      K12 = 0.36            ; (/hrs) sc absorption
      ALAG1 = 0.06          ; lag-time after sc
ENDIF

K23 = 0.81        ; (/hrs) distribution into peripheral
K32 = 0.55        ; (/hrs) redistribution from peripheral
CL = 8.4          ; (L/hrs)
V2 = 9.6          ; (L)
K20 = CL/V2       ; (/hrs) elimination from plasma


; COVARIATES
IF(NEWIND.LE.1) THEN
      BMI = WEIGHT/((HEIGHT/100)**2)        ; BMI
ENDIF

IF(NEWIND.LE.1.AND.SEX.EQ.0) THEN           ; LBW females
      LBW = (9270*WEIGHT)/(8780+244*BMI)
ENDIF

IF(NEWIND.LE.1.AND.SEX.EQ.1) THEN           ; LBW males
      LBW = (9270*WEIGHT)/(6680+216*BMI)
ENDIF


; RBC PK PARAMETERS FOR PARENT MTXGLU1
TVK24 = THETA(1)            ; kin
TVCL4 = THETA(2)            ; CLGlu1
TVV4 = THETA(3)*(MCV/90)    ; VGlu1

; RBC PK PARAMETERS FOR METABOLITES MTXPGs
TVV5 = THETA(3)*(MCV/90)    ; VGlu2
TVV6 = THETA(3)*(MCV/90)    ; VGlu3
TVV7 = THETA(3)*(MCV/90)    ; VGlu4
TVV8 = THETA(3)*(MCV/90)    ; VGlu5
```

```
TVK45 = THETA(4)      ; kFPGS Glu1->2
TVK56 = THETA(5)      ; kFPGS Glu2->3
TVK67 = THETA(6)      ; kFPGS Glu3->4
TVK78 = THETA(7)      ; kFPGS Glu4->5

TVK54 = THETA(8)      ; kgGH Glu2->1
TVK65 = THETA(9)      ; kgGH Glu3->2
TVK76 = THETA(10)     ; kgGH Glu4->3
TVK87 = THETA(11)     ; kgGH Glu5->4


; RUV PARAMETERS
CV21 = THETA(12)      ; CV2prop - MTXGlu1
ADV1 = THETA(13)      ; ADV - MTXGlu1
CV22 = THETA(14)      ; CV2prop - MTXGlu2
ADV2 = THETA(15)      ; ADV - MTXGlu2
CV23 = THETA(16)      ; CV2prop - MTXGlu3
ADV3 = THETA(17)      ; ADV - MTXGlu3
CV24 = THETA(18)      ; CV2prop - MTXGlu4
ADV4 = THETA(19)      ; ADV - MTXGlu4
CV25 = THETA(20)      ; CV2prop - MTXGlu5
ADV5 = THETA(21)      ; ADV - MTXGlu5


; BSV ON RBC PK PARAMETERS FOR PARENT MTXGLU1
K24 = TVK24*EXP(ETA(1))
CL4 = TVCL4*EXP(ETA(2))
V4 = TVV4*EXP(ETA(3))
K40 = CL4/V4

; BSV ON RBC PK PARAMETERS FOR METABOLITES MTXPGs
V5 = TVV5*EXP(ETA(4))
V6 = TVV6
V7 = TVV7
V8 = TVV8*EXP(ETA(5))

K45 = TVK45
K56 = TVK56
K67 = TVK67
K78 = TVK78

K54 = TVK54
K65 = TVK65*EXP(ETA(6))
K76 = TVK76*EXP(ETA(7))
K87 = TVK87*EXP(ETA(8))
```

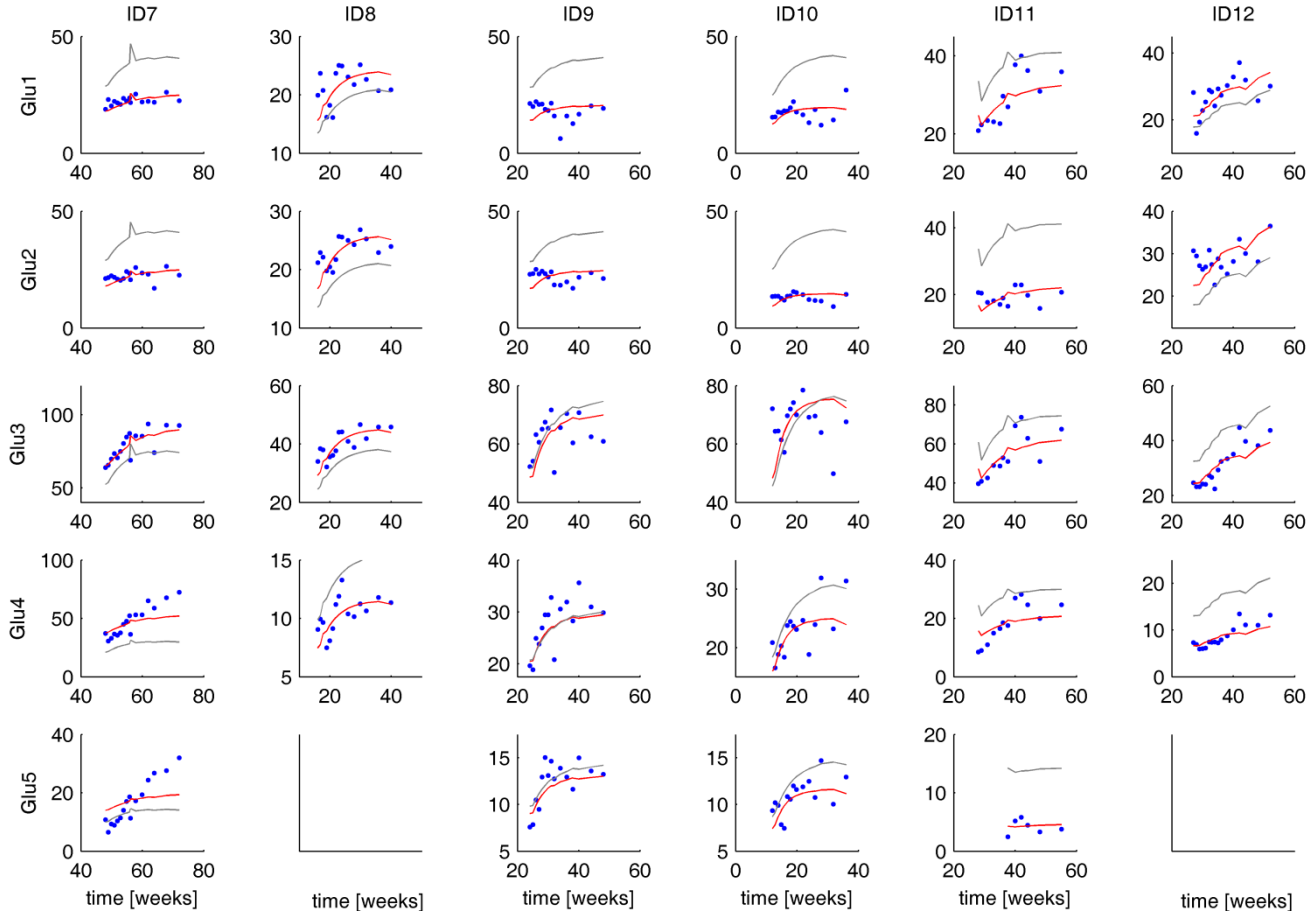
```
$ERROR

      ; CONCENTRATIONS
      CP = A(2)/V2          ; MTXGlu1 plasma concentration
      CGLU1 = A(4)/V4       ; MTXGlu1 concentration in RBCs
      CGLU2 = A(5)/V5       ; MTXGlu2 concentration in RBCs
      CGLU3 = A(6)/V6       ; MTXGlu3 concentration in RBCs
      CGLU4 = A(7)/V7       ; MTXGlu4 concentration in RBCs
      CGLU5 = A(8)/V8       ; MTXGlu5 concentration in RBCs


      ; RUV
      SD1 = SQRT(CV21*CGLU1**2+ADV1)   ; combined error MTXGlu1
      SD2 = SQRT(CV22*CGLU2**2+ADV2)   ; combined error MTXGlu2
      SD3 = SQRT(CV23*CGLU3**2+ADV3)   ; combined error MTXGlu3
      SD4 = SQRT(CV24*CGLU4**2+ADV4)   ; combined error MTXGlu4
      SD5 = SQRT(CV25*CGLU5**2+ADV5)   ; combined error MTXGlu5


      ; PARENT MTXGLU1

            Y=CGLU1+SD1*ERR(1)
            IPRED=CGLU1
            IRES=DV-CGLU1
            IWRES=IRES/SD1

      ; METABOLITE MTXGLU2
      IF(DVID.EQ.2.AND.BLQ.EQ.0) THEN

            Y=CGLU2+SD2*ERR(1)
            IPRED=CGLU2
            IRES=DV-CGLU2
            IWRES=IRES/SD2

      ENDIF

      ; METABOLITE MTXGLU3
      IF(DVID.EQ.3.AND.BLQ.EQ.0) THEN

            Y=CGLU3+SD3*ERR(1)
            IPRED=CGLU3
            IRES=DV-CGLU3
            IWRES=IRES/SD3

      ENDIF

      ; METABOLITE MTXGLU4
      IF(DVID.EQ.4.AND.BLQ.EQ.0) THEN

            Y=CGLU4+SD4*ERR(1)
            IPRED=CGLU4
            IRES=DV-CGLU4
            IWRES=IRES/SD4

      ENDIF
```

```
      ; METABOLITE MTXGLU5
      IF(DVID.EQ.5.AND.BLQ.EQ.0) THEN

            Y=CGLU5+SD5*ERR(1)
            IPRED=CGLU5
            IRES=DV-CGLU5
            IWRES=IRES/SD5

      ENDIF


$THETA     ; Fixed effects

      (0, 2.27E-04)    ; TVK24
      (0, 1.48E-03)    ; TVCL4
      0.3 FIX          ; TVV4

      0.171 FIX        ; TVK45 polyglutamation
      0.344 FIX        ; TVK56 polyglutamation
      0.097 FIX        ; TVK67 polyglutamation
      0.141 FIX        ; TVK78 polyglutamation
      (0, 0.174)       ; TVK54 deglutamation
      (0, 0.192)       ; TVK65 deglutamation
      (0, 0.243)       ; TVK76 deglutamation
      (0, 0.299)       ; TVK87 deglutamation

      (0, 0.0334)      ; CV21 MTXGLU1
      (0, 35.5)        ; ADV1 MTXGLU1
      (0, 0.0460)      ; CV22 MTXGLU2
      (0, 7.22)        ; ADV2 MTXGLU2
      (0, 0.0139)      ; CV23 MTXGLU3
      (0, 39.6)        ; ADV3 MTXGLU3
      (0, 0.0644)      ; CV24 MTXGLU4
      (0, 3.82)        ; ADV4 MTXGLU4
      (0, 0.0779)      ; CV25 MTXGLU5
      (0, 2.25)        ; ADV5 MTXGLU5

$OMEGA     ; Between subject variability
       0.482           ; BSVK24
       0.401           ; BSVCL4
       0.101           ; BSVV4
       0.104           ; BSVV5
       0.233           ; BSVV8
       0.284           ; BSVK65
       0.174           ; BSVK76
       0.059           ; BSVK87

$SIGMA  ; Residual unexplained variability
      1  FIX           ; EPS


$EST  NOABORT MAXEVAL=9990 SIG=5 PRINT=5 METHOD=COND
      INTERACTION
```

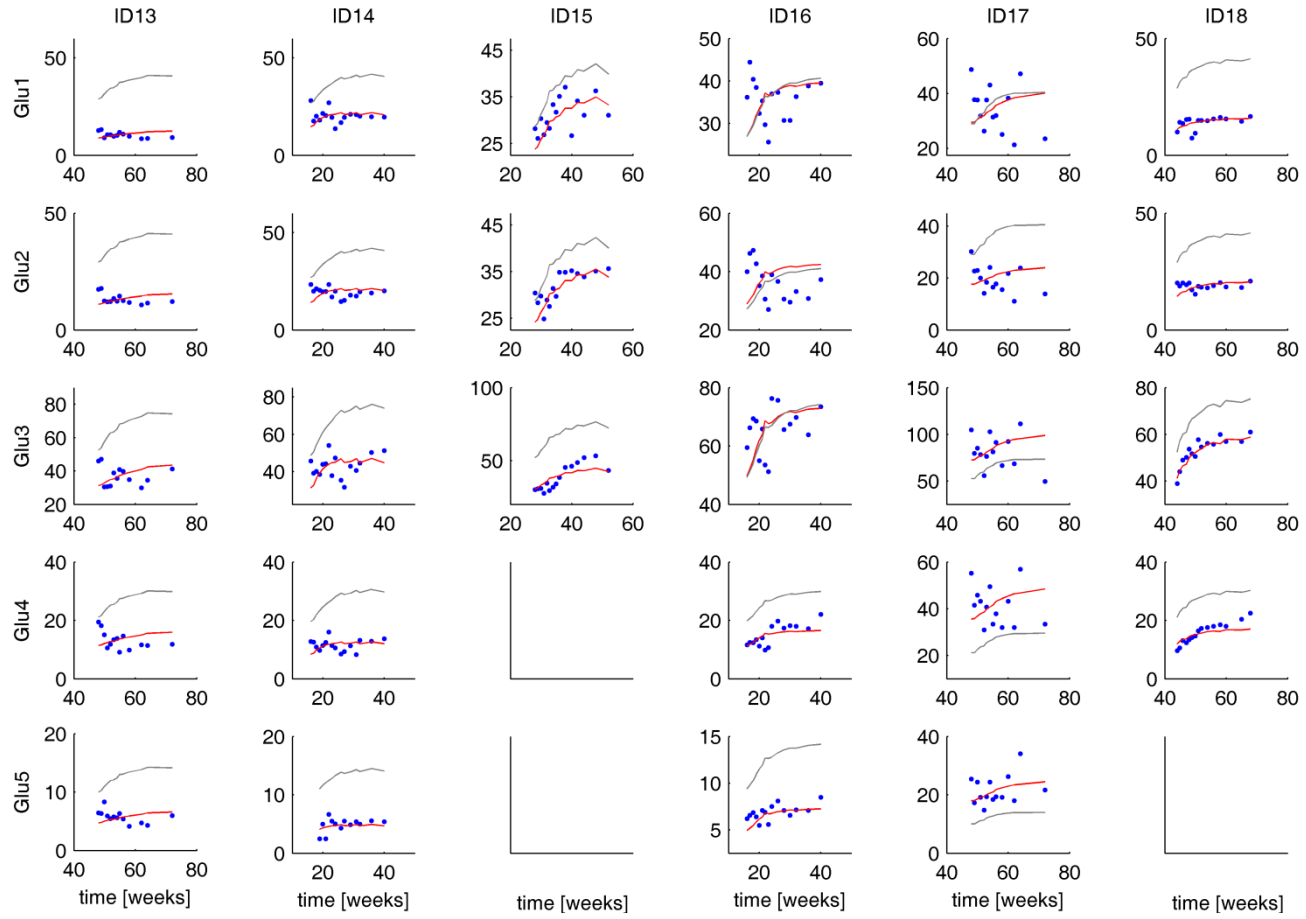
A.5.5. *Individual fits for the final reduced parent-metabolite model*



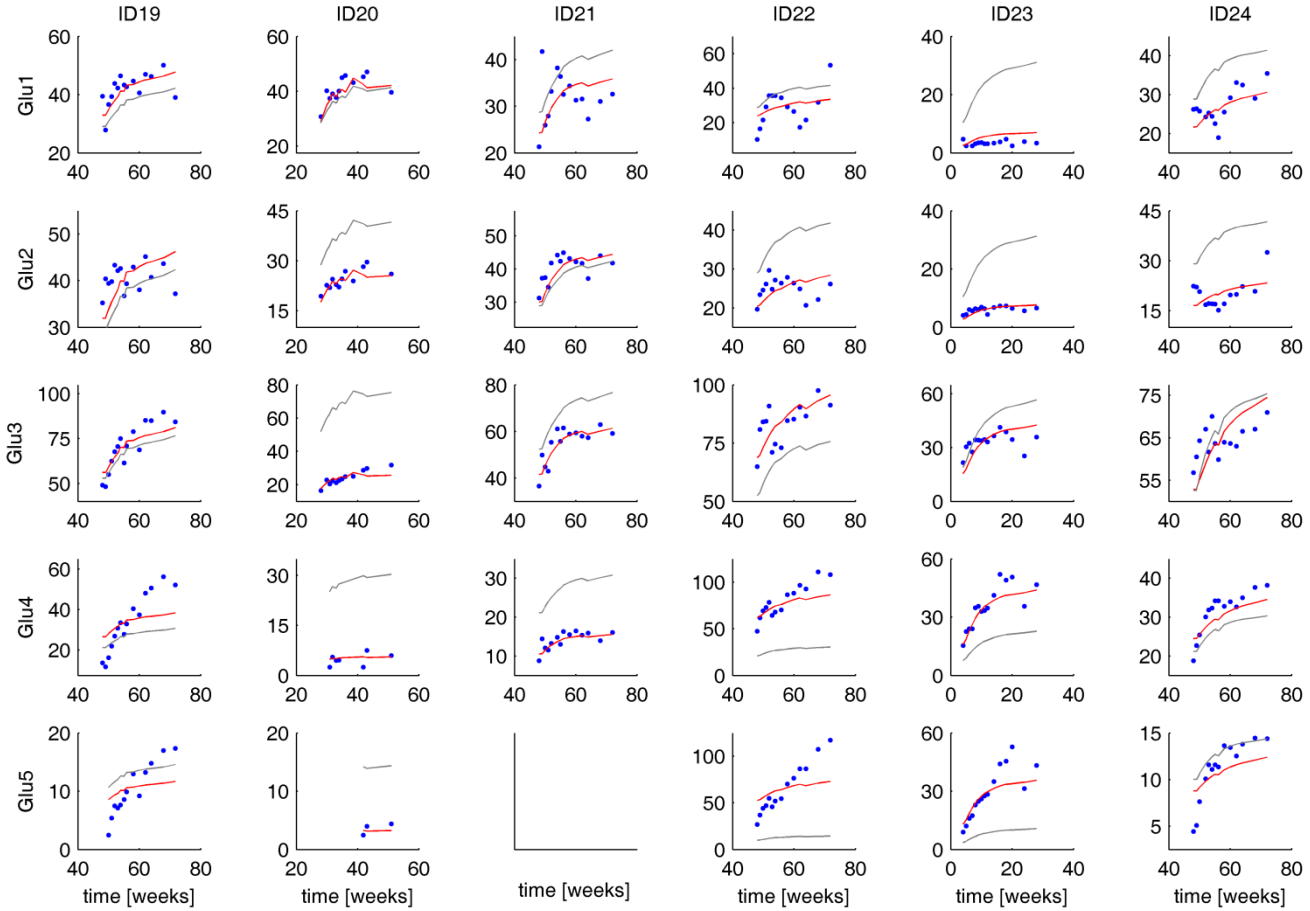
App Fig. 5.5.1: *Individual fits obtained with the final reduced PM model for all MTXGluX: ID1 to ID6.*



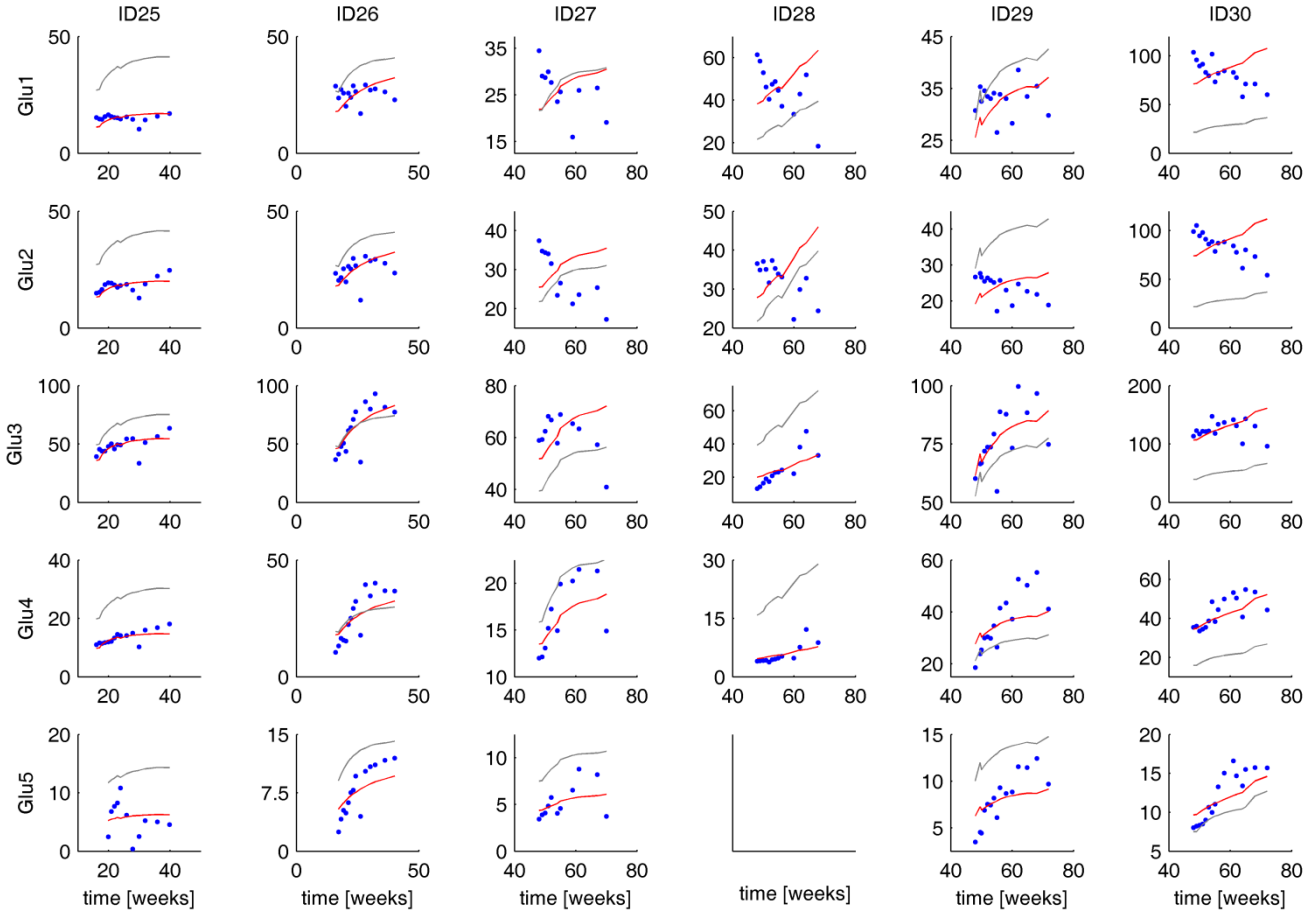
App Fig. 5.5.2: *Individual fits obtained with the final reduced PM model for all MTXGluX: ID7to ID12.*



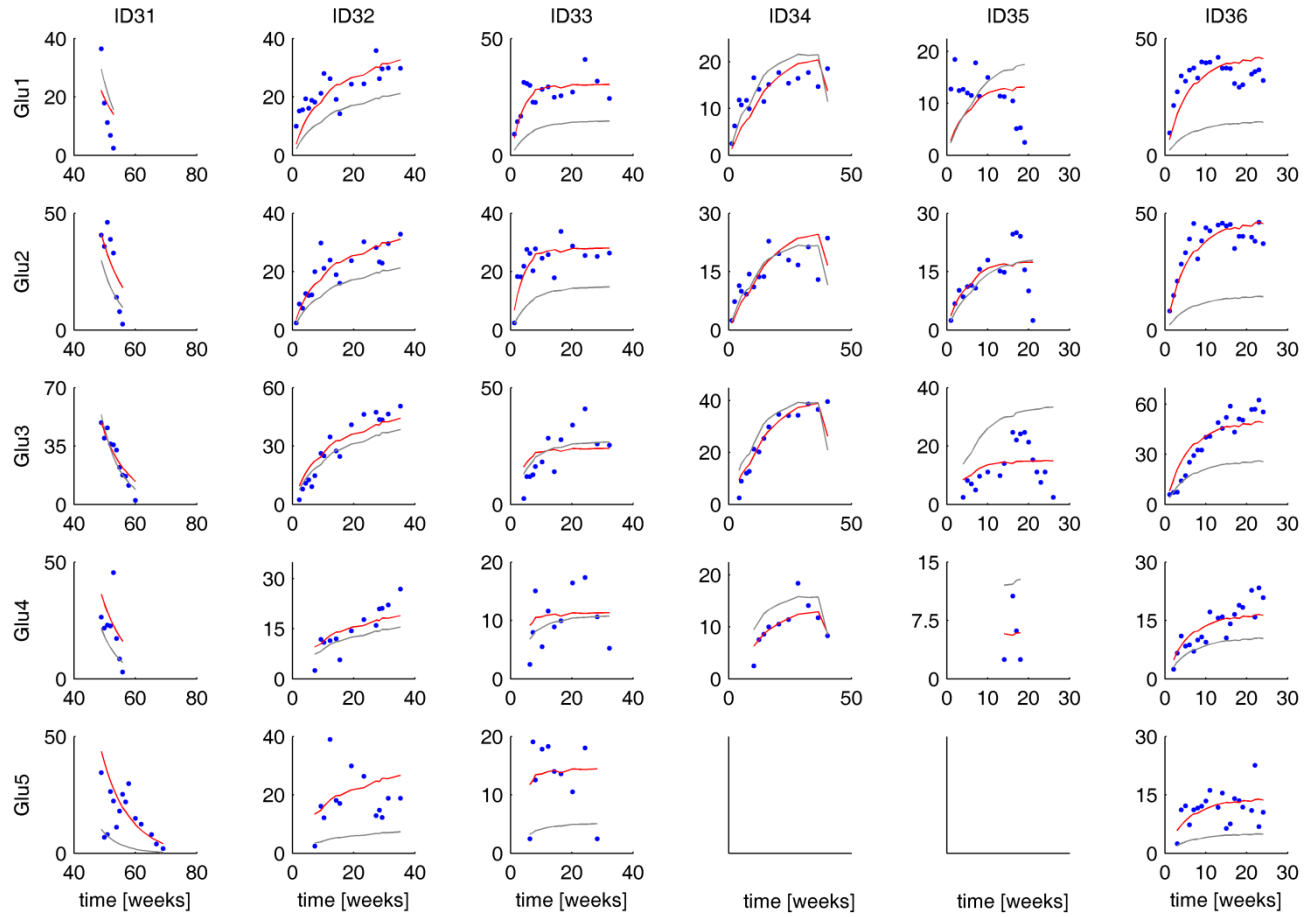
App Fig. 5.5.3: *Individual fits obtained with the final reduced PM model for all MTXGluX: ID13 to ID18.*



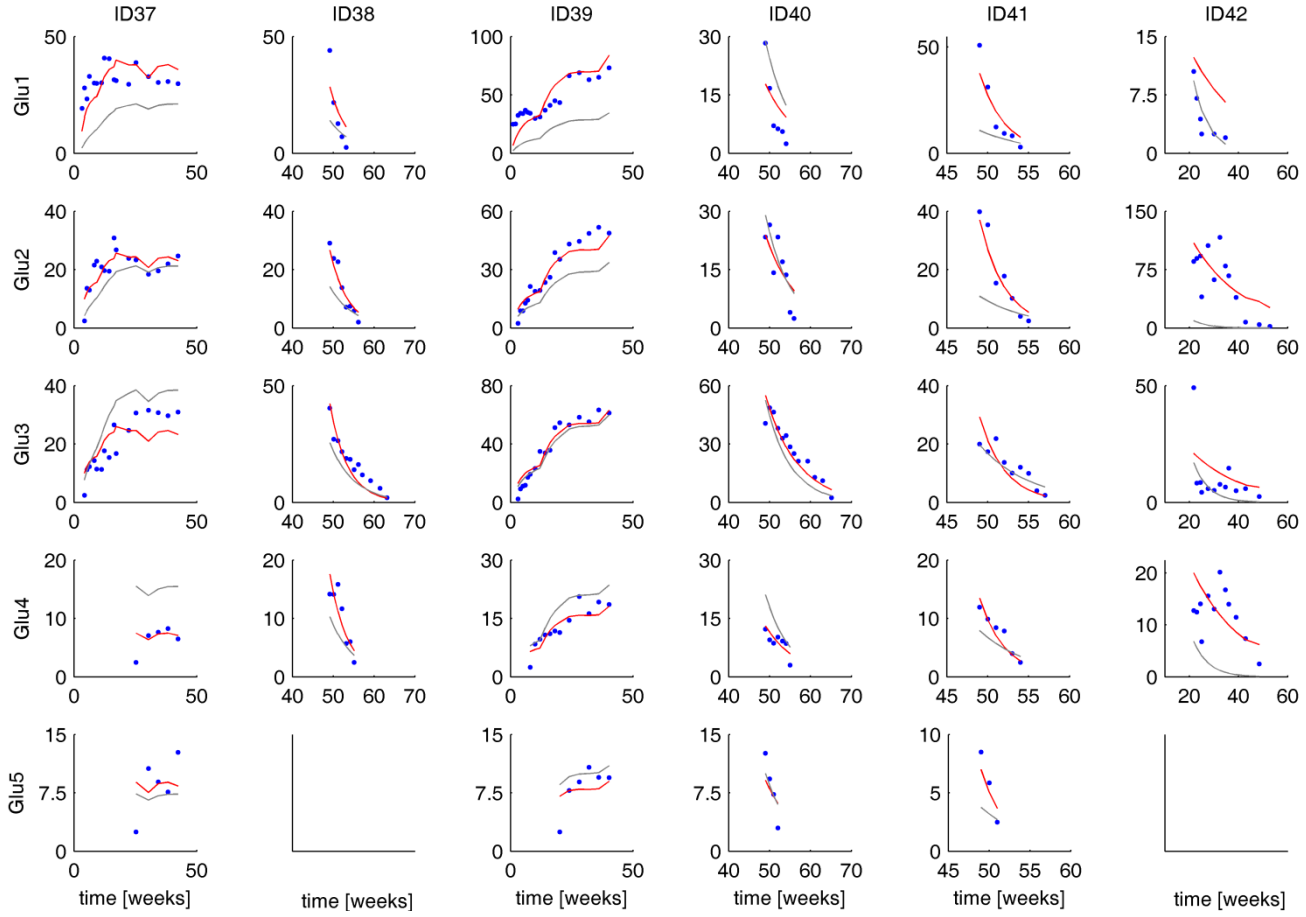
App Fig. 5.5.4: *Individual fits obtained with the final reduced PM model for all MTXGluX: ID19 to ID24.*



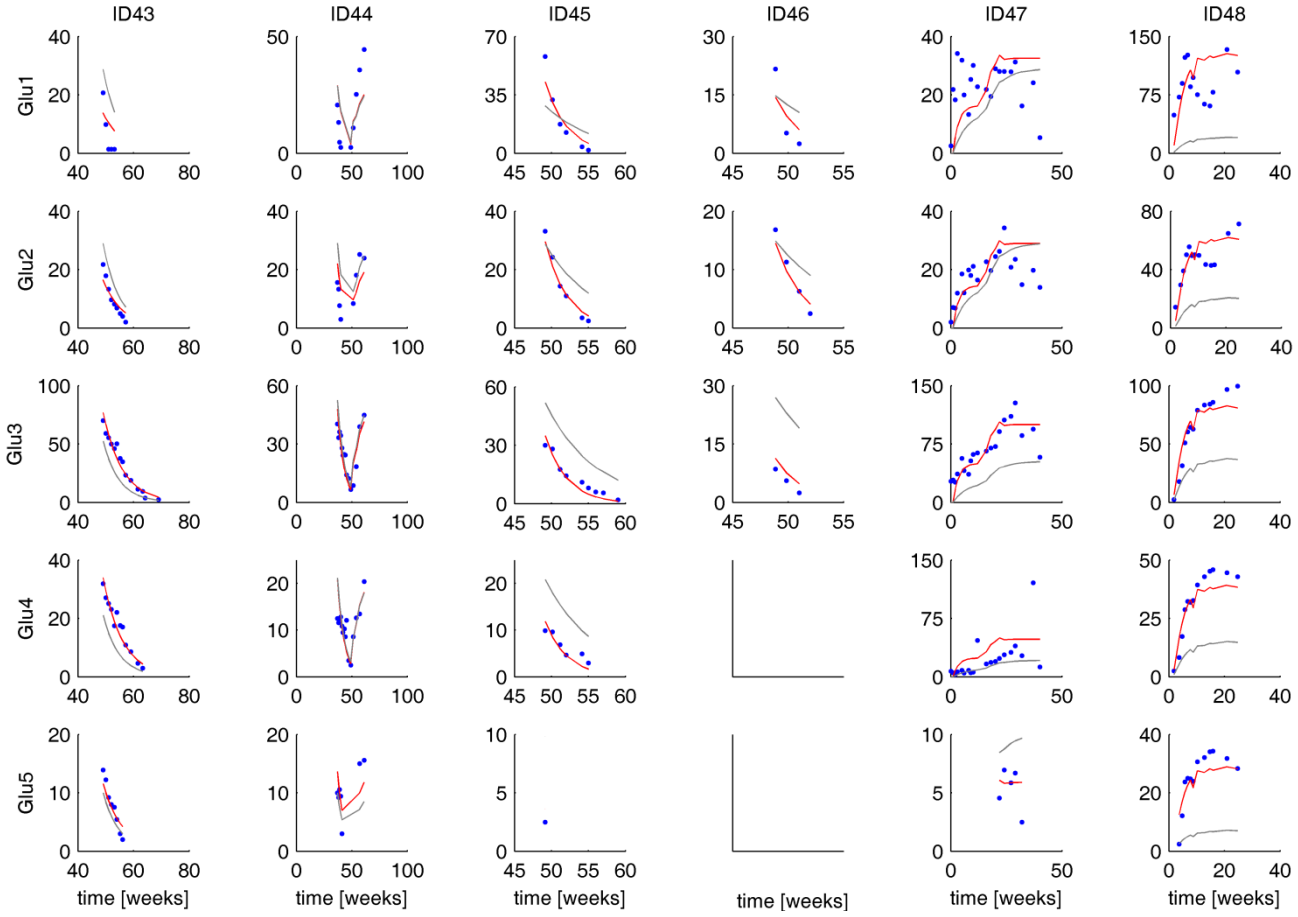
App Fig. 5.5.5: *Individual fits obtained with the final reduced PM model for all MTXGluX: ID25 to ID30.*



App Fig. 5.5.6: *Individual fits obtained with the final reduced PM model for all MTXGluX: ID31 to ID36.*



App Fig. 5.5.7: *Individual fits obtained with the final reduced PM model for all MTXGluX: ID37 to ID42.*



App Fig. 5.5.8: *Individual fits obtained with the final reduced PM model for all MTXGluX: ID43 to ID48.*

Appendix 6: Appendix to Chapter 6

A.6.1. *NONMEM® code for the final updated parent-metabolite model*

```
;RBC MTXPGs Simultaneous parent-metabolite model
;combined data set BASED ON AVERAGE of replicate measurements
;M6 to handle BLQ

;2cmpt PK model for MTX based on article by Hoekstra et al.
;(2004)
;combined error model
;first-order uptake and elimination from RBCs, elimination coded
;as clearance

;4 metabolites = MTXGlu2 - MTXGlu5
;first-order FPGS fixed to values obtained by Morrison & Allegra
;(1987) Model I, no BSV
;first-order gGH estimated with BSV on gGH3-5
;gGH cleaving one terminal glutamate moiety only
;elimination of all metabolites equal to elimination of parent
;drug
;same VRBCs for all MTXPGs fixed to parent value, with BSV on
;VGlu1, VGlu2 & VGlu5
;only MCV as covariate on VRBCs
;with BSV on CLGlu1

;eGFR and BSA as covariates on renal plasma clearance,
;allometrically scaled WT on non-renal plasma clearance


$PROBLEM    PMGlu15 fixed Hyp18h

$INPUT     ID TIME AMT ROUTE II ADDL DVID CMTX OCC DV=AVE MDV
EVID BLQ STUDY STARTSTOP RBC HB MCV HEIGHT WEIGHT AGE SEX
ABCB1rs10   SLC19A1rs10   ABCC2rs22   ABCC1rs35A1   ABCC1rs35A2
ABCC1rs37A1   ABCC1rs37A2   ABCC2rs41A1   ABCC2rs41A2   GGHrs11A1
GGHrs11A2 ABCG2rs17A1 ABCG2rs17A2


$DATA      MTX_data_genotypes.csv IGNORE=#
           IGNORE=(OCC.EQ.2,BLQ.EQ.1)


$SUBROUTINE ADVAN5

$MODEL      NCOMP=8
      COMP=(ABS,DEFDOSE)
      COMP=(PLASMA)
      COMP=(PERIPH)
      COMP=(RBCGLU1) ; Parent drug MTXGlu1 in RBCs
      COMP=(RBCGLU2) ; Metabolite MTXGlu2 in RBCs
      COMP=(RBCGLU3) ; Metabolite MTXGlu3 in RBCs
      COMP=(RBCGLU4) ; Metabolite MTXGlu4 in RBCs
      COMP=(RBCGLU5) ; Metabolite MTXGlu5 in RBCs
```

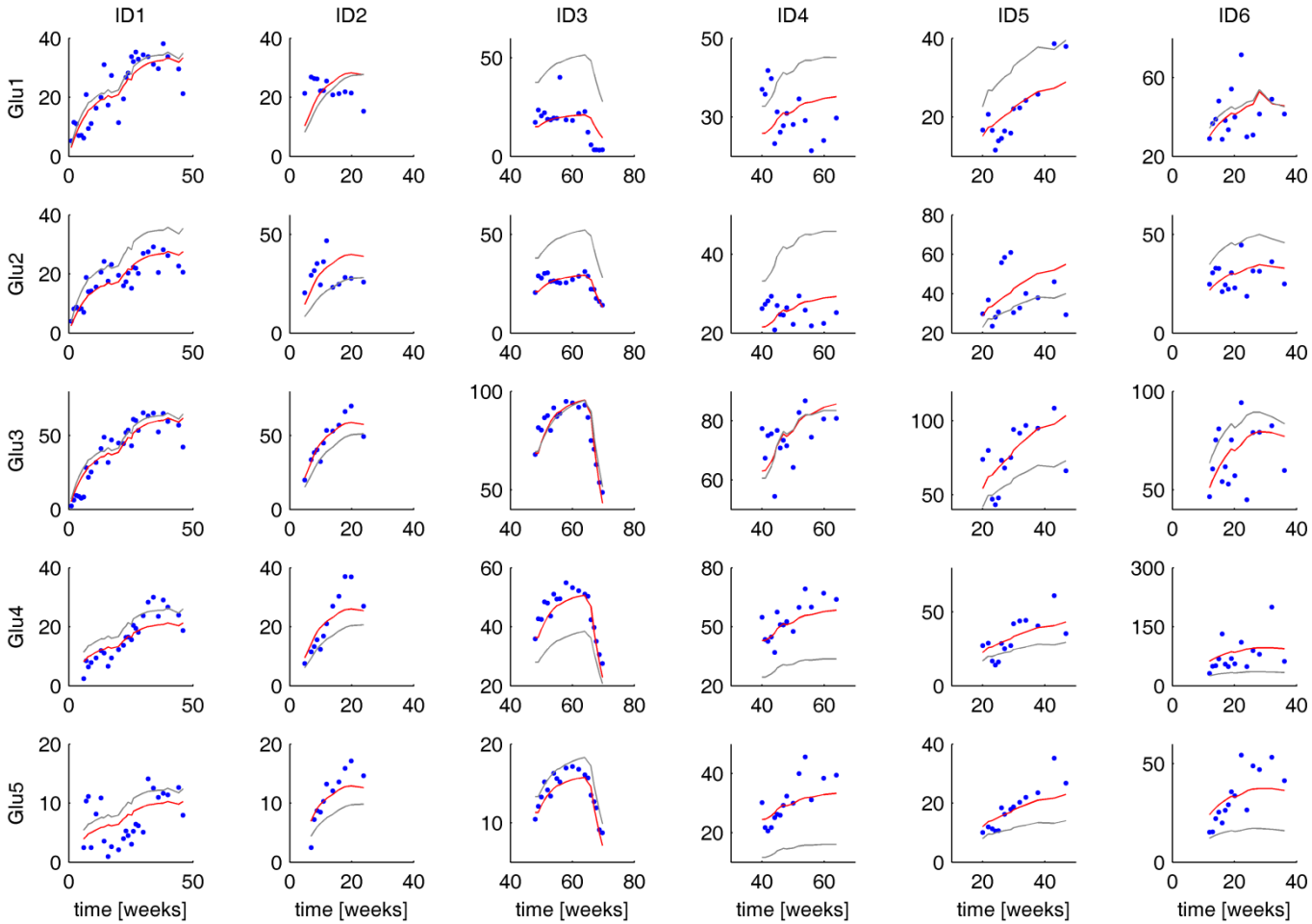
```
$PK
      ; COMPARTMENTS
      IF(DVID.EQ.0) THEN
            CMT=1       ; dose => absorption compartment
      ELSEIF(DVID.EQ.1) THEN
            CMT=4       ; observation => parent MTXGlu1 in RBCs
      ELSEIF(DVID.EQ.2) THEN
            CMT=5       ; observation => MTXGlu2 in RBCs
      ELSEIF(DVID.EQ.3) THEN
            CMT=6       ; observation => MTXGlu3 in RBCs
      ELSEIF(DVID.EQ.4) THEN
            CMT=7       ; observation => MTXGlu4 in RBCs
      ELSEIF(DVID.EQ.5) THEN
            CMT=8       ; observation => MTXGlu5 in RBCs
      ENDIF

      ; LIMIT OF QUANTIFICATION
      LOQ = 5           ; limit of quantification (nmol/LRBCs)

      ; COVARIATES
      IF(NEWIND.LE.1) THEN
            BMI = WEIGHT/((HEIGHT/100)**2)          ; BMI
            BSA = SQRT((WEIGHT*HEIGHT)/3600)        ; BSA (m^2)
      ENDIF

      IF(NEWIND.LE.1.AND.SEX.EQ.0) THEN             ; LBW females
            LBW = (9270*WEIGHT)/(8780+244*BMI)
      ENDIF

      IF(NEWIND.LE.1.AND.SEX.EQ.1) THEN             ; LBW males
            LBW = (9270*WEIGHT)/(6680+216*BMI)
      ENDIF

      FE = 0.81                   ; fraction renal elimination

      ; PLASMA PK FIXED: mean parameter estimates Hoekstra et al

      IF(ROUTE.EQ.0) THEN
            F1=0.7                ; oral dosing
            K12 = 0.87            ; (/hrs) oral absorption
            ALAG1 = 0.36          ; lag-time after oral
      ELSEIF(ROUTE.EQ.1) THEN
            F1=1                  ; sc dosing
            K12 = 0.36            ; (/hrs) sc absorption
            ALAG1 = 0.06          ; lag-time after sc
      ENDIF

      K23 = 0.81        ; (/hrs) distribution into peripheral
      K32 = 0.55        ; (/hrs) redistribution from peripheral

      CLRENAL = FE*8.4*(EGFR/100)*(BSA/1.73)*THETA(22)
      CLNONRENAL = (1-FE)*8.4*(WEIGHT/75)**0.75
      CL = CLRENAL + CLNONRENAL         ; total clearance (L/hrs)
      V2 = 9.6          ; (L)
      K20 = CL/V2       ; (/hrs) elimination from plasma
```

```
; RBC PK PARAMETERS FOR PARENT MTXGLU1
TVK24 = THETA(1)            ; kin
TVCL4 = THETA(2)            ; CLGlu1
TVV4 = THETA(3)*(MCV/90)    ; VGlu1

; RBC PK PARAMETERS FOR METABOLITES MTXPGs
TVV5 = THETA(3)*(MCV/90)    ; VGlu2
TVV6 = THETA(3)*(MCV/90)    ; VGlu3
TVV7 = THETA(3)*(MCV/90)    ; VGlu4
TVV8 = THETA(3)*(MCV/90)    ; VGlu5

TVK45 = THETA(4)      ; kFPGS Glu1->2
TVK56 = THETA(5)      ; kFPGS Glu2->3
TVK67 = THETA(6)      ; kFPGS Glu3->4
TVK78 = THETA(7)      ; kFPGS Glu4->5

TVK54 = THETA(8)      ; kgGH Glu2->1
TVK65 = THETA(9)      ; kgGH Glu3->2
TVK76 = THETA(10)     ; kgGH Glu4->3
TVK87 = THETA(11)     ; kgGH Glu5->4

; RUV PARAMETERS
CV21 = THETA(12)      ; CV2prop - MTXGlu1
ADV1 = THETA(13)      ; ADV - MTXGlu1
CV22 = THETA(14)      ; CV2prop - MTXGlu2
ADV2 = THETA(15)      ; ADV - MTXGlu2
CV23 = THETA(16)      ; CV2prop - MTXGlu3
ADV3 = THETA(17)      ; ADV - MTXGlu3
CV24 = THETA(18)      ; CV2prop - MTXGlu4
ADV4 = THETA(19)      ; ADV - MTXGlu4
CV25 = THETA(20)      ; CV2prop - MTXGlu5
ADV5 = THETA(21)      ; ADV - MTXGlu5

; BSV ON RBC PK PARAMETERS FOR PARENT MTXGLU1
K24 = TVK24*EXP(ETA(1))
CL4 = TVCL4*EXP(ETA(2))
V4 = TVV4*EXP(ETA(3))
K40 = CL4/V4

; BSV ON RBC PK PARAMETERS FOR METABOLITES MTXPGs
V5 = TVV5*EXP(ETA(4))
V6 = TVV6
V7 = TVV7
V8 = TVV8*EXP(ETA(5))

K45 = TVK45
K56 = TVK56
K67 = TVK67
K78 = TVK78

K54 = TVK54
K65 = TVK65*EXP(ETA(6))
K76 = TVK76*EXP(ETA(7))
K87 = TVK87*EXP(ETA(8))
```

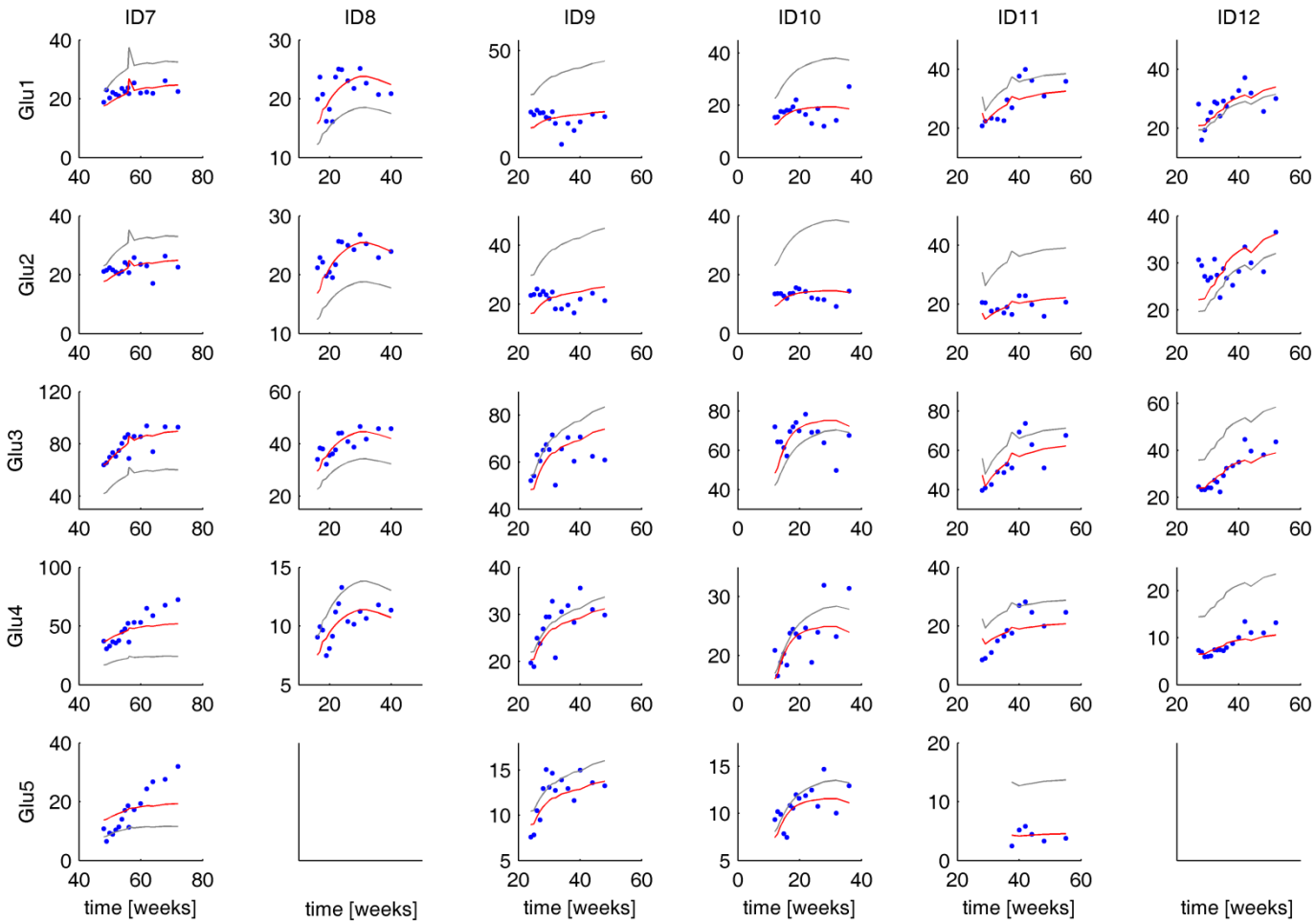
```
$ERROR

      ; CONCENTRATIONS
      CP = A(2)/V2          ; MTXGlu1 plasma concentration
      CGLU1 = A(4)/V4       ; MTXGlu1 concentration in RBCs
      CGLU2 = A(5)/V5       ; MTXGlu2 concentration in RBCs
      CGLU3 = A(6)/V6       ; MTXGlu3 concentration in RBCs
      CGLU4 = A(7)/V7       ; MTXGlu4 concentration in RBCs
      CGLU5 = A(8)/V8       ; MTXGlu5 concentration in RBCs

      ; RUV
      SD1 = SQRT(CV21*CGLU1**2+ADV1)   ; combined error MTXGlu1
      SD2 = SQRT(CV22*CGLU2**2+ADV2)   ; combined error MTXGlu2
      SD3 = SQRT(CV23*CGLU3**2+ADV3)   ; combined error MTXGlu3
      SD4 = SQRT(CV24*CGLU4**2+ADV4)   ; combined error MTXGlu4
      SD5 = SQRT(CV25*CGLU5**2+ADV5)   ; combined error MTXGlu5

      ; PARENT MTXGLU1

            Y=CGLU1+SD1*ERR(1)
            IPRED=CGLU1
            IRES=DV-CGLU1
            IWRES=IRES/SD1

      ; METABOLITE MTXGLU2
      IF(DVID.EQ.2.AND.BLQ.EQ.0) THEN

            Y=CGLU2+SD2*ERR(1)
            IPRED=CGLU2
            IRES=DV-CGLU2
            IWRES=IRES/SD2

      ENDIF

      ; METABOLITE MTXGLU3
      IF(DVID.EQ.3.AND.BLQ.EQ.0) THEN

            Y=CGLU3+SD3*ERR(1)
            IPRED=CGLU3
            IRES=DV-CGLU3
            IWRES=IRES/SD3

      ENDIF

      ; METABOLITE MTXGLU4
      IF(DVID.EQ.4.AND.BLQ.EQ.0) THEN

            Y=CGLU4+SD4*ERR(1)
            IPRED=CGLU4
            IRES=DV-CGLU4
            IWRES=IRES/SD4

      ENDIF

      ; METABOLITE MTXGLU5
      IF(DVID.EQ.5.AND.BLQ.EQ.0) THEN
```

```
            Y=CGLU5+SD5*ERR(1)
            IPRED=CGLU5
            IRES=DV-CGLU5
            IWRES=IRES/SD5

      ENDIF

$THETA     ; Fixed effects

      (0, 5.91E-05)    ; TVK24
      (0, 2.94E-04)    ; TVCL4
      0.3 FIX          ; TVV4

      0.171 FIX        ; TVK45 polyglutamation
      0.344 FIX        ; TVK56 polyglutamation
      0.097 FIX        ; TVK67 polyglutamation
      0.141 FIX        ; TVK78 polyglutamation
      (0, 0.169)       ; TVK54 deglutamation
      (0, 0.189)       ; TVK65 deglutamation
      (0, 0.241)       ; TVK76 deglutamation
      (0, 0.296)       ; TVK87 deglutamation

      (0, 0.0322)      ; CV21 MTXGLU1
      (0, 36.7)        ; ADV1 MTXGLU1
      (0, 0.0460)      ; CV22 MTXGLU2
      (0, 7.53)        ; ADV2 MTXGLU2
      (0, 0.0142)      ; CV23 MTXGLU3
      (0, 38.8)        ; ADV3 MTXGLU3
      (0, 0.0643)      ; CV24 MTXGLU4
      (0, 3.73)        ; ADV4 MTXGLU4
      (0, 0.0787)      ; CV25 MTXGLU5
      (0, 2.16)        ; ADV5 MTXGLU5

      0.106            ; covariate effect eGFR & BSA

$OMEGA     ; Between subject variability
       0.320           ; BSVK24
       0.169           ; BSVCL4
       0.104           ; BSVV4
       0.119           ; BSVV5
       0.268           ; BSVV8
       0.295           ; BSVK65
       0.177           ; BSVK76
       0.0238          ; BSVK87

$SIGMA  ; Residual unexplained variability
      1  FIX           ; EPS

$EST  NOABORT MAXEVAL=9990 SIG=5 PRINT=5 METHOD=COND
      INTERACTION
```

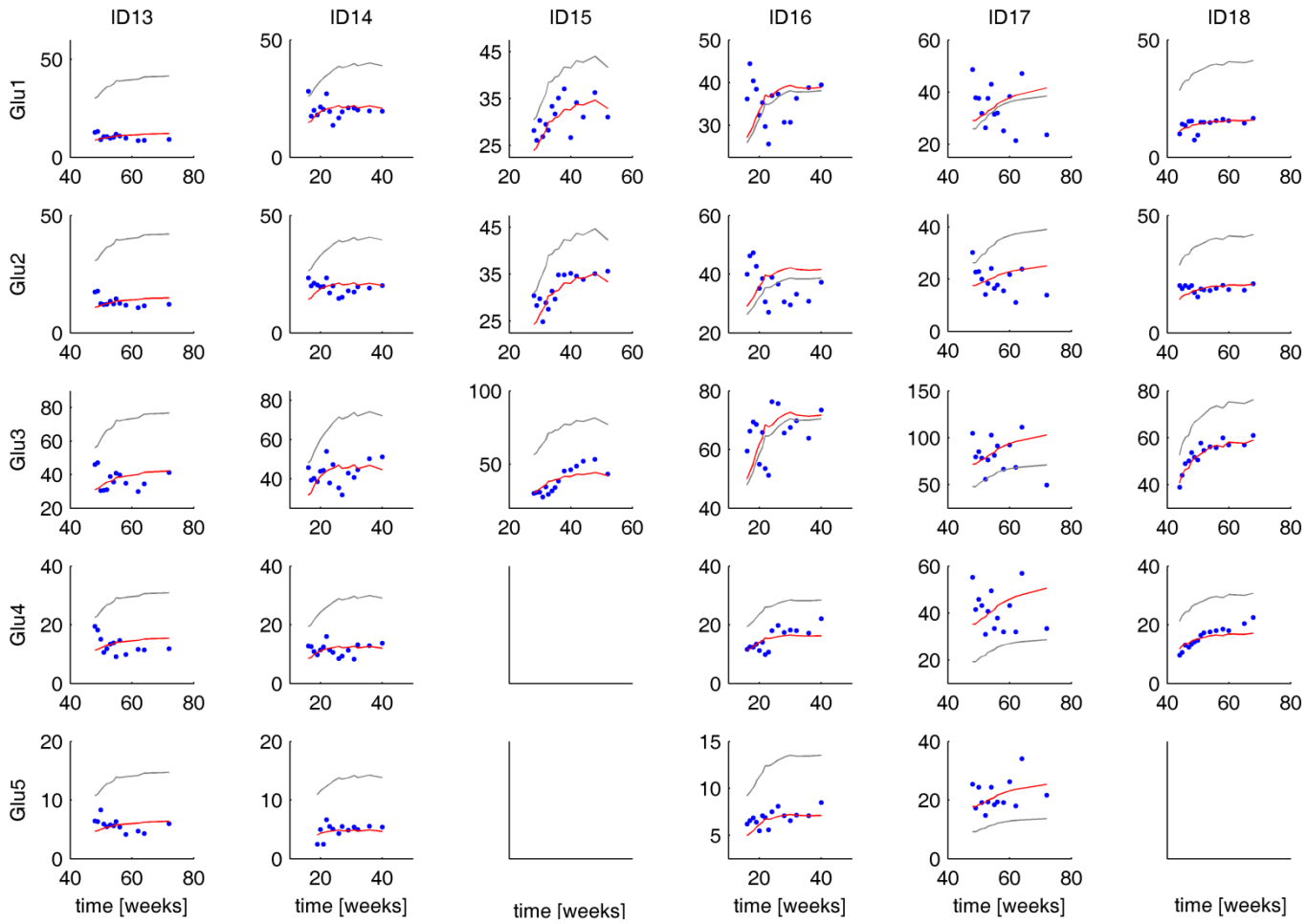
A.6.2. *Individual fits for the final updated parent-metabolite model*



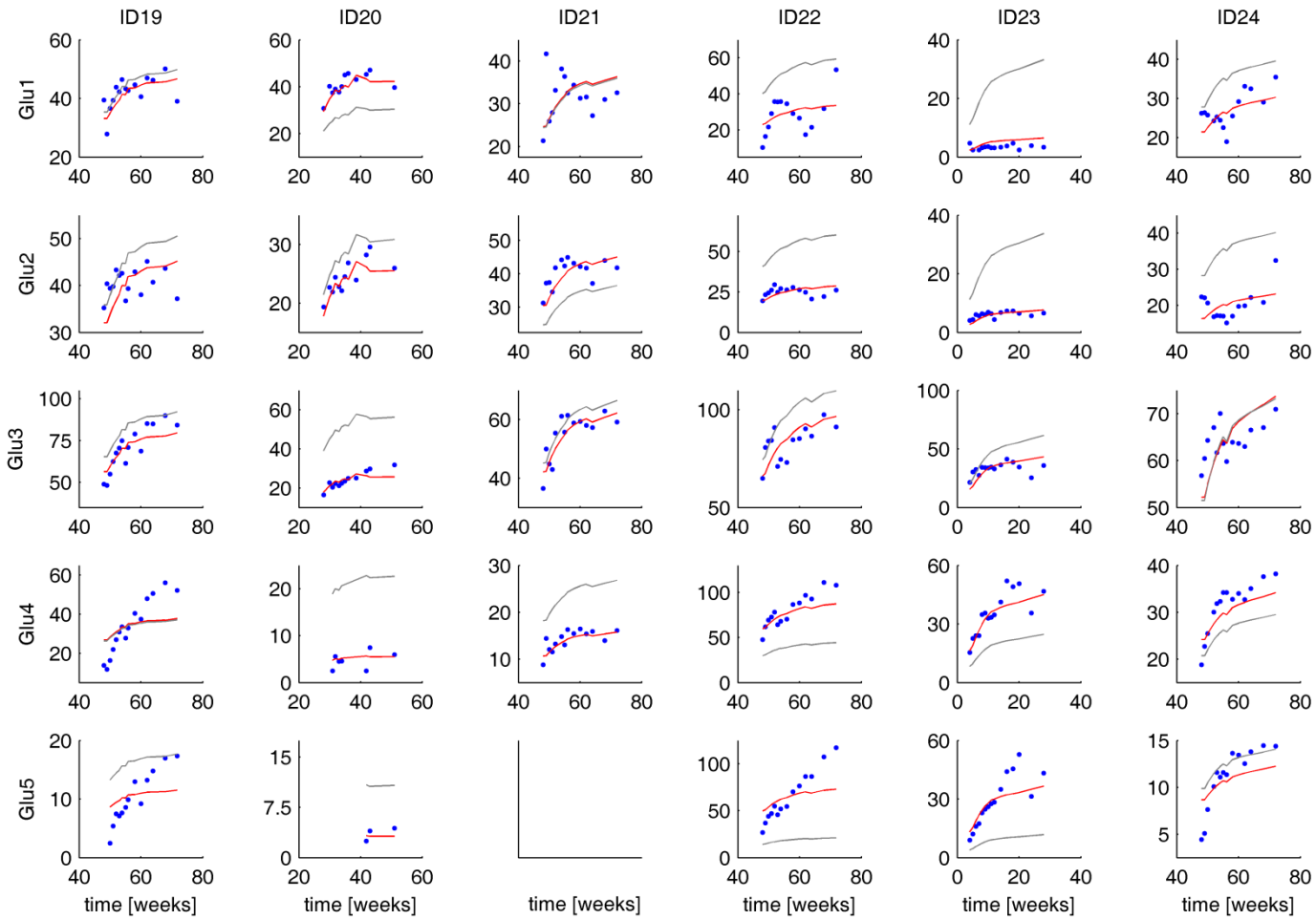
App Fig. 6.2.1: *Individual fits obtained with the final updated PM model for all MTXGluX: ID1 to ID6.*



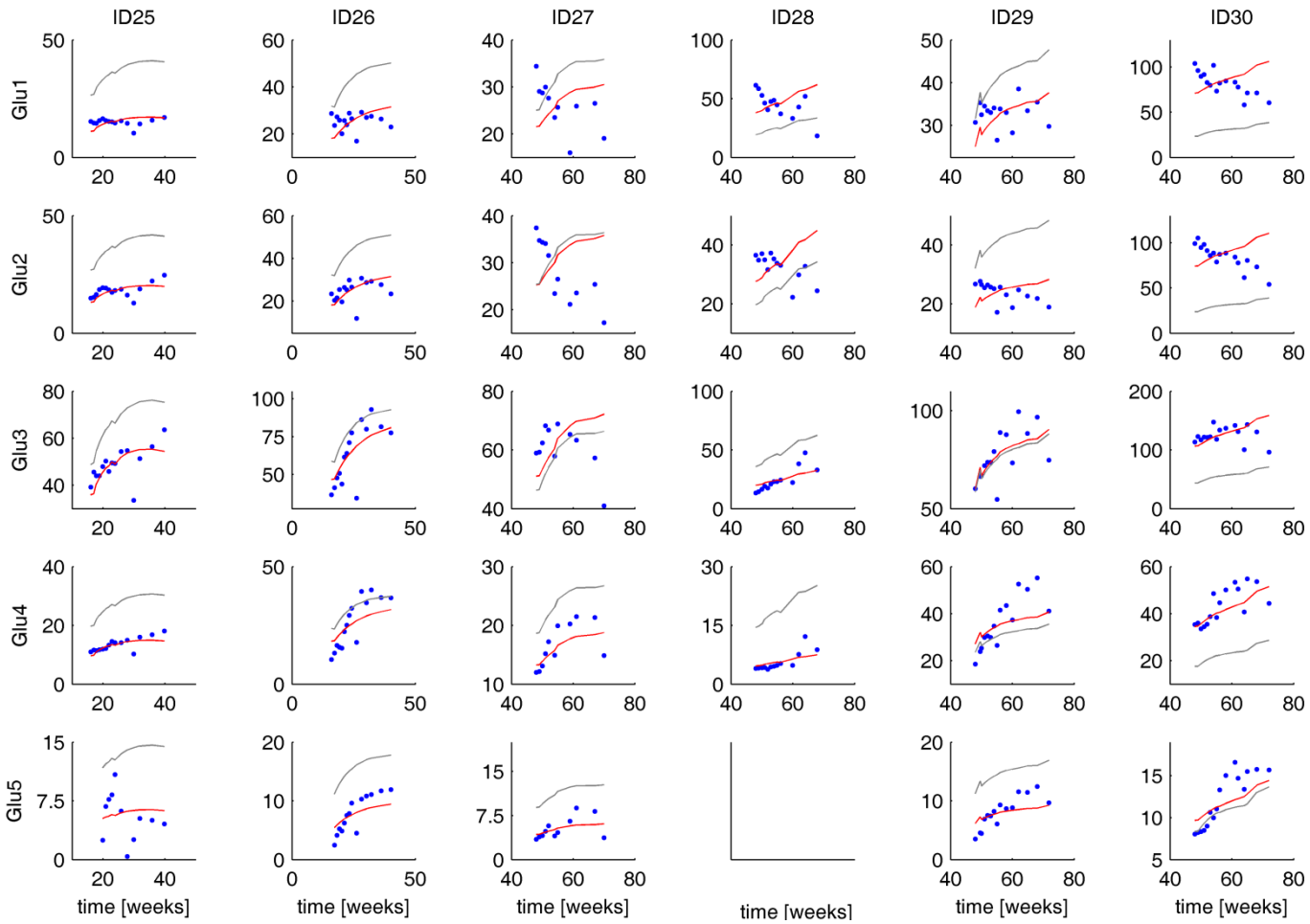
App Fig. 6.2.2: *Individual fits obtained with the final updated PM model for all MTXGluX: ID7 to ID12.*



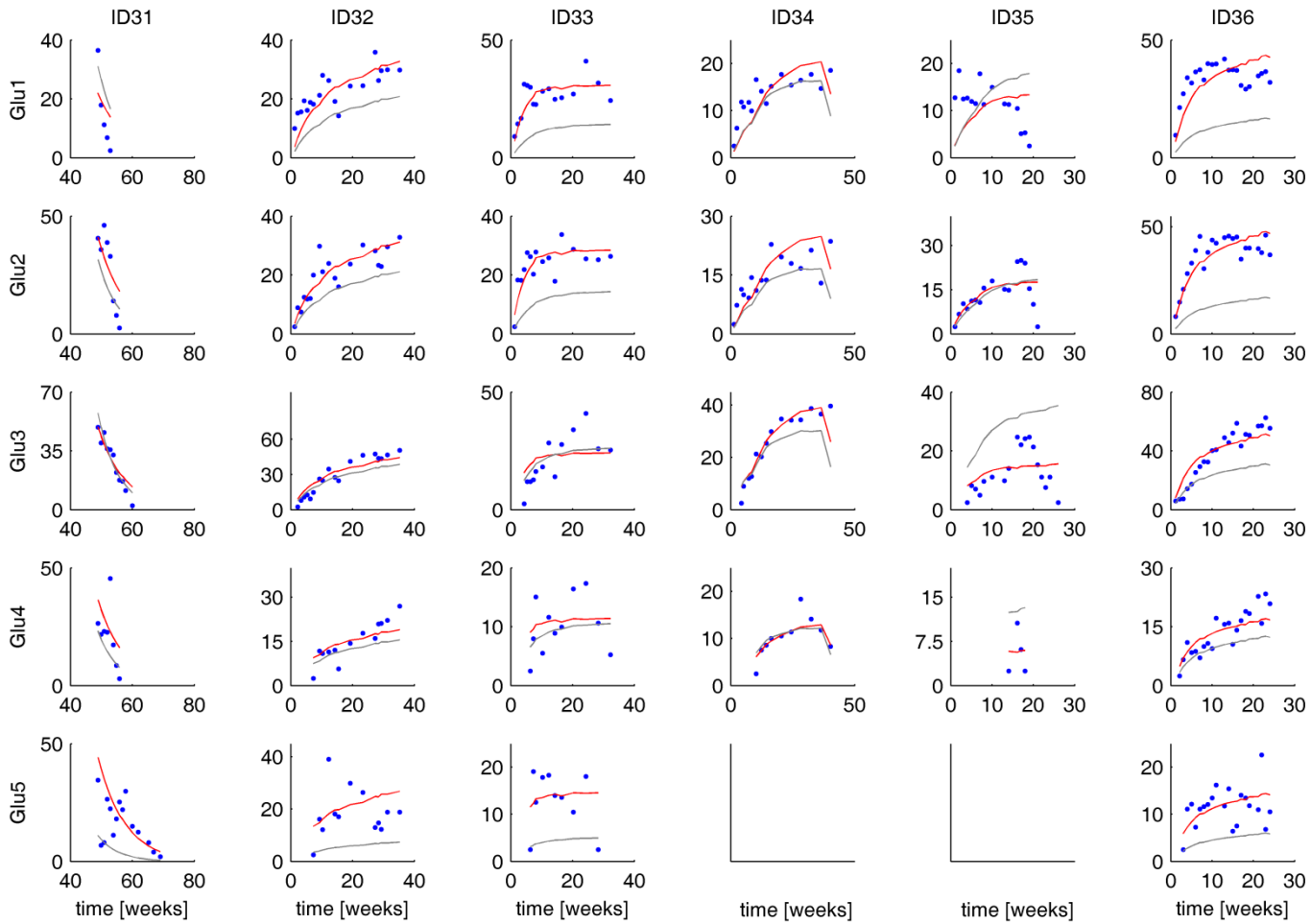
App Fig. 6.2.3: *Individual fits obtained with the final updated PM model for all MTXGluX: ID13 to ID18.*



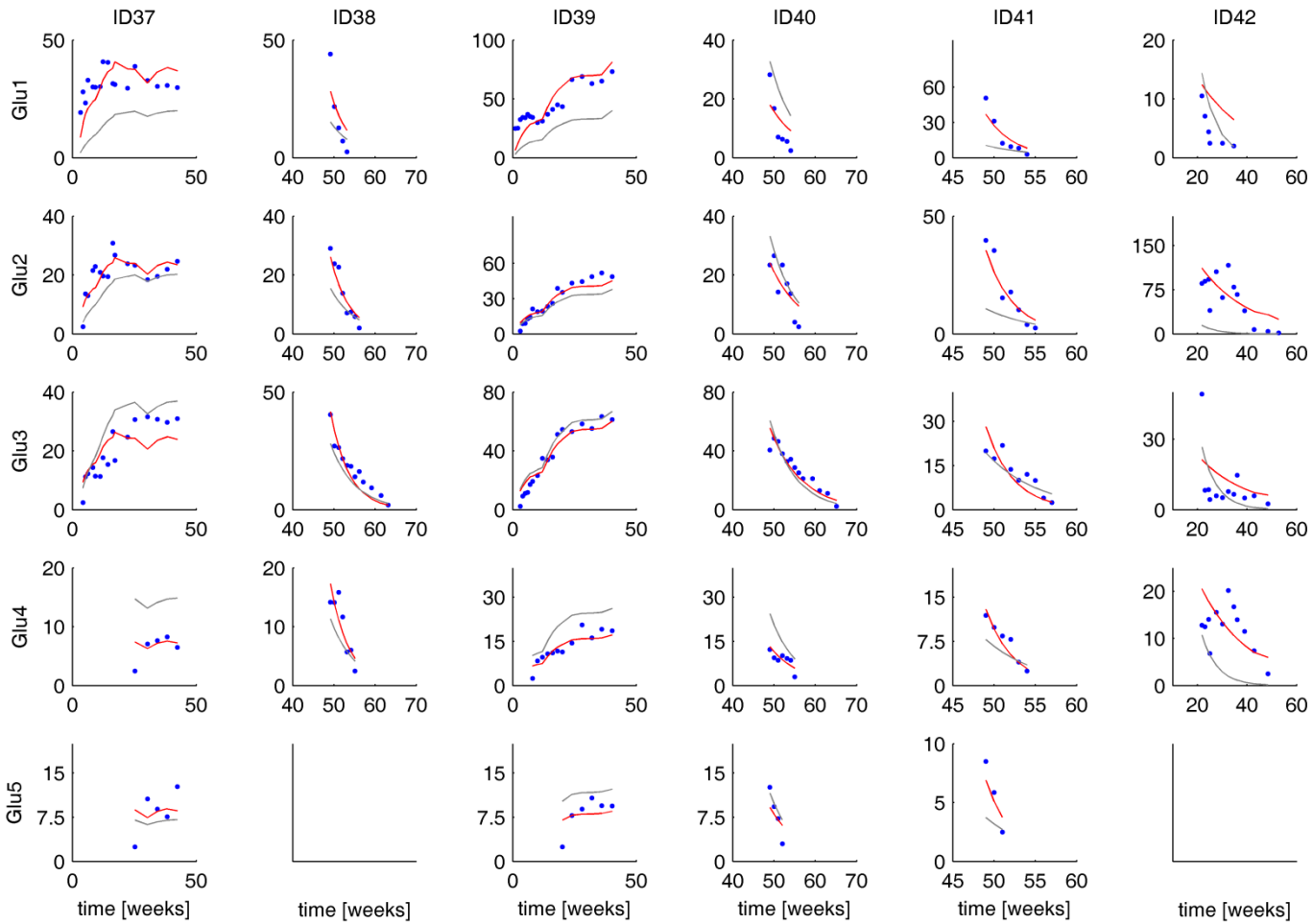
App Fig. 6.2.4: *Individual fits obtained with the final updated PM model for all MTXGluX: ID19 to ID24.*



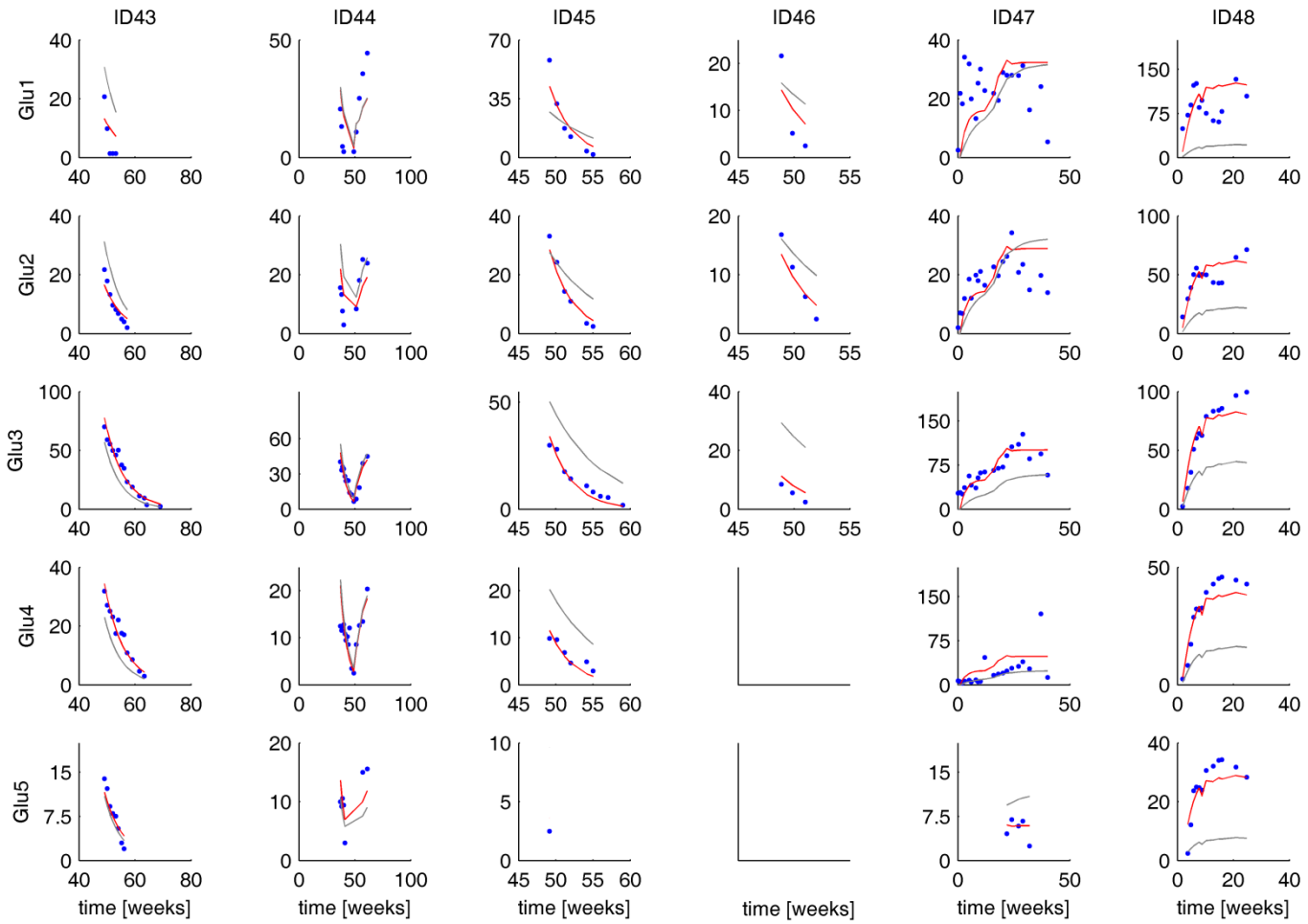
App Fig. 6.2.5: *Individual fits obtained with the final updated PM model for all MTXGluX: ID25 to ID30.*



App Fig. 6.2.6: *Individual fits obtained with the final updated PM model for all MTXGluX: ID31 to ID36.*



App Fig. 6.2.7: *Individual fits obtained with the final updated PM model for all MTXGluX: ID37 to ID42.*



App Fig. 6.2.8: *Individual fits obtained with the final updated PM model for all MTXGluX: ID43 to ID48.*

References

1. Bonate P (2006) *Pharmacokinetic-pharmacodynamic modeling and simulation*, 1st edn. Springer, New York (USA).
2. Gibaldi M, Perrier D (1982) *Pharmacokinetics*, 2nd edn. Marcel-Dekker, New York (USA).
3. Rowland M, Tozer T (2010) *Clinical pharmacokinetics and pharmacodynamics: concepts and applications*, 4th edn. Lippincott Williams & Wilkins, Baltimore (USA).
4. Rosenbaum S (2011) *Basic Pharmacokinetics and Pharmacodynamics - An integrated textbook and computer simulations*, 1st edn. Wiley, Hoboken, New Jersey (USA).
5. Box G, Draper N (1987) *Empirical Model Building and Response Surfaces*, 1st edn. Jon Wiley & Sons, Inc., New York (USA).
6. Gelman A, Carlin J, Stern H, Rubin D (1995) *Bayesian data analysis*, 1st edn. Chapman and Hall, London (UK).
7. Benet L (1984) *Pharmacokinetics: Basic Principles and Its Use as a Tool in Drug Metabolism*. In: Mitchell J, Horning M (eds.) Drug Metabolism and Drug Toxicity, 1st edn. Raven Press, New York (USA).
8. Morrison P, Allegra C (1987) *The kinetics of methotrexate polyglutamation in human breast cancer cells*. Arch. Biochem. Biophys. 254(2):597-610.
9. Savic R, Jonker D, Kerbusch T, Karlsson M (2004) *Evaluation of a transit compartment model versus a lag time model for describing drug absorption delay*. Population Analysis Group Europe (PAGE) 13, Abstract 513, Uppsala, Sweden.
10. Savic R, Jonker D, Kerbusch T, Karlsson M (2007) *Implementation of a transit compartment model for describing drug absorption in pharmacokinetic studies*. J. Pharmacokinet. Pharmacodyn. 34(5):711-726.
11. Hill A (1910) *The possible effects of the aggregation of haemoglobin on its dissociation curve*. J. Physiol. 40:4-7.
12. Chan P, Holford N (2001) *Drug treatment effects on disease progression*. Annu. Rev. Pharmacol. Toxicol. 41(1):625-659.
13. Jusko W, Ko H (1994) *Physiologic indirect response models characterize diverse types of pharmacodynamic effects*. Clin. Pharmacol. Ther. 56(4):406-419.
14. Holford N, Sheiner L (1982) *Kinetics of pharmacologic response*. Pharmacol. Ther. 16(2):143-166.
15. Holford N, Sheiner L (1981) *Understanding the dose-effect relationship: Clinical application of pharmacokinetic-pharmacodynamic models*. Clin. Pharmacokinet. 6:429-453.
16. Al-Sallami H, Kumar P, Landersdorfer C, Bulitta J, Duffull S (2009) *The time course of drug effects*. Pharm. Stat. 8(3):176-185.
17. Mager D, Wyska E, Jusko W (2003) *Diversity of mechanism-based pharmacodynamic models*. Drug Metab. Dispos. 31(5):510-518.
18. Collett D (2003) *Modelling survival data in medical research*, 2nd edn. Chapman & Hall/CRC, Boca Raton (USA).

19. Sheiner L, Beal S (1980) *Evaluation of methods for estimating population pharmacokinetic parameters. I. Michaelis-menten model: Routine clinical pharmacokinetic data*. J. Pharmacokinet. Pharmacodyn. 8(6):553-571.
20. Davidian M, Giltinan D (1995) *Nonlinear models for repeated measurement data*, 1st edn. Chapman & Hall/CRC, Boca Raton (USA).
21. Sheiner L, Beal S (1981) *Evaluation of methods for estimating population pharmacokinetic parameters II. Biexponential model and experimental pharmacokinetic data*. J. Pharmacokinet. Biopharm. 9(5):635-651.
22. Ette E, Williams P (2004) *Population pharmacokinetics II: estimation methods*. Ann. Pharmacother. 38(11):1907-1915.
23. Beal S, Sheiner L, Boeckmann A, Bauer R (1989-2009) *NONMEM User's Guides*. Icon Development Solutions, Ellicott City, MD (USA).
24. Lavielle M (2005) *MONOLIX 1.1 User manual*. Laboratoire de Mathematiques, Universite Paris-Sud, Orsay (France).
25. Peck C, Beal S, Sheiner L, Nichols A (1984) *Extended least squares nonlinear regression: A possible solution to the "choice of weights" problem in analysis of individual pharmacokinetic data*. J. Pharmacokinet. Pharmacodyn. 12(5):545-558.
26. Kuhn E, Lavielle M (2005) *Maximum likelihood estimation in nonlinear mixed effects models*. Comput. Stat. Data Anal. 49(4):1020-1038.
27. Lavielle M, Monolix Group (2004) *SAEM in MATLAB: an alternative to linearization*. Population Analysis Group Europe (PAGE) 13, Abstract 544, Uppsala, Sweden.
28. Foo L, Duffull S (2011) *Optimal Design of Pharmacokinetic-Pharmacodynamic Studies*. In: Bonate P, Howard D (eds.) Pharmacokinetics in Drug Development: Advances and Applications., 1st edn. Springer, New York (USA).
29. Atkinson A, Donev A (1992) *Optimum Experimental Designs*, 1st edn. Oxford University Press, Oxford (UK).
30. Pronzato L, Walter E (1985) *Robust experiment design via stochastic approximation*. Math. Biosci. 75(1):103-120.
31. D'argenio D (1990) *Incorporating prior parameter uncertainty in the design of sampling schedules for pharmacokinetic parameter estimation experiments*. Math. Biosci. 99(1):105-118.
32. Pronzato L, Walter E (1988) *Robust experiment design via maximin optimization*. Math. Biosci. 89(2):161-176.
33. Foo L, Duffull S (2010) *Methods of Robust Design of Nonlinear Models with an Application to Pharmacokinetics*. J. Biopharm. Stat. 20(4):886-902.
34. Greer J, Foerster J, Lukens J, Rodgers G, Paraskevas F, Glader B (eds.) (2004) *Wintrobe's Clinical Hematology*. 11th edn. Lippincott Williams & Wilkins, Philadelphia, USA.
35. Berlin N, Waldmann T, Weissmann S (1959) *Life span of red blood cell*. Physiol. Rev. 39(3):577-616.
36. Franco R (2009) *The measurement and importance of red cell survival*. Am. J. Hematol. 84(2):109-114.

37. Landaw S (1988) *Factors that accelerate or retard red blood cell senescence*. Blood Cells 14(1):47-59.
38. Eadie G, Brown I (1953) *Analytic review: Red blood cell survival studies*. Blood 8:1110-1136.
39. Giblett E, Coleman D, Pirzio-Biroli G, Donohue D, Motulsky A, Finch C (1956) *Erythrokinetics: Quantitative measurements of red cell production and destruction in normal subjects and patients with anemia*. Blood 11(4):291-309.
40. Finch C (1959) *Some quantitative aspects of erythropoiesis*. Ann. N. Y. Acad. Sci. 77(Hematopoietic Mechanisms):410-416.
41. Alfrey C, Rice L, Udden M, Driscoll T (1997) *Neocytolysis: Physiological down-regulator of red-cell mass*. Lancet 349(9062):1389-1390.
42. Rice L, Alfrey C, Driscoll T, Whitley C, Hachey D, Suki W (1999) *Neocytolysis contributes to the anemia of renal disease*. Am. J. Kidney Dis. 33(1):59-62.
43. Joske R, Mcalister J, Prankerd A (1956) *Isotope investigations of red cell production and destruction in chronic renal disease*. Clin. Sci. 15(4):511-522.
44. Ly J, Marticorena R, Donnelly S (2004) *Red blood cell survival in chronic renal failure*. Am. J. Kidney Dis. 44(4):715-719.
45. Rifkind R (1966) *Destruction of injured red cells in vivo*. Am. J. Med. 41:711-723.
46. Piomelli S, Seaman C (1993) *Mechanism of red blood cell aging: Relationship of cell density and cell age*. Am. J. Hematol. 42:46-52.
47. Gardos G (1958) *The function of calcium in the potassium permeability of human erythrocytes*. Biochim. Biophys. Acta 30(3):653-654.
48. Trial J, Rice L, Alfrey C (2001) *Erythropoietin withdrawal alters interactions between young red blood cells, splenic endothelial cells, and macrophages: an in vitro model of neocytolysis*. J. Investig. Med. 49(4):335-345.
49. Trial J, Rice L (2004) *Erythropoietin withdrawal leads to the destruction of young red cells at the endothelial-macrophage interface*. Curr. Pharm. Des. 10(2):183-190.
50. Rice L, Alfrey C (2005) *The negative regulation of red cell mass by neocytolysis: Physiologic and pathophysiologic manifestations*. Cell. Physiol. Biochem. 15(6):245-250.
51. Alfrey C, Fishbane S (2007) *Implications of neocytolysis for optimal management of anaemia in chronic kidney disease*. Nephron Clin. Pract. 106 (4):c149-c156.
52. Berlin N (1964) *Determination of red blood cell life span*. JAMA 188:375-378.
53. Ross J, Chapin M (1943) *Effect of storage of citrated blood on the survival of transfused erythrocytes*. JAMA 123(13):827-829.
54. Shemin D, Rittenberg D (1946) *The life span of the human red blood cell*. J. Biol. Chem. 166:627-636.
55. Ashby W (1919) *The determination of the length of life of transfused blood corpuscles in man*. J. Exp. Med. 29:267-282.
56. Gray S, Sterling K (1950) *The tagging of red cells and plasma proteins with radioactive chromium*. J. Clin. Invest. 29(12):1604-1613.

57. Ebaugh F, Emerson C, Roos J (1953) *The use of radioactive chromium 51 as an erythrocyte tagging agent for the determination of red cell survival in vivo*. J. Clin. Invest. 29(12):1260-1276.
58. Mollison P, Veall N (1955) *The use of the isotope ^{51}Cr as a label for red blood cells*. Br. J. Haematol. 1(1):62-74.
59. Cohen J, Warringa M (1954) *The fate of P^{32} labelled diisopropylfluorophosphonate in the human body and its use as a labelling agent in the study of the turnover of blood plasma and red cells*. J. Clin. Invest. 33(3):459-467.
60. Eernisse J, Van Rood J (1961) *Erythrocyte survival-time determination with the aid of $DF^{32}P$*. Br. J. Haematol. 7(3):382-404.
61. Garby L (1962) *Analysis of red cell survival curves in clinical practice and the use of di-iso-propylfluorophosphonate ($DF^{32}P$) as a label for red cells in man*. Br. J. Haematol. 8(1):15-27.
62. Brouillard R (1974) *Measurement of red blood cell life-span*. JAMA 230(9):1304-1305.
63. Strocchi A, Schwartz S, Ellefson M, Engel R, Medina A, Levitt M (1992) *A simple carbon monoxide breath test to estimate erythrocyte turnover*. J. Lab. Clin. Med. 120(3):392-399.
64. Medina A, Ellis C, Levitt M (1994) *Use of alveolar carbon monoxide measurements to assess red blood cell survival in hemodialysis patients*. Am. J. Hematol. 46(2):91-94.
65. Krishnan S, Dixit N (2009) *Estimation of red blood cell lifespan from alveolar carbon monoxide measurements*. Transl. Res. 154(1):15-17.
66. International Committee for Standardization in Haematology (1980) *Recommended method for radioisotope red-cell survival studies*. British Journal of Haematology 45(4):659-666.
67. Cline M, Berlin N (1963) *The red cell chromium elution rate in patients with some hematologic diseases*. Blood 21(1):63-69.
68. Willekens F, Roerdinkholder-Stoelwinder B, Groenen-Doepp Y, Bos H, Bosman G, Van Den Bos A, Verkleij A, Werre J (2003) *Hemoglobin loss from erythrocytes in vivo results from spleen-facilitated vesiculation*. Blood 101(2):747-751.
69. Dornhorst A (1951) *Analytic review: The interpretation of red cell survival curves*. Blood 6:1284-1292.
70. Woo S, Krzyzanski W, Jusko W (2006) *Pharmacokinetic and pharmacodynamic modeling of recombinant human erythropoietin after intravenous and subcutaneous administration in rats*. J. Pharmacol. Exp. Ther. 319(3):1297-1306.
71. Krzyzanski W, Perez Ruixo J (2012) *Lifespan based indirect response models*. J. Pharmacokinet. Pharmacodyn. 39(1):109-123.
72. Krzyzanski W, Ramakrishnan R, Jusko W (1999) *Basic pharmacodynamic models for agents that alter production of natural cells*. J. Pharmacokinet. Biopharm. 27(5):467-489.

73. Krzyzanski W, Perez-Ruixo J, Vermeulen A (2008) *Basic pharmacodynamic models for agents that alter the lifespan distribution of natural cells*. J. Pharmacokinet. Pharmacodyn. 35(3):349-377.
74. Freise K, Widness J, Schmidt R, Veng-Pedersen P (2007) *Pharmacodynamic analysis of time-variant cellular disposition: Reticulocyte disposition changes in phlebotomized sheep*. J. Pharmacokinet. Pharmacodyn. 34(4):519-547.
75. Krzyzanski W, Woo S, Jusko W (2006) *Pharmacodynamic models for agents that alter production of natural cells with various distributions of lifespans*. J. Pharmacokinet. Pharmacodyn. 33(2):125-166.
76. Freise K, Widness J, Schmidt R, Veng-Pedersen P (2008) *Modeling time variant distributions of cellular lifespans: Increases in circulating reticulocyte lifespans following double phlebotomies in sheep*. J. Pharmacokinet. Pharmacodyn. 35(3):285-323.
77. Krzyzanski W (2011) *Interpretation of transit compartments pharmacodynamic models as lifespan based indirect response models* J. Pharmacokinet. Pharmacodyn. 38(2):179-204.
78. Hamren B, Bjoerk E, Karlsson M (2006) *Mechanism-based pharmacokinetic and pharmacodynamic modelling of tesaglitazar in type 2 diabetes patients*. Population Approach Group in Europe (PAGE) 15, Abstract 961, Bruges, Belgium.
79. Ribbing J, Hamren B, Svensson M, Karlsson M (2010) *A model for glucose, insulin, and beta-cell dynamics in subjects with insulin resistance and patients with type 2 diabetes*. J. Clin. Pharmacol. 50(8):861-872.
80. Kalicki R, Lledo R, Karlsson M (2009) *Modeling of red blood cell life-span in a hematologically normal population*. Population Approach Group in Europe (PAGE) 18, Abstract 1677, St. Petersburg, Russia.
81. Korell J (2009) *A model for red blood cell survival time*. Dissertation towards a Postgraduate Certificate in Pharmacy, University of Otago.
82. Korell J, Coulter C, Duffull S (2011) *A statistical model for red blood cell survival*. Journal of Theoretical Biology 268(1):39-49.
83. Bebbington M, Lai C, Zitikis R (2007) *Modeling human mortality using mixtures of bathtub shaped failure distributions*. Journal of Theoretical Biology 245(3):528-538.
84. Smith A, Gelfand A (1992) *Bayesian statistics without tears: A sampling-resampling perspective*. Amer. Statist. 46(2):84-88.
85. Press W, Teukolsky S, Vetterling W, Flannery B (1996) *Chapter 7.3: Rejection method: Gamma, Poisson, Binominal deviates*. In: Numerical Recipes in Fortran 77: The art of scientific computing, 2nd edn. Cambridge University Press, Cambridge (UK).
86. Hardman J, Limbird L (eds.) (2001) *Goodman & Gilman's: The Pharmacological Basis of Therapeutics*. 10th edn. McGraw-Hill, New York, USA.
87. Kurnik D, Loebstein R, Fishbein E, Almog S, Halkin H, Bar-Meir S, Chowers Y (2003) *Bioavailability of oral vs. subcutaneous low-dose*

methotrexate in patients with Crohn's disease. Aliment. Pharmacol. Ther. 18(1):57-63.

88. Herman R, Veng Pedersen P, Hoffman J, Koehnke R, Furst D (1989) *Pharmacokinetics of low dose methotrexate in rheumatoid arthritis patients*. J. Pharm. Sci. 78(2):165-171.
89. Godfrey C, Sweeney K, Miller K, Hamilton R, Kremer J (1998) *The population pharmacokinetics of long term methotrexate in rheumatoid arthritis*. Br. J. Clin. Pharmacol. 46(4):369-376.
90. Panetta J, Yanishevski Y, Pui C, Sandlund J, Rubnitz J, Rivera G, Ribeiro R, Evans W, Relling M (2002) *A mathematical model of in vivo methotrexate accumulation in acute lymphoblastic leukemia*. Cancer Chemother. Pharmacol. 50(5):419-428.
91. Hoekstra M, Haagsma C, Neef C, Proost J, Knuif A, Van De Laar M (2004) *Bioavailability of higher dose methotrexate comparing oral and subcutaneous administration in patients with rheumatoid arthritis*. J. Rheumatol. 31(4):645-648.
92. Johansson A, Hill N, Perisoglou M, Whelan J, Karlsson M, Standing J (2011) *A Population Pharmacokinetic/Pharmacodynamic Model of Methotrexate and Mucositis Scores in Osteosarcoma*. Ther. Drug Monit. 33(6):711-718.
93. Sirotnak F, Tolner B (1999) *Carrier-mediated membrane transport of folates in mammalian cells*. Annu. Rev. Nutr. 19(1):91-122.
94. Ritari S, Sakami W, Black C, Rzepka J (1975) *The determination of folylpolyglutamate synthetase*. Anal. Biochem. 63(1):118-129.
95. Chabner B, Allegra C, Curt G, Clendeninn N, Baram J, Koizumi S, Drake J, Jolivet J (1985) *Polyglutamation of methotrexate. Is methotrexate a prodrug?* J. Clin. Investig. 76(3):907-912.
96. Moran R, Colman P, Forsch R, Rosowsky A (1984) *A mechanism for the addition of multiple moles of glutamate by folylpolyglutamate synthetase*. J. Med. Chem. 27(10):1263-1267.
97. Rhee M, Lindau-Shepard B, Chave K, Galivan J, Ryan T (1998) *Characterization of Human Cellular Î³-Glutamyl Hydrolase*. Mol. Pharmacol. 53(6):1040-1046.
98. Mcguire J, Bertino J (1981) *Enzymatic synthesis and function of folylpolyglutamates*. Mol. Cell. Biochem. 38(1):19-48.
99. Zeng H, Chen Z, Belinsky M, Rea P, Kruh G (2001) *Transport of Methotrexate (MTX) and Folates by Multidrug Resistance Protein (MRP) 3 and MRP1*. Cancer Res. 61(19):7225-7232.
100. White J, Goldman I (1976) *Mechanism of Action of Methotrexate*. Mol. Pharmacol. 12(5):711-719.
101. Allegra C, Drake J, Jolivet J, Chabner B (1985) *Inhibition of phosphoribosylaminoimidazolecarboxamide transformylase by methotrexate and dihydrofolic acid polyglutamates*. Proc. Natl. Acad. Sci. U. S. A. 82(15):4881-4885.
102. Szeto D, Cheng Y, Rosowsky A, Yu C, Modest E, Piper J, Temple C, Elliott R, Rose J, Montgomery J (1979) *Human thymidylate synthetase - III:*

Effects of methotrexate and folate analogs. Biochem. Pharmacol. 28(17):2633-2637.

103. Baggott J, Vaughn W, Hudson B (1986) *Inhibition of 5-aminoimidazole-4-carboxamide ribotide transformylase, adenosine deaminase and 5'-adenylate deaminase by polyglutamates of methotrexate and oxidized folates and by 5-aminoimidazole-4-carboxamide riboside and ribotide*. Biochem. J. 236(1):193-200.
104. Cronstein B, Naime D, Ostad E (1993) *The antiinflammatory mechanism of methotrexate. Increased adenosine release at inflamed sites diminishes leukocyte accumulation in an in vivo model of inflammation*. J. Clin. Investig. 92(6):2675-2682.
105. Cutolo M, Sulli A, Pizzorni C, Seriolo B, Straub R (2001) *Anti-inflammatory mechanisms of methotrexate in rheumatoid arthritis*. Ann. Rheum. Dis. 60(8):729-735.
106. Stamp L, Roberts R (2011) *Effect of genetic polymorphisms in the folate pathway on methotrexate therapy in rheumatic diseases*. Pharmacogenomics 12(10):1449-1463.
107. Allegra C, Chabner B, Drake J, Lutz R, Rodbard D, Jolivet J (1985) *Enhanced inhibition of thymidylate synthase by methotrexate polyglutamates*. J. Biol. Chem. 260(17):9720-9726.
108. Stamp L, Roberts R, Kennedy M, Barclay M, O'donnell J, Chapman P (2006) *The use of low dose methotrexate in rheumatoid arthritis - are we entering a new era of therapeutic drug monitoring and pharmacogenomics?* Biomed. Pharmacother. 60(10):678-687.
109. Saibeni S, Bollani S, Losco A, Michielan A, Sostegni R, Devani M, Lupinacci G, Pirola L, Cucino C, Meucci G, Basilisco G, D'inca R, Bruno S (2012) *The use of methotrexate for treatment of inflammatory bowel disease in clinical practice*. Dig. Liver Dis. 44(2):123-127.
110. Visser K, Katchamart W, Loza E, Martinez-Lopez J, Salliot C, Trudeau J, Bombardier C, Carmona L, Van Der Heijde D, Bijlsma J (2009) *Multinational evidence-based recommendations for the use of methotrexate in rheumatic disorders with a focus on rheumatoid arthritis: integrating systematic literature research and expert opinion of a broad international panel of rheumatologists in the 3E Initiative*. Ann. Rheum. Dis. 68(7):1086-1093.
111. Visser K, Van Der Heijde D (2009) *Optimal dosage and route of administration of methotrexate in rheumatoid arthritis: a systematic review of the literature*. Ann. Rheum. Dis. 68(7):1094-1099.
112. Sinnett M (1989) *Methotrexate pharmacokinetics in patients with rheumatoid arthritis*. J. Rheumatol. 16(6):745-748.
113. Tishler M, Caspi D, Graff E, Segal R, Peretz H, Yaron M (1989) *Synovial and serum levels of methotrexate during methotrexate therapy of rheumatoid arthritis*. Rheumatology 28(5):422-423.
114. Dervieux T, Furst D, Lein D, Capps R, Smith K, Walsh M, Kremer J (2004) *Polyglutamation of methotrexate with common polymorphisms in reduced folate carrier, aminoimidazole carboxamide ribonucleotide transformylase, and thymidylate synthase are associated with methotrexate effects in rheumatoid arthritis*. Arthritis. Rheum. 50(9):2766-2774.

115. Angelis-Stoforidis P, Vajda F, Christophidis N (1999) *Methotrexate polyglutamate levels in circulating erythrocytes and polymorphs correlate with clinical efficacy in rheumatoid arthritis*. Clin. Exp. Rheumatol. 17(3):313-320.
116. Dervieux T, Furst D, Lein D, Capps R, Smith K, Caldwell J, Kremer J (2005) *Pharmacogenetic and metabolite measurements are associated with clinical status in patients with rheumatoid arthritis treated with methotrexate: results of a multicentred cross sectional observational study*. Ann. Rheum. Dis. 64(8):1180-1185.
117. Stamp L, O'donnell J, Chapman P, Zhang M, James J, Frampton C, Barclay M (2010) *Methotrexate polyglutamate concentrations are not associated with disease control in rheumatoid arthritis patients receiving long-term methotrexate therapy*. Arthritis. Rheum. 62(2):359-368.
118. Krzyzanski W, Jusko W (2002) *Multiple-pool cell lifespan model of hematologic effects of anticancer agents*. J. Pharmacokinet. Pharmacodyn. 29(4):311-337.
119. Clark M, Shohet S (1985) *Red cell senescence*. Clin. Haematol. 14(1):223-257.
120. Schiodt E (1938) *On the duration of life of the red blood corpuscles*. Acta Med. Scand. 95:49-79.
121. Cline M, Berlin N (1963) *The reticulocyte count as an indicator of the rate of erythropoiesis*. Am. J. Clin. Pathol. 39(2):121-128.
122. Sandberg S, Rustard P, Johannesen B, Stolsnes B (1998) *Within-subject biological variation of reticulocytes and reticulocyte-derived parameters*. Eur. J. Haematol. 61:42-48.
123. Seip M (1953) *Reticulocyte studies*. Acta Med. Scand. Suppl. 282.
124. Cohen R, Franco R, Khera P, Smith E, Lindsell C, Ciraolo P, Palascak M, Joiner C (2008) *Red cell life span heterogeneity in hematologically normal people is sufficient to alter HbA1c*. Blood 112(10):4284-4291.
125. London I, Shemin D, West R, Rittenberg D (1949) *Heme synthesis and red blood cell dynamics in normal humans and in subjects with polycythemia vera, sickle-cell anemia, and pernicious anemia*. J. Biol. Chem. 179(1):463-484.
126. Wilson C, Hopkins P, Cabello-Inchausti B, Melnick S, Robinson M (2000) *The peripheral blood smear in patients with sickle cell trait: A morphologic observation*. Lab. Medicine 31(8):445-447.
127. Uehlinger D, Gotch F, Sheiner L (1992) *A pharmacodynamic model of erythropoietin therapy for uremic anemia*. Clin. Pharmacol. Ther. 51(1):76-89.
128. Goldstein D, Little R, Lorenz R, Malone J, Nathan D, Peterson C, Sacks D (2004) *Tests of glycemia in diabetes*. Diabetes Care 27(7):1761-1773.
129. Cavill I (2002) *Red cell lifespan estimation by 51Cr labelling*. Br. J. Haematol. 117(4):997-997.
130. Furne J, Springfield J, Ho S, Levitt M (2003) *Simplification of the end-alveolar carbon monoxide technique to assess erythrocyte survival*. J. Lab. Clin. Med. 142(1):52-57.

131. Virtue M, Furne J, Nuttall F, Levitt M (2004) *Relationship between GHb concentration and erythrocyte survival determined from breath carbon monoxide concentration*. Diabetes Care 27(4):931-935.
132. Garby L, Groth T, Schneider W (1969) *Determination of kinetic parameters of red blood cell survival by computer simulation*. Comput. Biomed. Res. 2(3):229-241.
133. Corbett G, O'flaherty E, Liang J, Throop L, Finley B, Kerger B (1998) *In vitro reduction kinetics of hexavalent chromium in human blood*. Environ. Res. 78(Section A):7-11.
134. Willekens F, Bosch F, Roerdinkholder-Stoelwinder B, Groenen-Doepp Y, Werre J (1997) *Quantification of loss of haemoglobin components from the circulating red blood cell in vivo*. Eur. J. Haematol. 58(4):246-250.
135. Bosman G, Willekens F, Werre J (2005) *Erythrocyte aging: A more than superficial resemblance to apoptosis?* Cell. Physiol. Biochem. 16:1-8.
136. Werre J, Willekens F, Bosch F, De Haan L, Van Der Vegt S, Van Den Bos A, Bosman G (2004) *The red cell revisited - Matters of life and death*. Cell. Mol. Biol. 50(2):139-145.
137. Bosman G, Werre J, Willekens F, Novotny V (2008) *Erythrocyte ageing in vivo and in vitro: Structural aspects and implications for transfusion*. Transfus. Med. 18(6):335-347.
138. Audi G, Bersillon O, Blachot J, Wapstra A (1997) *The NUBASE evaluation of nuclear and decay properties*. Nucl. Phys. A 624(1):1-124.
139. Hughes Jones N, Mollison P (1956) *The interpretation of measurements with ^{51}Cr-labelled red cells*. Clin. Sci. 15(2):207-218.
140. Eadie G, Brown I, Curtis W (1955) *The potential life span and ultimate survival of fresh red blood cells in normal healthy recipients as studied by simultaneous Cr^{51} tagging and differential hemolysis*. J. Clin. Invest. 34(4):629-636.
141. Currie L (1968) *Limits for qualitative detection and quantitative determination. Application to radiochemistry*. Anal. Chem. 40(3):586-593.
142. Ayres R, Shaw J, Mills C, Coleman R, Neuberger J (1991) *A 51Cr release cytotoxicity assay for use with human intrahepatic biliary epithelial cells*. J. Immunol. Methods 141(1):117-122.
143. Freud A, Canfi A, Ben-Hur E (1993) *Validity of Chromium-51 Labelling of Human Erythrocytes as an Assay for Their Survival after Photodynamic Therapy in Vitro for Blood Sterilization*. Int. J. Radiat. Biol. 63(5):651-653.
144. Ramsay W, Fulton J (1964) *A new method for the assay of small amounts of radioactive iron (59Fe)*. Anal. Biochem. 8(3):319-324.
145. Gifford S, Derganc J, Shevkoplyas S, Yoshida T, Bitensky M (2006) *A detailed study of time-dependent changes in human red blood cells: From reticulocyte maturation to erythrocyte senescence*. Br. J. Haematol. 135(3):395-404.
146. Greenwalt T (2006) *The how and why of exocytic vesicles*. Transfusion 46:143-152.

147. Szymanski I, Valeri C (1970) *Factors influencing chromium elution from labelled red cells in vivo and the effect of elution on red-cell survival measurements*. Br. J. Haematol. 19(3):397-409.
148. Amatuzio D, Evans R (1953) *Isotope Curves in Red Blood Cell Studies*. Nature 171:797-798.
149. Mock D, Lankford G, Widness J, Burmeister L, Kahn D, Strauss R (1999) *Measurement of red cell survival using biotin-labeled red cells: Validation against 51Cr-labeled red cells*. Transfusion 39(2):156-162.
150. Suzuki T, Dale G (1987) *Biotinylated erythrocytes: in vivo survival and in vitro recovery*. Blood 70(3):791-795.
151. Hsu C, Mcculloch C, Curhan G (2002) *Epidemiology of anemia associated with chronic renal insufficiency among adults in the United States: results from the Third National Health and Nutrition Examination Survey*. J. Am. Soc. Nephrol. 13(2):504-510.
152. Mcclellan W, Aronoff S, Bolton W, Hood S, Lorber D, Tang K, Tse T, Wasserman B, Leiserowitz M (2004) *The prevalence of anemia in patients with chronic kidney disease*. Curr. Med. Res. Opin. 20(9):1501-1510.
153. Eschbach J, Funk D, Adamson J, Kuhn I, Scribner B, Finch C (1967) *Erythropoiesis in patients with renal failure undergoing chronic dialysis*. N. Engl. J. Med. 276(12):653-658.
154. Loge J, Lange R, Moore C (1958) *Characterization of the anemia associated with chronic renal insufficiency*. Am. J. Med. 24:4-18.
155. Desforge J, Dawson J (1958) *The anemia of renal failure*. Arch. Intern. Med. 101:326-332.
156. Schwartz A, Keich B, Terzian L, Prior J, Kim K, Pequinot E, Kahn S (1990) *One year of rHuEPO therapy prolongs RBC survival and may stabilize RBC membranes despite natural progression of chronic renal failure to uremia and need for dialysis*. ASAIO Journal 36(3):M691-M696.
157. Polenakovic M, Sikole A (1996) *Is erythropoietin a survival factor for red blood cells?* J. Am. Soc. Nephrol. 7(8):1178-1182.
158. Cotes P, Pippard M, Reid C, Winearls C, Oliver D, Royston J (1989) *Characterization of the anaemia of chronic renal failure and the mode of its correction by a preparation of human erythropoietin (r-HuEPO). An investigation of the pharmacokinetics of intravenous erythropoietin and its effects on erythrokinetics*. Q. J. Med. 70(262):113-137.
159. Pavlovic-Kentera V, Clemons G, Djukanovic L, Biljanovic-Paunovic L (1987) *Erythropoietin and anemia in chronic renal failure*. Exp. Hematol. 15(7):785-789.
160. Necheles T, Weinstein I, Leroy G (1953) *Radioactive sodium chromate for the study of survival of red blood cells: I. The effect of radioactive sodium chromate on red cells*. J. Lab. Clin. Med. 42(3):358-367.
161. Saha G (2010) *Fundamentals of nuclear pharmacy*, 6th edn. Springer Verlag, New York, USA.
162. Woo S, Krzyzanski W, Jusko W (2007) *Target-mediated pharmacokinetic and pharmacodynamic model of recombinant human erythropoietin (rHuEPO)*. J. Pharmacokinet. Pharmacodyn. 34(6):849-868.

163. Bonomi F, Stefanelli M (1980) *REDCE: A computer program for the analysis of red cell survival data*. Comput. Programs Biomed. 12(2-3):172-182.
164. Korell J, Coulter C, Duffull S (2011) *Evaluation of red blood cell labelling methods based on a statistical model for red blood cell survival*. J. Theor. Biol. 291(0):88-98.
165. Vos F, Schollum J, Coulter C, Doyle T, Duffull S, Walker R (2011) *Red Blood Cell Survival in Long-term Dialysis Patients*. Am. J. Kidney Dis. 58(4):591-598.
166. Sheiner L, Beal S (1985) *Pharmacokinetic parameter estimates from several least squares procedures: Superiority of extended least squares*. J. Pharmacokinet. Pharmacodyn. 13(2):185-201.
167. Lavielle M, Mentré F (2007) *Estimation of Population Pharmacokinetic Parameters of Saquinavir in HIV Patients with the MONOLIX Software*. J. Pharmacokinet. Pharmacodyn. 34(2):229-249.
168. Shaw A (1967) *Haemolysis in chronic renal failure*. Br. Med. J. 2:213-216.
169. Fauci A, Braunwald E, Kasper D, Hauser S, Longo D, Loscalzo J (2008) *Harrison's Principles of Internal Medicine*, 17th edn. McGraw-Hill Companies Inc., New York, USA
170. Znojil V (1983) *An ignored error in the mathematical formulation of erythrocyte survival curves*. J. Theor. Biol. 102(4):625-628.
171. Bentley S, Glass H, Lewis S, Szur L (1974) *Elution correction in ^{51}Cr red cell survival studies*. Br. J. Haematol. 26(2):179-184.
172. Taylor H, Whitley P, Heaton A (2006) *A historical perspective on platelet radiolabeling techniques*. Transfusion 46(s3):53S-58S.
173. International Committee for Standardization in Haematology (1977) *Recommended methods for radioisotope platelet survival studies: by the panel on Diagnostic Application of Radioisotopes in Hematology*. Blood 50(6):1137-1144.
174. Eyre H, Rosen P, Perry S (1970) *Relative Labeling of Leukocytes, Erythrocytes and Platelets in Human Blood by 51Chromium*. Blood 36(2):250-253.
175. Schwartz A, Kahn S, Kelch B, Kim K, Pequignot E (1992) *RBC improved survival due to recombinant human erythropoietin explains effectiveness of less frequent, low dose subcutaneous therapy*. Clin. Nephrol. 38(5):283-289.
176. Najean Y, Moynot A, Deschryver F, Zins B, Naret C, Jacquot C, Drueke T (1989) *Kinetics of erythropoiesis in dialysis patients receiving recombinant erythropoietin treatment*. Nephrol. Dial. Transplant. 4(5):350-355.
177. Ait-Oudhia S, Scherrmann J, Krzyzanski W (2010) *Simultaneous Pharmacokinetics/Pharmacodynamics Modeling of Recombinant Human Erythropoietin upon Multiple Intravenous Dosing in Rats*. J. Pharmacol. Exp. Ther. 334(3):897-910.
178. Osman A, Mulherin D (2001) *Is parenteral methotrexate worth trying?* Ann. Rheum. Dis. 60(4):432.
179. Burbage G, Gupta R, Lim K (2001) *Intramuscular methotrexate in inflammatory rheumatic disease*. Ann. Rheum. Dis. 60(12):1156.

180. Bingham S, Buch M, Lindsay S, Pollard A, White J, Emery P (2003) *Parenteral methotrexate should be given before biological therapy*. Rheumatology 42(8):1009-1010.
181. Wegrzyn J, Adeleine P, Miossec P (2004) *Better efficacy of methotrexate given by intramuscular injection than orally in patients with rheumatoid arthritis*. Ann. Rheum. Dis. 63(10):1232-1234.
182. Brooks P, Spruill W, Parish R, Birchmore D (1990) *Pharmacokinetics of methotrexate administered by intramuscular and subcutaneous injections in patients with rheumatoid arthritis*. Arthritis. Rheum. 33(1):91-94.
183. Jundt J, Browne B, Fiocco G, Steele A, Mock D (1993) *A comparison of low dose methotrexate bioavailability: oral solution, oral tablet, subcutaneous and intramuscular dosing*. J. Rheumatol. 20(11):1845-1849.
184. Braun J, Kästner P, Flaxenberg P, Währisch J, Hanke P, Demary W, Von Hinüber U, Rockwitz K, Heitz W, Pichlmeier U (2008) *Comparison of the clinical efficacy and safety of subcutaneous versus oral administration of methotrexate in patients with active rheumatoid arthritis: Results of a six-month, multicenter, randomized, double-blind, controlled, phase IV trial*. Arthritis. Rheum. 58(1):73-81.
185. Hamilton R, Kremer J (1997) *Why intramuscular methotrexate may be more efficacious than oral dosing in patients with rheumatoid arthritis*. Rheumatology 36(1):86-90.
186. Dalrymple J, Stamp L, O'donnell J, Chapman P, Zhang M, Barclay M (2008) *Pharmacokinetics of oral methotrexate in patients with rheumatoid arthritis*. Arthritis. Rheum. 58(11):3299-3308.
187. Stamp L, Barclay M, O'donnell J, Zhang M, Drake J, Frampton C, Chapman P (2011) *Effects of Changing from Oral to Subcutaneous Methotrexate on Red Blood Cell Methotrexate Polyglutamate Concentrations and Disease Activity in Patients with Rheumatoid Arthritis*. J. Rheumatol. 38(12):2540-2547
188. Sheiner L, Stanski D, Vozeh S, Miller R, Ham J (1979) *Simultaneous modeling of pharmacokinetics and pharmacodynamics: application to d-tubocurarine*. Clin. Pharmacol. Ther. 25(3):358-371.
189. Beal S (2001) *Ways to fit a PK model with some data below the quantification limit*. J. Pharmacokinet. Pharmacodyn. 28(5):481-504.
190. Stamp L, O'donnell J, Chapman P, Zhang M, Frampton C, James J, Barclay M (2009) *Determinants of red blood cell methotrexate polyglutamate concentrations in rheumatoid arthritis patients receiving long term methotrexate treatment*. Arthritis. Rheum. 60(8):2248-2256.
191. Janmahasatian S, Duffull S, Ash S, Ward L, Byrne N, Green B (2005) *Quantification of lean bodyweight*. Clin. Pharmacokinet. 44(10):1051-1065.
192. Holford N, Kirkpatrick C, Duffull S (2006) *NONMEM termination status is not an important indicator of the quality of bootstrap parameter estimates*. Population Analysis Group Europe (PAGE) 15, Abstract 992, Brugge, Belgium.
193. Savic R, Karlsson M (2009) *Importance of Shrinkage in Empirical Bayes Estimates for Diagnostics: Problems and Solutions*. AAPS J. 11(3):558-569.

194. Bergstrand M, Plan E, Kjellsson M, Karlsson M (2007) *A comparison of methods for handling of data below the limit of quantification in NONMEM VI*. Population Analysis Group Europe (PAGE) 16, Abstract 1201, Copenhagen, Denmark.
195. Ahn J, Karlsson M, Dunne A, Ludden T (2008) *Likelihood based approaches to handling data below the quantification limit using NONMEM VI*. J. Pharmacokinet. Pharmacodyn. 35(4):401-421.
196. Xu X, Dunne A, Kimko H, Nandy P, Vermeulen A (2011) *Impact of low percentage of data below the quantification limit on parameter estimates of pharmacokinetic models*. J. Pharmacokinet. Pharmacodyn. 38:423-432.
197. Schalhorn A, Sauer H, Wilmanns W, Stupp-Poutot G (1982) *Pharmacokinetics of erythrocyte methotrexate after high-dose methotrexate*. Cancer Chemother. Pharmacol. 9(1):65-69.
198. Lloyd M, Carr M, Mcelhatton P, Hall G, Hughes R (1999) *The effects of methotrexate on pregnancy, fertility and lactation*. Q. J. Med. 92(10):551-563.
199. Galivan J, Ryan T, Chave K, Rhee M, Yao R, Yin D (2000) *Glutamyl hydrolase: pharmacological role and enzymatic characterization*. Pharmacol. Ther. 85(3):207-215.
200. Dervieux T, Kremer J, Lein D, Capps R, Barham R, Meyer G, Smith K, Caldwell J, Furst D (2004) *Contribution of common polymorphisms in reduced folate carrier and γ-glutamylhydrolase to methotrexate polyglutamate levels in patients with rheumatoid arthritis*. Pharmacogenet. Genomics 14(11):733-739.
201. Faganel Kotnik B, Grabnar I, Bohanec Grabar P, Dolžan V, Jazbec J (2011) *Association of genetic polymorphism in the folate metabolic pathway with methotrexate pharmacokinetics and toxicity in childhood acute lymphoblastic leukaemia and malignant lymphoma*. Eur. J. Clin. Pharmacol. 67:993-1006.
202. Stamp L, Chapman P, O'donnell J, Zhang M, James J, Frampton C, Barclay M, Kennedy M, Roberts R (2010) *Polymorphisms within the folate pathway predict folate concentrations but are not associated with disease activity in rheumatoid arthritis patients on methotrexate*. Pharmacogenet. Genomics 20(6):367-376.
203. Cariaso M, Lennon G (2012) *SNPedia: a wiki supporting personal genome annotation, interpretation and analysis* Nucleic Acids Res. 40(Database issue):D1308-D1312.
204. Levey A, Bosch J, Lewis J, Greene T, Rogers N, Roth D, For the Modification of Diet in Renal Disease Study Group* (1999) *A More Accurate Method To Estimate Glomerular Filtration Rate from Serum Creatinine: A New Prediction Equation*. Ann. Intern. Med. 130(6):461-470.
205. Mosteller R (1987) *Simplified Calculation of Body Surface Area*. N. Engl. J. Med. 317(17):1098.
206. Holford N (1996) *A size standard for pharmacokinetics*. Clin. Pharmacokinet. 30(5):329-332.

207. Bergstrand M, Hooker A, Wallin J, Karlsson M (2011) *Prediction-corrected visual predictive checks for diagnosing nonlinear mixed-effects models*. AAPS J. 13(2):143-151.
208. Zhang L, Beal S, Sheiner L (2003) *Simultaneous vs. Sequential Analysis for Population PK/PD Data I: Best-Case Performance*. J. Pharmacokinet. Pharmacodyn. 30(6):387-404.
209. Bergstrand M, Karlsson M (2009) *Handling data below the limit of quantification in mixed effect models*. AAPS J. 11(2):371-380.
210. International Expert Committee (2009) *International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes*. Diabetes Care 32(7):1327-1334.
211. Hamren B, Bjoerk E, Sunzel M, Karlsson M (2008) *Models for plasma glucose, HbA1c, and hemoglobin interrelationships in patients with type 2 diabetes following tesaglitazar treatment*. Clin. Pharmacol. Ther. 84(2):228-235.
212. Lledo R, Mazer N, Karlsson M (2010) *A mechanistic model of the steady-state relationship between HbA1c and average glucose levels in a mixed population of healthy volunteers and diabetic subjects*. Population Approach Group in Europe (PAGE) 19, Abstract 1783, Berlin, Germany.
213. Alskär O, Korell J, Duffull S (2012) *A pharmacokinetic model for the glycation of albumin*. J. Pharmacokinet. Pharmacodyn. 39(3):273-282.
214. San George R, Nagel R, Fabry M (1984) *On the mechanism for the red-cell accumulation of mefloquine, an antimalarial drug*. Biochim. Biophys. Acta, Mol. Cell Res. 803(3):174-181.
215. Chladek J, Martinkova J, Simikova M, Vaneckova J, Koudelkova V, Nozickova M (1998) *Pharmacokinetics of low doses of methotrexate in patients with psoriasis over the early period of treatment*. Eur. J. Clin. Pharmacol. 53(6):437-444.
216. Brooks A, Begg E, Zhang M, Frampton C, Barclay M (2007) *Red blood cell methotrexate polyglutamate concentrations in inflammatory bowel disease*. Ther. Drug Monit. 29(5):619-625.
217. Van Haandel L, Leeder J, Becker M (2011) *Measurement of Methotrexathe Metabolites in Peripheral Blood Mononuclear Cells in Juvenile Idiopathic Arthritic: A More Relevant Cellular Biomarker for Drug Response?* Arthritis. Rheum. 63(10 Supplement):S95.
218. Olsen N, Murray L (1989) *Antiproliferative effects of methotrexate on peripheral blood mononuclear cells*. Arthritis. Rheum. 32(4):378-385.
219. Seitz M, Loetscher P, Dewald B, Towbin H, Rordorf C, Gallati H, Baggiolini M, Gerber N (1995) *Methotrexate action in rheumatoid arthritis: stimulation of cytokine inhibitor and inhibition of chemokine production by peripheral blood mononuclear cells*. Rheumatology 34(7):602-609.
220. Sheiner L (1997) *Learning versus confirming in clinical drug development*. Clin. Pharmacol. Ther. 61(3):275-291.
221. Mikkelsen T, Thorn C, Yang J, Ulrich C, French D, Zaza G, Dunnenberger H, Marsh S, Mcleod H, Giacomini K (2011) *PharmGKB*

summary: methotrexate pathway. Pharmacogenet. Genomics 21(10):679-686.

222. Mcdonagh E, Whirl-Carrillo M, Garten Y, Altman R, Klein T (2011) *From pharmacogenomic knowledge acquisition to clinical applications: the PharmGKB as a clinical pharmacogenomic biomarker resource*. Biomarkers 5(6):795-806.
223. Korell J, Vos F, Coulter C, Schollum J, Walker R, Duffull S (2011) *Modeling red blood cell survival data*. J. Pharmacokinet. Pharmacodyn. 38(6):787-801.